ISSN 1847-2001 (Print)
ISSN 1848-2295 (Online)

UNIVERSITY OF ZAGREB
FACULTY OF ORGANIZATION AND INFORMATICS
VARAŽDIN

Central European CONFERENCE on Information and Intelligent Systems

Proceedings

33rd International Scientific Conference

CECIIS

Organized by:
Faculty of Organization and Informatics
University of Zagreb

September 21-23/2022
Dubrovnik, Croatia

ISSN 1847-2001 (Print)
ISSN 1848-2295 (Online)

UDC 004:007

33rd International Scientific Conference

Central European Conference on Information and Intelligent Systems

CECIIS

Proceedings

September 21st - 23rd, 2022
Dubrovnik, Croatia

Organized by:
Faculty of Organization and Informatics
University of Zagreb

Publisher
Faculty of Organization and Informatics, University of Zagreb
Pavlinska 2, HR-42000 Varaždin, Croatia

Contact
E-mail: ceciis@foi.unizg.hr
Phone: +385 42 390 890
Fax: +385 42 213 413
Web: ceciis.foi.hr
Fb: facebook.com/ceciis.foi

For Publisher
Nina Begičević Ređep, Faculty of Organization and Informatics, University of Zagreb, Croatia

Editors
Neven Vrček, Faculty of Organization and Informatics, University of Zagreb, Croatia
Lourdes Guàrdia, Open University of Catalonia, Spain
Petra Grd, Faculty of Organization and Informatics, University of Zagreb, Croatia

Editorial Board
Victoria Appatova, University of Cincinnati, USA
Igor Balaban, University of Zagreb, Croatia
Nina Begičević Ređep, University of Zagreb, Croatia
Marko Bohanec, Jožef Stefan Institute, Slovenia
Leo Budin, University of Zagreb, Croatia
Tonči Carić, University of Zagreb, Croatia
Carlos Carrascosa, Universitat Politècnica de València, Spain
Ivica Crnković, Chalmers University of Technology, Sweden
Anna Czarczyńska, Kozminski University, Poland
Luis de Marcos Ortega, University of Alcalá, Spain
Blaženka Divjak, University of Zagreb, Croatia
Jasminka Dobša, University of Zagreb, Croatia
Matjaž Gams, Jožef Stefan Institute, Slovenia
Gordan Gledec, University of Zagreb, Croatia
Andrina Granić, University of Split, Croatia
Marjan Heričko, University of Maribor, Slovenia
Dieter Hertweck, Herman Hollerith Research Centre, Germany
Edouard Ivanjko, University of Zagreb, Croatia
Dimitris Karagiannis, University of Vienna, Austria
Valentina Kirinić, University of Zagreb, Croatia
Marina Klačmer Čalopa, University of Zagreb, Croatia
Michal Kohani, University of Zilina, Slovakia
Mario Konecki, University of Zagreb, Croatia
Melita Kozina, University of Zagreb, Croatia
Marjan Krašna, University of Maribor, Slovenia
Alen Lovrenčić, University of Zagreb, Croatia
Sandra Lovrenčić, University of Zagreb, Croatia
Ivan Luković, University of Belgrade, Serbia
Ivan Magdalenić, University of Zagreb, Croatia
Aleksandar Marković, University of Belgrade, Serbia
Marta Mas Machuca, UIC Barcelona, Spain
Dunja Mladenić, Jožef Stefan Institute, Slovenia
Oliver Moravčik, Slovak University of Technology, Slovakia
Enrique Mu, Carlow University, USA
Karlo Nyarko, University of Osijek, Croatia
Jan Paralič, Technical University of Košice, Slovakia
Elisabeth Pergler, University of Applied sciences, Austria
Wolf Rauch, University of Graz, Austria
Sonja Ristić, University of Novi Sad, Serbia
Markus Schatten, University of Zagreb, Croatia
Zlatko Stapić, University of Zagreb, Croatia
William Steingartner, Technical University of Košice, Slovakia
Nebojša Stojčić, University of Dubrovnik, Croatia
Vjeran Strahonja, University of Zagreb, Croatia
Diana Šimić, University of Zagreb, Croatia
Katarina Tomičić-Pupek, University of Zagreb, Croatia
Jan Tretmans, Radboud University Nijmegen, The Netherlands
Wim Van Petegem, KU Lueven, Belgium
Violeta Vidaček Hainš, University of Zagreb, Croatia
Sirje Virkus, Tallinn University, Estonia
Mladen Vouk, North Carolina State University, USA
Ksenija Vuković, University of Zagreb, Croatia

Publishing Board
Igor Tomičić, Faculty of Organization and Informatics, University of Zagreb, Croatia
Bernarda Kos, Faculty of Organization and Informatics, University of Zagreb, Croatia

All rights reserved. No parts of this publication may be reproduced, stored in a retrieval system or transmitted in any form by any means, electronic, mechanical photocopying, recording or otherwise without the prior permission of the copyright owner.
CECIIS Proceedings have been indexed in the following databases: INSPEC, ProQuest (SciTech Premium Collection, ProQuest Technology Collection and ProQuest Advanced Technologies & Aerospace database) and Web of Science.

ISSN 1847-2001 (Print)
ISSN 1848-2295 (Online)

Creative Commons Licence
Central European Conference on Information and Intelligent Systems Proceedings is licensed under a Creative Commons Attribution-NonCommercialNoDerivatives 4.0 International License.
http://creativecommons.org/licenses/by-nc-nd/4.0/

Contents

Index of Authors	VII
List of Reviewers	IX
Preface	XI

Computer Games

Markus Schatten, Tomislav Peharda, Igor Tomičić **Towards an Orchestrated Game Development Approach to Digital Twinning in Autonomous Vehicles**	3
Bogdan Okreša Đurić **Gamification in Smart and Sustainable Mobility - A Review and Roadmap**	9
Markus Schatten, Bogdan Okreša Đurić, Tomislav Peharda **Towards a Streamed Holographic 3D Game Engine**	17

Data and Knowledge Bases

Andrea Toic, Patrizia Poščić, Danijela Jaksic **Analysis of Selected Business Intelligence Data Visualization Tools**	25
Anita Ciunova-Shuleska, Nikolina Palamidovska-Sterjadovska, Violeta Cvetkoska, Irena Bogoevska-Gavrilova **Ten Years of Customer Engagement Research in Tourism and Hospitality: A Bibliometric Analysis**	33
Martina Asenbrener Katic, Sanja Candrlic, Marina Rauker Koch **Adverb and Preposition Representation in Croatian and English using the Node of Knowledge Method**	41
Jakob Jelenčič, Jože M. Rožanec, Dunja Mladenić **KL-ADWIN: enhanced concept drift detection over multiple time windows**	49
Jelena Petrovački, Dejana Gladić, Sara Havzi, Darko Stefanović, Sonja Ristić **NoSQL Databases Deployment in Various Application Domains: A Systematic Literature Review**	55
Ines Vlahović **Global Repeat Map algorithm as genomic technology for Higher Order Repeat identification**	65

Digital Response to Pandemic

Vjeran Bušelić **Teaching Information Literacy and Critical Thinking Skills in Coronavirus Pandemic Time**	85
Katarina Tomičić-Pupek, Ana Kutnjak, Vjeran Strahonja **Insights into understanding the perspective of digital platform providers**	95
Nikolina Žajdela Hrustek, Diana Šimić, Žarko Čižmar **An instrument for measuring needs of vulnerable groups in terms of digital inclusion**	107
Azra Staščik, Amanda Glavaš, Filip Glavaš **Regulation of the legal protection of individual in digital environment within the Croatian educational system**	117
Ema Novak, Ana Kutnjak, Larisa Hrustek **Opportunities for a Digital Transformation of the Global Supply Chains: Case Study of the Suez Canal Blockade**	137
Goran Bubaš, Antonela Čižmešija **Possible Changes in Academic Teaching as a Consequence of the COVID-19 Pandemic: International Findings and a Case Study from Croatia**	149

Digital Transformation and Corporate Culture

Albena Antonova, Yavor Dankov **Digital Platforms Building Blocks: Focusing on Agile Organizations and Product-Service Systems**	167
Tea Mijač, Mario Jadrić, Maja Ćukušić **Information Quality of Digital Services in Higher Education: Higher Order Construct Analysis**	175
Rudi Pleteš, Martina Tomičić Furjan, Željko Dobrović **Factors influencing project manager's ability to deliver a software engineering project**	181
Charalampos Alexopoulos, Igor Pihir, Martina Tomičić Furjan **Automatic End-to-End Decomposition and Semantic Annotation of Laws Using High-Performance-Computing and Open Data as a Potential Driver for Digital Transformation**	189
Olivia Doina Negoita, Mirona Ana-Maria Popescu, Albena Antonova, Dominik Bork **A comparison of business platforms used by SMEs to digitalize management activities**	195
Clara Viganego, Anna Akhmedova, Marta Mas-Machuca **Digital skills gap: Industry demand vs. Higher Education Institutions offerings**	205

Education and Learning Analytics

Barbi Svetec, Lea Oksanen, Blaženka Divjak, Damir Horvat
Digital Teaching in Higher Education during the Pandemic: Experiences in Four Countries 215

Francisco Ortin, Juan Ramón Pérez-Pérez, David Cabielles-Hernández, Miguel Sánchez-Santillán, MPuerto Paule-Ruiz
Mining Interaction Patterns of Children with and without Communication Disorders in the use of Tablets Apps 223

Stefan Svetsky, Oliver Moravcik
Design of All-In-One Technology-Enhanced Learning Software for Supporting Teachers' Personal Activities 233

Lourdes Guardia Ortiz, Maite Fernandez-Ferrer, Nati Cabrera, Federica Mancini, Marcelo Maina
Enhancing Self-regulation Strategies through Peer Feedback in Online Higher Education 241

Marcelo Maina, Lourdes Guàrdia, Nati Cabrera, Maite Fernandez-Ferrer, Maria Luisa Malerba, Carolina García
Evaluation of a Study Programme for the Implementation of Blended Learning at the Maldives National University 249

Melita Milić, Blaženka Divjak
Digital Maturity of Schools - Explanation, Literature Review and Analysis 255

Gordana Borotić, Gordan Gledec, Tomislav Jagušt
Exploring the Usage of an Adaptive Learning System for Elementary School Math Classes 263

Antonela Čižmešija, Goran Bubaš, Igor Balaban
A Preliminary Investigation of the Effects of Communication Quality and Use of Communication Platforms on Students' Perception of Virtual Team Performance 269

Snježana Babić, Darko Etinger, Monika Lenić
Implementing Educational Humanoid Robots in Croatian Secondary Schools: Potential Challenges and Benefits 279

Dragana Kupres, Nati Cabrera Lanzo, Albert Sangrà Morer
Digitally Mature Education: The Role of Digital Maturity and Educational Leadership in Meaningful Digital Transformation of Schools 287

ICT Entrepreneurship and Innovation

Miroslav Mandić, Iva Gregurec, Ema Galović
Employee's Perception of CRM in the Croatian Telecommunications Industry 297

Eva Malichová, Peter Márton, Emese Tokarčíková
Blended Mobility on Entrepreneurship – Intensive Programme for ICT Students in the HEI Curriculum 309

Intelligent Energy

Zvonimir Šimić, Danijel Topić, Anto Rogalo
Modeling and Simulation of AC, DC, and Hybrid AC-DC Microgrid with Battery Energy Storage 317

Intelligent Systems

Saša Brdnik, Boštjan Šumak **Adaptive and intelligent user interfaces and how they adapt to users: A systematic mapping study**	327
Ivan Magdalenić, Neven Vrček, Boris Tomaš, Karl Sölvi Guðmundsson **Localization of drone operator using Unmanned Aerial Vehicles: Preliminary research**	335
Ena Barcic, Igor Tomičić, Mario Harjač **A survey on face template protection methods**	341
Marko Horvat, Ljudevit Jelečević, Gordan Gledec **A Comparative study of YOLOv5 models performance for image localization and classification**	349
Jože M. Rožanec, Patrik Zajec, Jelle Kaizer, Elena Trajkova, Blaž Fortuna, Bor Brecelj, Beno Šircelj, Dunja Mladenić **Enhancing manual revision in manufacturing with AI-based defect hints**	357
Jože Rožanec, Tadej Krivec, Vid Keršič, Larsen Cundrič, Blaž Stojanovič, Marko Zeman, Ivan Bratko **Bicycle Sharing Systems meet AI: forecasting bicycles availability and decision-making**	365
Josip Kunsabo, Jasminka Dobša **A Systematic Literature Review on Topic Modelling and Sentiment Analysis**	371
Ljiljana Zekanović-Korona, Jurica Grzunov **Facilitating Smart Tourism Destination Development: Towards a Model of Information System for Smart Tourism**	381

Quality in ICT

Renata Mekovec, Ruben Picek, Sandra Lovrenčić, Marina Klačmer Čalopa **Creating jobs for future computer professionals**	391

Strategic Planning and Decision Making

Ivana Dvorski Lacković **Learning analytics implementation in higher education institutions: Enterprise risk management approach**	401
Maja Gligora Marković, Tena Jagačić, Nikola Kadoić **Analyzing Decision-Making Styles: Army vs. Business students in Croatia**	409
Merve Güler, Esin Mukul, Gülçin Büyüközkan **Digital Transformation Technology Evaluation in a post COVID-19 World with an Integrated Fuzzy SAW-ARAS Methodology**	417
Esin Mukul, Merve Güler, Gülçin Büyüközkan **Evaluation of Supply Chain Tools with Fuzzy AHP-COPRAS Methodology for Intelligent Supply Chain Management**	423
Jan Fauser, Dieter Hertweck **Business model development and business ecosystem evaluation in the energy domain**	431

Systems for Smart Mobility

Boris Tomaš, Lovro Posarić **WiFi access point localization in urban multi-sensor environment**	443
Filip Ribić, Tomislav Erdelić, Tonči Carić, Martina Erdelić **Solving Capacitated Time-Dependent Vehicle Routing Problem**	449
Filip Vrbanić, Edouard Ivanjko, Slaven Gašpar **Virtual Reality and Transport: A Short Review**	457

Trends in Software Engineering

Matija Novak, Iva Levak **Using graph databases in source code plagiarism detection**	465
Luis de-Marcos, José-María Gutiérrez-Martínez, Adrián Domínguez-Díaz, Sergio Caro-Álvaro, Daniel Rodríguez **Evaluating the UX of a Microservice Web App: Cyberranges in the µDevOps Project**	471
Nadica Hrgarek Lechner, Vjeran Strahonja, Zlatko Stapić **Threat Modeling Methods in the Medical Device Industry: An Integrative Literature Review**	479
Jelica Stanojević, Uroš Šošević, Miroslav Minović, Miloš Milovanović **An Overview of Modern Cross-platform Mobile Development Frameworks**	489
Žana Zekić, Zlatko Stapić **Software Quality Assessment Standards and Metrics: A Systematic Literature Review**	499
Robert Logozar, Matija Mikac, Danijel Radošević **Exploring the Access to the Static Array Elements via Indices and via Pointers — the Introductory C++ Case**	507

Author Guidelines	521
CECIIS 2022 Sponsors	525
CECIIS 2022 Financial Support	532
CECIIS 2022 Program and Organizing Committees' Members, Research, Professional Tracks' Chairs and Student Poster and Presentation Section Chair	533

Index of Authors

Akhmedova, Anna 205
Alexopoulos, Charalampos 189
Antonova, Albena 167, 195
Asenbrener Katic, Martina 41
Balaban, Igor 269
Babić, Snježana 279
Barcic, Ena 341
Bogoevska-Gavrilova, Irena 33
Bork, Dominik 195
Borotić, Gordana 263
Bratko, Ivan 365
Brdnik, Saša 327
Brecelj, Bor 357
Bubaš, Goran 149, 269
Bušelić, Vjeran 85
Büyüközkan, Gülçin 417, 423
Cabielles-Hernández, David 223
Cabrera, Nati 241, 249, 287
Candrlic, Sanja 41
Carić, Tonči 449
Caro-Álvaro, Sergio 471
Ciunova-Shuleska, Anita 33
Cundrič, Larsen 365
Cvetkoska, Violeta 33
Čižmar, Žarko 107
Čižmešija, Antonela 149, 269
Ćukušić, Maja 175
Dankov, Yavor 167
de-Marcos, Luis
Divjak, Blaženka 215, 255
Dobrović, Željko 181
Dobša, Jasminka 371
Domínguez-Díaz, Adrián 471
Dvorski Lacković, Ivana 401
Erdelić, Tomislav 449
Erdelić, Martina 449
Etinger, Darko 279
Fauser, Jan 431
Fernandez-Ferrer, Maite 241, 249
Fortuna, Blaž 357
Galović, Ema 297
García, Carolina 249
Gašpar, Slaven 457
Gladić, Dejana 55
Glavaš, Amanda 117
Glavaš, Filip 117
Gledec, Gordan 263, 349
Gligora Marković, Maja 409
Gregurec, Iva 297
Grzunov, Jurica 381
Guardia Ortiz, Lourdes 241, 249
Güler, Merve 417, 423
Gutiérrez-Martínez, José-María 471
Harjač, Mario 341
Havzi, Sara 55
Horvat, Marko 349
Horvat, Damir 215
Hrgarek Lechner, Nadica 479
Hrustek, Larisa 137
Ivanjko, Edouard 457
Jadrić, Mario 175
Jagačić, Tena 409
Jagušt, Tomislav 263
Jaksic, Danijela 25
Jelečević, Ljudevit 349
Jelenčič, Jakob 49
Kadoić, Nikola 409
Kaizer, Jelle 357
Keršič, Vid 365
Klačmer Čalopa, Marina 391
Krivec, Tadej 365
Kunsabo, Josip 371
Kupres, Dragana 287
Kutnjak, Ana 95, 137
Lenić, Monika 279
Levak, Iva 465
Logozar, Robert 507
Lovrenčić, Sandra 391
Magdalenić, Ivan 335
Maina, Marcelo 241, 249
Malerba, Maria Luisa 249
Malichová, Eva 309
Mancini, Federica 241
Mandić, Miroslav 297
Márton, Peter 309
Mas-Machuca, Marta 205
Mekovec, Renata 391
Mijač, Tea 175
Mikac, Matija 507
Milić, Melita 255
Milovanović, Miloš 489
Minović, Miroslav 489
Mladenić, Dunja 49, 357
Moravcik, Oliver 233
Mukul, Esin 417, 423
Negoita, Olivia Doina 195
Novak, Matija 465
Novak, Ema 137
Okreša Đurić, Bogdan 9, 17
Oksanen, Lea 215
Ortin, Francisco 223
Palamidovska-Sterjadovska, Nikolina 33
Paule-Ruiz, MPuerto 223
Peharda, Tomislav 3, 17
Pérez-Pérez, Juan Ramón 223
Petrovački, Jelena 55
Picek, Ruben 391
Pihir, Igor 189
Pleteš, Rudi 181
Popescu, Ana-Maria Mirona 195
Posarić, Lovro 443
Poščić, Patrizia 25
Radošević, Danijel 507
Rauker Koch, Marina 41
Ribic, Filip 449
Ristić, Sonja 55
Rodríguez, Daniel 471
Rogalo, Anto 317
Rožanec, Jože M. 49, 357, 365
Sánchez-Santillán, Miguel 223
Sangrà Morer, Albert 287
Schatten, Markus 3, 17
Sölvi Guðmundsson, Karl 335
Stanojević, Jelica 489
Stapić, Zlatko 479, 499
Staščik, Azra 117
Stefanović, Darko 55
Stojanovič, Blaž 365
Strahonja, Vjeran 95, 479
Svetec, Barbi 215
Svetsky, Stefan 233
Šimić, Zvonimir 317
Šimić, Diana 107
Šircelj, Beno 357
Šošević, Uroš 489
Šumak, Boštjan 327
Toic, Andrea 25
Tokarčíková, Emese 309
Tomaš, Boris 335, 443
Tomičić, Igor 3, 341
Tomičić Furjan, Martina 181, 189
Tomičić-Pupek, Katarina 95
Topić, Danijel 317
Trajkova, Elena 357
Viganego, Clara 205
Vlahović, Ines 65
Vrbanić, Filip 457
Vrček, Neven 335
Zajec, Patrik 357
Zekanović-Korona, Ljiljana 381
Zekić, Žana 499
Zeman, Marko 365
Žajdela Hrustek, Nikolina 107

List of Reviewers

Alexopoulos, Charalampos, Greece
Antonova, Albena, Bulgaria
Appatova, Victoria, USA
Babić, Snježana, Croatia
Balaban, Igor, Croatia
Banjanović-Mehmedović, Lejla, Bosnia and Herzegovina
Barać, Dušan, Serbia
Beliga, Slobodan, Croatia
Beranič, Tina, Slovenia
Bohanec, Marko, Slovenia
Brank, Janez, Slovenia
Cagáňová, Dagmar, Slovakia
Carić, Tonči, Croatia
Caro-Álvaro, Sergio, Spain
Carrascosa, Carlos, Spain
Chodarev, Sergej, Slovakia
Cojocar, Grigoreta Sofia, Romania
Cukusic, Maja, Croatia
Cvetokoska, Violeta, North Macedonia
Czarczyńska, Anna, Poland
Delibašić, Boris, Serbia
Dmitrovic, Veljko , Serbia
Dobrović, Željko, Croatia
Dohnal, Vlastislav, Czech Republic
Domínguez-Díaz, Adrián, Spain
Emruli, Blerim, Sweden
Erjavec, Zlatko, Croatia
Felderer, Michael, Austria
Ferencik, Norbert, Slovakia
Gazda, Juraj, Slovakia
Guardia Ortiz, Lourdes, Spain
Hajdin, Goran, Croatia
Holzinger, Andreas, Austria
Horvat, Marko, Croatia
Hruska, Domagoj, Croatia
Hunjak, Tihomir, Croatia
Ivanjko, Edouard, Croatia
Ivkovic, Nikola, Croatia
Jakšić, Danijela, Croatia
Jelenčič, Jakob, Slovenia
Julian, Vincente, Spain
Kadoić, Nikola, Croatia
Koltovska-Nechoska, Daniela, North Macedonia
Konecki, Mario, Croatia
Kovacic, Bozidar, Croatia
Krašna, Marjan, Slovenia
Kumar Cherukuri, Aswani , India
Kvassay, Miroslav, Slovakia
Lampoltshammer, Thomas, Austria
Lovrenčić, Sandra, Croatia
Luković, Ivan, Serbia
Magdalenic, Ivan, Croatia
Maina, Marcelo Fabián, Spain
Marković, Aleksandar, Serbia
Martinčić - Ipšić, Sanda, Croatia
Marton, Peter, Slovakia
Mekovec, Renata, Croatia
Melichercik, Miroslav, Slovakia
Mijač, Marko, Croatia
Milić Keresteš, Neda, Serbia
Mladenic, Dunja, Slovenia
Mu, Enrique, USA
Oreški, Dijana, Croatia
Oreški, Predrag, Croatia
Orešković, Marko, Croatia
Paralič, Jan, Slovakia
Pataki, Norbert, Hungary
Peras, Dijana, Croatia
Pergler, Elisabeth, Austria
Pihir, Igor, Croatia
Plantak Vukovac, Dijana, Croatia
Poščić, Patrizia, Croatia
Rabuzin, Kornelije, Croatia
Radaković, Davorka, Serbia
Radošević, Danijel, Croatia
Radovanović, Sandro, Serbia
Rienties, Bart, United Kingdom
Ristić, Sonja, Serbia
Rodríguez García, Daniel, Spain
Rožanec, Jože, Slovenia
Schatten, Markus, Croatia
Seljan, Sanja, Croatia
Seme, Sebastijan, Slovenia
Slavuj, Vanja, Croatia
Smaguc, Tamara, Croatia
Stapic, Zlatko, Croatia
Steingartner, William, Austria
Stremy, Maximilian, Slovakia
Suknovic, Milija, Serbia
Svetsky, Stefan, Slovakia
Šarac, Vasilija, North Macedonia
Šestak, Martina, Slovenia
Škorput, Pero, Croatia
Tomaš, Boris, Croatia
Tomičić Furjan, Martina, Croatia
Tretmans, Jan, Netherlands
Vagac, Michal, Slovakia
Van Petegem, Wim, Belgium
Varga, Michal, Slovakia
Veigend, Petr, Czech Republic
Virkus, Sirje, Estonia
Vrcek, Neven, Croatia
Zdravkova, Katerina, North Macedonia
Zlatovic, Miran, Croatia
Zornić, Nikola, Serbia
Žugec, Bojan, Croatia

Preface

It is our great honour to present to you the Proceedings of the Central European Conference on Information and Intelligent Systems 2022. We hope that you will find it useful, exciting and inspiring.

These Proceedings contain full text conference papers presented at the annual international Central European Conference on Information and Intelligent Systems (CECIIS), organized by the Faculty of Organization and Informatics Varaždin, Croatia and hosted by the University of Dubrovnik, in Dubrovnik, on September 21-23, 2022.

CECIIS has been held for 33 years with the aim to be the interface between researchers and professionals within the field of information and intelligent systems and related areas and the official language of the Conference is English.

The special topic of CECIIS 2022 Conference – "Smart Industry in post Covid-19 society", was designed to provide our participants with an opportunity to research and discuss the possibilities and challenges of new industrial and social changes brought by Covid-19 pandemic. Smart industry is a very important topic since it brings new momentum to the troubled economies that face many challenges. We are building a new generation of enterprises that will be future ready, resilient and capable of utilizing advanced technologies including robotics and artificial intelligence. In that process, interdisciplinarity is the key to success. We must creatively combine technology, management, business processes, business models, digital ecosystems, regulatory framework and many other aspects. At the same time, there is a rising fear that certain part of the population will be left behind and that all these advancements will create a deep social divide. These issues must be considered and analysed by the scientific community and that is why events like CECIIS are important for the open exchange of opinions and ideas. It is our ambition to gather critical thinkers who can contribute to the challenges of the present. We also would like to mention that the convergence of Program and Organizing Committee members, academics and experts of different industries allows us to create a cross-disciplinary approach of the conference, moving on the future topics that will be trending very soon, but also the possibility of facilitating networking between people with different areas of expertise. Bringing together thought members from 13 different countries and 30 different institutions and sectors to share their views on driving change in a new era of global connectivity gives added value to CECIIS.

The Proceedings of CECIIS 2022 include 57 scientific papers, peer-reviewed by at least two independent reviewers and accepted by the international Program Committee, within the following sections: Computer Games, Data and Knowledge Bases, Digital Response to Pandemic, Digital Transformation and Corporate Culture, Education and Learning Analytics, ICT Entrepreneurship and Innovation, Intelligent Energy, Intelligent Systems, Quality in ICT, Strategic Planning and Decision Making, Systems for Smart Mobility and Trends in Software Engineering. The papers were accepted for publication based on their interest, relevance, innovation and application to the information and intelligent systems.

The review process has been conducted with the dedication of 106 reviewers (42 reviewers from Croatia and 64 from abroad) who conducted 191 reviews which resulted with 57 accepted papers. The papers were prepared by 153 authors from 15 countries and presented during the conference. After the conference, a number of selected papers will be further reviewed with the possibility of being extended and published within the Journal of Organizational and Information Sciences (JIOS), the international scientific journal published by the Faculty of Organization and Informatics or within the European Journal of Open, Distance and E-Learning (EURODL).

The special value of this year's conference is featured by the outstanding invited keynote speakers who have left a deep mark in the scientific and professional areas of their interest. We are happy that we gathered such prominent speakers to highlight our conference. Within CECIIS 2022, the following invited keynote speakers presented their work:

- Professor Nenad Medvidović, Ph.D.: "Software Archipelagos as Impediments to Scientific Progress "- Professor Medvidović is a Director Emeritus of the Center for Systems and Software Engineering and is a Professor in the Computer Science Department at the University of Southern California. He is a faculty associate of the Institute for Software Research (ISR) at the University of California, Irvine. Medvidovic is a recipient of the National Science Foundation CAREER (2000) and ITR (2003) awards, as well as the

Okawa Foundation Research Grant (2005), IBM Real-Time Initiative Faculty Award (2007), Orange County Engineering Council's Distinguished Engineering Merit Award (2018), and ACM SIGSOFT's Distinguished Service Award (2020).

- Professor Ozan Tonguz, Ph.D.: "Pervasive use of AI and Machine Learning in Smart Transportation" - Professor Tonguz is a Full Professor in the Electrical and Computer Engineering Department of Carnegie Mellon University (CMU), Founder and CEO of Virtual Traffic Lights, LLC, and Co-Founder, CEO of Reflex AI and a pioneer in vehicular communications, V2X, wireless communications, and biologically inspired design. His work has been funded by several agencies including NSF, DARPA, US DoT, ABB, Harris RF Communications, Intel, General Motors (for 14 years), Texas Instruments, NeuroKinetics, and the Public Investment Fund (PIF) of Saudi Arabia. His current research is focused on AI and Machine Learning (ML), wireless communications, V2V and V2X communications, Computer Vision, and Autonomous Driving. Professor Tonguz has 30 US patents and is the author or co-author of more than 300 technical papers in the IEEE and ACM journals and conferences.

This year's conference featured two poster sections and gathered 12 scientific research posters and 11 student posters. The posters were prepared by 67 authors from 11 countries and presented during the conference.

The conference program also included a panel discussion "Smart University for Smart Industry" moderated by professor Blaženka Divjak, PhD (Faculty of Organization and Informatics, University of Zagreb, Croatia). The panelists were invited to share their visions of smart universities. They also discussed the position of smart universities in relation to industry in terms of new, AI-related technologies, as well as opportunities, benefits and challenges, and the contribution of smart universities to the society and industry.

We would like to thank all participants for their contributions to the Conference program and these Proceedings. Our special thanks goes to the invited speakers, authors, reviewers, session chairs, as well as Program and Organizing Committee members for their time and effort. The organizational and supporting assistance of the Faculty of Organization and Informatics was crucial for the overall success of the conference.
Finally, we are especially grateful to all our academic and industrial partners and sponsors for their support. CECIIS 2022 was organized in partnership with the EDEN Digital Learning Europe, Faculty of Transport and Traffic Sciences of the University of Zagreb, Faculty of Electrical Engineering and Computing of the University of Zagreb, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek and the University of Dubrovnik.
CECIIS 2022 was financially supported by the following sponsors: InfoDom d.o.o., Mobilisis d.o.o., Oracle d.o.o., DignetSoftware d.o.o., Factory X d.o.o., IGEA d.o.o. and the Ministry of Science and Education Republic of Croatia.

Finally, it is our wish that all the participants of CECIIS 2022 had a wonderful and fruitful time at the conference and find relevant and interesting papers in the following pages. We are extremely proud that many CECIIS participants have become our friends during all these years and that CECIIS has a circle of dedicated participants that confirm to us that every effort to make the conference better is worth it.

We are looking forward to the following Central European Conference on Information and Intelligent Systems to be held in autumn 2023.

September, 2022

On the behalf of the Program and Organizing Committees

Neven Vrček, Lourdes Guàrdia and Petra Grd
Editors

Computer Games

Markus Schatten, Tomislav Peharda, Igor Tomičić
Towards an Orchestrated Game Development Approach to Digital Twinning in Autonomous Vehicles

Bogdan Okreša Đurić
Gamification in Smart and Sustainable Mobility - A Review and Roadmap

Markus Schatten, Bogdan Okreša Đurić, Tomislav Peharda
Towards a Streamed Holographic 3D Game Engine

Towards an Orchestrated Game Development Approach to Digital Twinning in Autonomous Vehicles

Markus Schatten, Tomislav Peharda, Igor Tomičić
Faculty of Organization and Informatics, Artificial Intelligence Laboratory
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{markus.schatten,tomislav.peharda,igor.tomicic}@foi.unizg.hr

Abstract. *In this paper we provide a game development perspective on Digital twins (DTs) with emphasis on smart and autonomous vehicles. A DT which comprises a virtual representation of some systems' throughout its lifecycle and used real-time input, simulation, reasoning and machine learning to allow for decision making about the actual system is described as an agent or actor inside a simulated (game) environment. We provide a conceptual model for the development of DTs within an orchestration platform for hybrid artificial intelligence services that is being developed for complex game engines.*

Keywords. game engine, interactive fiction, massively multiplayer on-line games, role playing games, multi agent systems

1 Introduction

Digital twins (DTs) is the concept of having a virtual model that describes an observed physical system as accurately as possible. The difference between a DT and a plain simulation is in robustness of a DT in terms of variety of processes that may be analyzed, as well as DTs often working on real-world data from the simulated system. For example, a simulation may be targeting only a single parameterized use-case in order to conduct an analysis, whereas a DT serves as sort of a virtual environment that is capable to run all use-cases that a physical model may engage in (Rassõlkin et al., 2019; Wang et al., 2022). Classically defined, a digital twin is "a virtual replica of a real-world product, system, being, communities, even cities, that are continuously updated with data from its physical counterpart, as well as its environment" (Jiang et al., 2021). Compared to a "classical" model, which is defined as a simplification/abstraction of the real system, DT can be viewed as a complete mirror image model. A DT model could be developed for the purpose of any complex system. Some examples might include wind turbines, autonomous vehicles (AVs), aircraft tracking etc. Several authors view the practice of DT as the backbone of industry (Colombo et al., 2017; Jiang et al., 2021; Jones et al., 2020; Melesse et al., 2020). Although the term was coined about 20 years ago, it draw significant attention more recently, due to digital infrastructure being more embedded throughout the industry, cities, communities and everyday activities (Batty, 2018).

In this paper we will present benefits and caveats of using an orchestrated game development platform for the implementation of DTs of AVs, including the elimination of duplicated and "boilerplate" code as well as the reuse potential of the original AV system code due to the introduced modularity and proces distribution. However, such benefits might come with the price of additional computing resources.

2 Digital Twins in Autonomous Vehicles

Recently, the number of research papers on the subject of using DTs for AVs has been rising constantly. For example, in (Rassõlkin et al., 2019) it has been investigated how the DT concept can be beneficial on handling propulsion in an electrical vehicle. The idea is that a physical vehicle, which is equipped with plenty of various sensors as well as measurement systems, shares the vehicle propulsion information with the service system which is connected to the DT environment. On demand, the service system conducts simulations inside the DT environment for which it gets a feedback. The service system then uses the feedback to control the vehicle. On a similar note, (Wang et al., 2022) propose the usage of the DT concept, however, with the focus on traffic modelling. They suggest that proper traffic simulations are crucial to AVs safety and performance as different traffic dynamics will require the vehicle to apply adequate driving style depending on it. The DT concept may as well be utilized for better economics (Bhatti et al., 2021) of vehicle aspects in terms of gas consumption, driver assistance systems, battery management systems, vehicle power electronics, etc. Vehicle safety and security is another domain that requires high-level of attention (Almeaibed et al.,

2021). An example that is being brought up in the research is related to insuring consistent behaviour of a vehicle in case of a cyber attack that may harm the vehicles data integrity and consistency. Obviously, such a scenario could even lead to a traffic accident. In that sense, DTs are proposed to simulate data collection, data processing, and analytics.

It is evident that in the domain of autonomous vehicles solely, there is a wide range of aspects that require proper monitoring and forecasting, which could be supported by the usage DTs. This implies that the development of DTs is a challenging task thus requiring proper planning and system architecture. artificial intelligence (AI) and especially machine learning (ML) are greatly advocated for in this domain as they help predicting and proposing future actions based on the current circumstances. Obviously, usage of such methods adds another layer of complexity. As the components may change, it is also significant to make the DTs components extendable and modular in order to easily adjust them to new requirements. For example, with a presence of more AVs every day, that may indicate a different development paradigm shall be applied (Almeaibed et al., 2021; Bhatti et al., 2021). In the following we will try to address these problems related to the need of complexity reduction as well modularization of AV and DT architectures.

3 Awkward Π-nguin Orchestration Infrastructure

In (Schatten, Okreša Đurić, et al., 2020) we have introduced a high-level conceptual model of our microservice orchestration platform shown on figure 1. We have partially implemented[1] this platform in form of APi which is a declarative agent-based programming language based on Π calculus (Milner, 1999) (a process calculus) that allows us to model communication flows between microservices.

APi is agent based (i.e. microservices are represented as autonomous agents) and in fact holonic (Rodriguez et al., 2011) (i.e. agents can be organized hierarchically as holons). APi also features a declarative engine implemented in Python and especially Smart Python Agent Development Environment (SPADE) (Palanca et al., 2020) with parts implemented in ANTLR (syntax parser) (Rajan, 2022) and BASH shell scripts (for inter-process communication). It is programming language agnostic (i.e. microservices can be implemented in any programming language that can be executed under UNIX-friendly environments) and communication protocol agnostic (i.e. services can communicate using stdin/stdout process communication, files, HTTP, TCP, UDP or WebSockets whilst other protocols can be added through a partially implemented plug-in system) which is accomplished by creating wrappers around each microservice to be used in an ensemble. These wrappers then behave as agents inside a holonic multi agent system (HMAS) and can be used as building blocks for the creation of complex architectures. In a way, APi is a high-level network implementation of the UNIX inter-process communication system based on input/output redirection in which special programs (called filters) can be used to create complex chains of processing whereas each process reads the standard output of a previous process as its input and "filters" it producing new output to be used in a consecutive process. The same basic idea applies to APi but with the addition that these processes (in our case microservices or agents in the end effect) can be distributed among various servers and communicate across a network regardless of their implementation.

The syntax of the APi programming language is influenced by Python and Π-calculus. The following excerpt (listing 1) shows some of the main features of the language.

Listing 1: Example APi syntax

```
environment :
 input1 => { 'val1': ?x, 'val2': ?y }
 output1 <= { 'action': ?act }

channel c :
 { 'data'->?x } -> { 'payload' -> ?x }

agent a ( x ) :
 input1 -> self
 self -> ?x

agent b :
 c -> self
 self -> output1

start a( c ) & b
```

In this excerpt a holon is defined which communicates with the environment through channels `input1` and `output1`. A local channel (`c`) is defined which transforms accepts messages of the form `{ 'data'->?x }` (JSON format with logic variables) defined on the left-hand side, and emits messages of the format `{ 'payload' -> ?x }` defined on the right-hand side. Additionally two agents (microservice wrappers) are defined (`a` and `b`). Agents' definitions correspond to corresponding service-descriptors which describe how to start the service, how it communicates, what type of service it is etc., and can take arguments. In each agent definition there are one or more channel mapping definitions, for example agent `a` will direct messages from `input1` to its own input (keyword `self`) and redirect its outputs to channel held in the variable `?x` (an instantiation argument). Both agents are started consecutively (when agent `a` finishes successfully agent `b` will be started). The de-

[1] Awkward Π-nguin (APi) is open source and available at https://github.com/AILab-FOI/APi

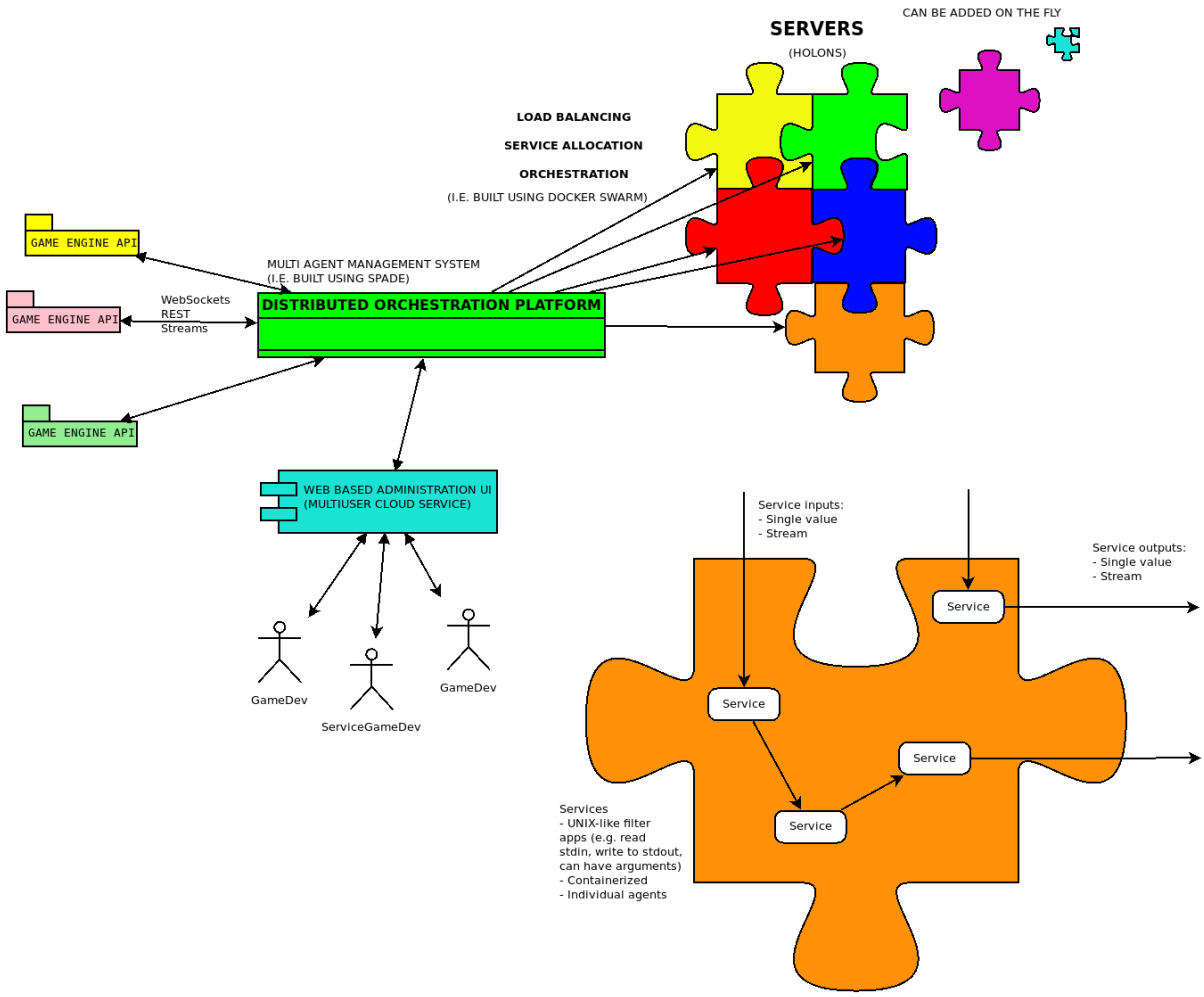


Figure 1: High-level Conceptual Model (Schatten, Okreša Đurić, et al., 2020)

fined holon can then be imported into other holons to create complex orchestration architectures. For example if the above holon had been defined in a file called `holon1.api` we could import and use its output as shown in listing 2

Listing 2: Example APi syntax

```
from holon1 import output1 as h1

agent a ( ... ) :
     h1 -> self
```

The platform is intended for the orchestration of especially AI and computer game related microservices as outlined in (Schatten, Okreša Đurić, et al., 2019) and (Schatten, Tomičić, et al., 2020). In that regard, it allows us to connect and assemble complex data streaming applications which is a prerequisite for the implementation of AV systems and consequently DTs.

4 Digital Twins as Game Actors

A common software development pattern in computer game design is the actor model (Lindley, 2002; Silva et al., 2000) which corresponds to the multiagent systems (MASs) model or more precisely to intelligent virtual environments (IVEs) (Okreša Đurić et al., 2019; Rincon et al., 2014). In such environments actors (agents) interact with their environment which consists of other (potentially intelligent and or autonomous) agents as well as dynamic or inanimate, static objects.

It has been argued that IVEs provide us with the necessary toolset to develop, test and monitor AVs architectures (Ferreira et al., 2002; Schatten, Okreša Đurić, et al., 2019; Zhang et al., 2008). Herein we argue additionally that a serious game development environment with an orchestration platform as described in section 3 can provide additional benefits. Firstly, game development environments usually feature an integrated game engine that provide basic components necessary for the development of IVEs including but not limited to 3D physics, various sensors, a simulation engine, AI methods, actor templates, cameras, logging facilities etc. Additionally, it provides us with instant visual monitoring of the environment. Secondly, an orchestration platform allows us to implement complex data streaming ensembles that might consist of various ML cloud services that are usually to resource consuming

to be executed on one computer during a simulation and/or real-time monitoring.

In order to implement a DT of an AV real-time monitoring of the vehicle in sense of sensor inputs, actuator outputs as well as internal decision making mechanisms is crucial. With an orchestration platform in which multiple cloud-based servers can be employed to handle various complex tasks like pattern recognition, automated planning or reasoning; such a workload can be distributed and handled. In the following we will show a simple proof-of-concept implementation of a DT of an AV using the APi language.

An implementation of a AV system usually consists of a number of sensors (which might include LiDAR, radar, cameras etc.), a number of decision making services (e.g. ML models, complex AI algorithms including but not limited to search and planning techniques), a number of communication services (communicating with relevant on-line services like traffic congestion and other information, routing services etc.), as well as a number of actuators (e.g. steering, car safety monitoring like tire pressure, fuel or battery life, infotainment systems etc.). All these subsystems can be wrapped into mutually communicating agents inside a holon as described in section 3.

For sake of simplicity, let us assume that the AV is a holon (i.e. a multi-agent system of various component) that communicates by reading its sensors and appropriate external services and writing commands to appropriate actuators. An implementation of such a holon in APi would look similar to the code in listing 3

Listing 3: AV holon implementation - `AVHolon.api`

```
environment :
     sensor_1 => ...
     ...
     sensor_n => ...
     external_service_1 => ...
     ...
     external_service_k => ...
     actuator_1 <= ...
     ...
     actuator_m <= ...

agent AV :
     sensor_1 -> self
     ...
     sensor_n -> self
     external_service_1 -> self
     ...
     external_service_k -> self
     self -> actuator_1
     ...
     self -> actuator_m
```

It is important to note that by defining the AV system in this way we abstract away all internal mechanisms of communication, meaning that we now have standard communication interfaces (channels) for all included components and that these interfaces can be reused, i.e. we can redirect them at will to other agents or holons. This possibility of redirection is a special feature of the declarative engine of APi - it allows us to listen to or write to any channel in the given scope (i.e. the current holon which in this case acts as a namespace). This feature is most important for the implementation of a DT since it allows us to "tap into" communication flows and process real time data.

In order to have benefits from a DT we should additionally add a corrective input, i.e. an additional communication channel to allow for intervention if the AV system does not function as intended. The environment would then look as shown in listing 4.

Listing 4: Adding a corrective to the AV holon

```
environment :
     corrective => ...
     sensor_1 => ...
     ...
...
```

Having the prerequisites defined, the implementation of a DT using APi is now straightforward. We firstly have to import all communication channels and redirect them to a new agent which will represent the DT. In the end we start the DT in parallel with the AV holon defined above as shown in listing 5.[2]

Listing 5: Digital twin implementation

```
from holonAV import sensor_1 as s1
...
from holonAV import sensor_n as sn

from holonAV import external_service_1 as
    ↪ e1
...
from holonAV import external_service_k as
    ↪ ek

from holonAV import actuator_1 as a1
...
from holonAV import actuator_m as am

from holonAV import corrective as cr

agent AVtwin :
     s1 -> self
     ...
     sn -> self
     e1 -> self
     ...
     ek -> self
     a1 -> self
     ...
     am -> self
     self -> cr
```

[2]The | character between two agents denotes parallel execution.

```
start holonAV | AVtwin
```

In this way the DT can monitor all inputs and all outputs of the original AV system and use this data to update its IVE, simulate and analyze its decision making process as well as send corrective information if needed.

5 Discussion

The main benefit of an orchestrated implementation as shown in section 4 is the modular approach that minimizes code duplication especially in "boilerplate" communication code. In a non-orchestrated implementation each service instance from the original AV system which communicates in any way (which is in fact most instances) would have to be either rewritten or at least extended to forward communicated data to the DT. Also, for the implementation of the DT each component simulation of the original would have to be implemented anew in an IVE, game engine or other simulation environment using native development tools and languages. By using agent wrappers around these components most (not hardware specific) code could be reused and work in a more or less unchanged way as it would in the original setting.

Still, there are caveats of course. By wrapping the original AV system components as agents additional overhead code is introduced which might use additional computing and resources in order to function and communicate with the orchestration platform. Such additional resources might not be available or it might be not desirable to use them in real-time and mission critical settings. Additionally, a problem that isn't solved using the given approach is the need to implement the update process of the IVE by using real-time data from the original AV system. These update components need to be implemented as additional services which will update the simulated environment based on sensory data and integrated with the other (possible reused) services. On the other hand, the integration process is fairly trivial, due to the modularity of APi, i.e. they can be introduces without intervention into the code of the original services, which might not be the case in a non-orchestrated setting.

6 Conclusion

In this paper we have argued the benefits and caveats of using an orchestrated game development platform for the implementation of DTs of AVs. By using the APi platform that we are actively developing we have shown how to implement a basic DT ensemble. The most important advantages of using an orchestrated game development platform, besides the various building blocks provided by game development software (especially game engines), include the elimination of duplicated and "boilerplate" code as well as the reuse potential of the original AV system code due to the introduced modularity and proces distribution. These benefits come with an expense of introducing the need of additional computing resources on the AV system side.

Our future research is aimed towards the implementation of a proof-of-concept digital twinning environment for AVs in which we shall address some of the issues outlined above.

Acknowledgement

This work has been fully supported by the Croatian Science Foundation under the project number IP-2019-04-5824.

References

Almeaibed, S., Al-Rubaye, S., Tsourdos, A., & Avdelidis, N. P. (2021). Digital twin analysis to promote safety and security in autonomous vehicles. *IEEE Communications Standards Magazine*, *5*(1), 40–46.

Batty, M. (2018). Digital twins.

Bhatti, G., Mohan, H., & Singh, R. R. (2021). Towards the future of smart electric vehicles: Digital twin technology. *Renewable and Sustainable Energy Reviews*, *141*, 110801.

Colombo, A. W., Karnouskos, S., Kaynak, O., Shi, Y., & Yin, S. (2017). Industrial cyberphysical systems: A backbone of the fourth industrial revolution. *IEEE Industrial Electronics Magazine*, *11*(1), 6–16.

Ferreira, F. P., Gelatti, G., & Musse, S. R. (2002). Intelligent virtual environment and camera control in behavioural simulation. *Proceedings. XV Brazilian Symposium on Computer Graphics and Image Processing*, 365–372.

Jiang, Y., Yin, S., Li, K., Luo, H., & Kaynak, O. (2021). Industrial applications of digital twins. *Philosophical Transactions of the Royal Society A*, *379*(2207), 20200360.

Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterising the digital twin: A systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, *29*, 36–52.

Lindley, C. A. (2002). The gameplay gestalt, narrative, and interactive storytelling. *CGDC Conf.*

Melesse, T. Y., Di Pasquale, V., & Riemma, S. (2020). Digital twin models in industrial operations: A systematic literature review. *Procedia Manufacturing*, *42*, 267–272.

Milner, R. (1999). *Communicating and mobile systems: The pi calculus*. Cambridge university press.

Okreša Đurić, B., Rincon, J., Carrascosa, C., Schatten, M., & Julian, V. (2019). Mambo5: A new ontology approach for modelling and managing intelligent virtual environments based on multi-agent systems. *Journal of Ambient Intelligence and Humanized Computing*, *10*(9), 3629–3641.

Palanca, J., Terrasa, A., Julian, V., & Carrascosa, C. (2020). Spade 3: Supporting the new generation of multi-agent systems. *IEEE Access*, *8*, 182537–182549.

Rajan, H. (2022). Antlr: A brief review.

Rassõlkin, A., Vaimann, T., Kallaste, A., & Kuts, V. (2019). Digital twin for propulsion drive of autonomous electric vehicle. *2019 IEEE 60th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON)*, 1–4.

Rincon, J., Garcia, E., Julian, V., & Carrascosa, C. (2014). Developing adaptive agents situated in intelligent virtual environments. *International conference on hybrid artificial intelligence systems*, 98–109.

Rodriguez, S., Hilaire, V., Gaud, N., Galland, S., & Koukam, A. (2011). Holonic multi-agent systems. In *Self-organising software* (pp. 251–279). Springer.

Schatten, M., Okreša Đurić, B., & Tomičić, I. (2019). Towards simulation of ambient intelligence in autonomous vehicles using car racing games. *Central European Conference on Information and Intelligent Systems*, 3–6.

Schatten, M., Okreša Đurić, B., & Tomičić, I. (2020). Orchestration platforms for hybrid artificial intelligence in computer games-a conceptual model. *2020 Central European Conference on Information and Intelligent Systems*, 3–8.

Schatten, M., Tomičić, I., & Okreša Đurić, B. (2020). Towards application programming interfaces for cloud services orchestration platforms in computer games. *2020 Central European Conference on Information and Intelligent Systems*, 9–14.

Silva, D. R., Siebra, C. A., Valadares, J. L., Almeida, A. L., Frery, A. C., da Rocha Falcão, J., & Ramalho, G. L. (2000). A synthetic actor model for long-term computer games. *Virtual Reality*, *5*(2), 107–116.

Wang, S.-H., Tu, C.-H., & Juang, J.-C. (2022). Automatic traffic modelling for creating digital twins to facilitate autonomous vehicle development. *Connection Science*, *34*(1), 1018–1037.

Zhang, T., Liu, X., Mei, T., Tang, G., Li, B., & Wang, X. (2008). A novel platform for simulation and evaluation of intelligent behavior of driverless vehicle. *2008 IEEE International Conference on Vehicular Electronics and Safety*, 237–240.

Gamification in Smart and Sustainable Mobility - A Review and Roadmap

Bogdan Okreša Đurić
Faculty of Organization and Informatics, Artificial Intelligence Laboratory
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
dokresa@foi.unizg.hr

Abstract. *Gamification represents the introduction of computer game related concepts into other, not game related processes in order to foster motivation of participants. There have been numerous studies which have shown that gamification can and should be applied to various aspects of mobility and have provided promising results in encouraging smart and sustainable forms of transportation. Herein, we provide an overview of these studies. Nevertheless, only particular aspects of gamification have been applied so far and there are still many more methods that gamification has to offer. Thus, we provide a roadmap of possible application areas in future developments.*

Keywords. gamification, sustainable mobility, smart mobility, smart city

Sustainable mobility is one of the key aspects of smart cities, sustainable development, and sustainable living (Lombardi and Giordano, 2015; Salat et al., 2011). Moreover, developing a society that is engaged and invested in sustainable development, a part of which is sustainable mobility, is a process that requires willingness, commitment, and investment from all the stakeholders. The most numerous out of those are citizens, i.e. the general public. In order to change the developed behaviour patterns of the general public, the change must be motivated, either by the demand, the supply, or by some other factor.

Sustainable mobility options have existed for decades (e.g. bicycles), and some of the historically big pollutants are being developed into sustainable units (e.g. electric vehicles). The comfort and convenience of using personal vehicles is one of the significant obstacles to the public shift to some more sustainable and environment-friendly means of transportation, especially in large cities. Therefore, many attempts have been undertaken in order to incentivise and motivate the shift of the public behaviour patterns in favour of sustainable mobility options.

Gamification (Deterding et al., 2011) is a wonderful approach to encouraging engagement, motivation, and feeling of success of participants in various types of activities. Although the set of game-inspired methods is rich and varied, the existing and past gamified approaches to encouraging smart and sustainable forms of transportation use only a small subset of those. Further planning and research is needed, in order to identify the most useful gamification methods in the domain of forming and altering public opinion. Apart from the latter, gamification can be used in many more scenarios in the context of smart and sustainable mobility, as is presented in the following sections.

This paper is structured as follows. Section 2 is gives a short definition and introduction to gamification. Section 3 provides a review of a select set of examples and research on various implementation of gamification in the context of smart and sustainable mobility. Section 4 consists of a short discussion related to the contents of Section 3, followed by Section 5: a roadmap of possible application areas of gamification on sustainable mobility.

1 Gamification

Gamification can, in general terms, be described as a process of applying game-bound elements, design techniques and mechanics to non-game context (Deterding et al., 2011; Kalmpourtzis, 2018; Matallaoui et al., 2017). One of the more prominent goals of using gamification, i.e. gamifying a certain non-game context, is motivating users to "achieve certain behavioural or psychological outcomes." (Matallaoui et al., 2017) There is a multitude of domains whereupon gamification is applicable, including, but not limited to, education, sport, medicine, etc.

Although the first documented use of the concept of gamification is observed in 2008 (Deterding et al., 2011), or possibly as early as 2002, the idea of applying game-related concepts to non-game context has evolved various forms, spanning from simple implementation of leader boards to advanced role-playing and serious games. Yet, even though the science of gamification sparked interest in both academia and industry, it is still in early stages of development and research in many of its application domains (Schatten et

al., 2021; Smith et al., 2021).

Game mechanisms used most often in gamifying a select process or experience include score boards, badges, various kinds of points, levels, progress bars, virtual currency, and player avatars (Dicheva et al., 2015; Smith et al., 2021). Most of these concepts can be grouped together and classified as reward mechanisms, but the others are vital for the full gamified experience as well. Additionally, three categories of game elements were identified in (Werbach and Hunter, 2012) as relevant to gamification: dynamics, mechanics, and components, wherein the category of dynamics comprises the most abstract game elements, while the category of components encompasses the most specific elements that are visible to the users. The main goal of gamification through such game mechanics and related concepts is increased user engagement, higher user motivation, and increased interaction frequency with the gamified experience.

Even though gamification is concentrated on using game elements in application domains which lack game elements by default, the recurring result is a designed game-like context that utilises select ideas from both the game and the non-game domains. In other words, it is often found that domain-specific games are being developed in order to foster how game elements are used towards achieving a certain goal, e.g. a location-based game is developed in order to motivate people to move more. Therefore, a set of categories of games emerges, based on their set of functions and expected types of application, although most of these depend on mobile devices. The most interesting of those, considering the scope of this paper, and the contents of Section 3, are location-based mobile games (Frith, 2013), and activity-based mobile games (such as fitness and step-tracking applications).

2 Literature Review

The following is an overview of use cases and application domains of gamification in the context of mobility, encompassing personal transportation, public transportation, and mobility as a concept describing how people behave while moving through the physical space. The second part of this section gives an overview of use of gamification in the domain of personal vehicles and related contexts.

2.1 Gamification and Mobility

Considering the popularity of wearable devices, such as smart bands and smart watches, one of the most used ways, and arguably the easiest, of introducing gamification to the general audience in the domain of mobility, according to the amount of published research, are fitness applications and step trackers, i.e. applications that count the number of steps performed in a calendar day (Bucher et al., 2016; S. R. Greysen et al., 2021; Nakashima et al., 2017). Such applications feature, in the very least, various kinds of reward mechanisms, for example:

- leader boards that include users on a local, national and regional level,
- badges that users are awarded with for achieving specific challenges, such as a predefined number of steps, consecutive days of satisfying activity, or other specific achievements,
- scores based on the steps walked within a day, possibly enriched with other physical activities.

Two studies are selected to be presented here in favour of the above application domain of gamification in the context of mobility. The first (Nakashima et al., 2017) is a smartphone application designed to foster and promote Mobility Management, a set of policies and strategies with the goal of promoting a modal shift towards the more sustainable means of transportation. The developed application motivated users to walk more, counted their steps and spent calories, and even featured images for each of the specific four categories of number of steps, so as to encourage its users to walk. In the end, the conducted survey showed that there are, on average, no benefits of using the developed application, or the methods contained within. However, a statistically significant increase in the number of steps was identified in users who had a below-average step count before using the application.

A similar result was presented by (S. R. Greysen et al., 2021). The system presented in this paper was developed with the intention of increasing mobility after hospital discharge, i.e. it counted the number of steps performed by former patients. One of the elements to foster motivation here is social engagement, i.e. the developed system is connected to social media, and a specific close contact was selected for each user – this close contact was to be used as a personal motivator for each user. Furthermore, the system allowed for interaction with the user's social circle using interfaces towards various social network application. The results of this research are similar to the one described above – users who used the system, along with its full gamified capabilities, have not performed better, i.e. made more steps within a calendar day, than the users in the control group. However, the difference was statistically significant when a subset of participants was observed – in particular those with higher levels of social engagement prior to participating in the study.

An interesting use case of gamification in public transportation domain presents itself in the context of motivating users to use the available means of transportation in specific periods in time (Cardoso et al., 2019). An example of the former is promoting use of public transportation out of rush hour by providing greater rewards when users use the available public infrastructure in the goal time slots, or providing incen-

tive to users who use a city-wide bicycle network instead of driving to work in the morning or using other means of public transportation if they live within a certain distance of their work destination.

Mobility, i.e. moving in traffic, can itself be observed as a game, although such a game might be described as serious and pervasive, meaning that it is not created with the goal of having fun, and it expands out of the social, temporal, and physical boundaries of regular games (McCall et al., 2013; Montola, 2005). One such system described in the paper from 2013 featured instant rewards, status rewards, and basic collaboration and competition mechanisms, such as grouping and inter-group competing. The results of the research presented in (McCall et al., 2013) indicate that the monetary value of incentive that can induce change in behaviour can be relatively small. Another approach is identified by Frith in the location-based social network application Foursquare (Frith, 2013). Frith argues that "[...] earning the badge is often enough reward in itself to encourage behaviors." Furthermore, a gamified approach that overlays virtual context over the real world, such as is provided by Foursquare, is argued to be influential in building a person's image, affect consumption, and alter the way game participants observed and experienced their non-virtual surroundings.

Providing users of gamified applications with incentive to change their mobility behaviour and sway it in the more sustainable, or *green* as it is often called, direction is a recurring topic in research on gamification and its application in the domain of mobility, e.g. (Broll et al., 2012; Cellina et al., 2019; Di Dio et al., 2019; Ferron et al., 2019; Froehlich et al., 2009; Kazhamiakin, Loria, et al., 2021; Khoshkangini et al., 2021; Lee et al., 2013; Marconi et al., 2018; Zinke-Wehlmann and Friedrich, 2019). This observation grows in importance when put into the context of smart cities, since sustainable urban mobility is one of the key issues for sustainability (Lombardi and Giordano, 2015; Salat et al., 2011). One solution towards sustainable urban mobility is the gamification framework presented in (Kazhamiakin, Marconi, et al., 2015), which explores gamification mechanisms that can be utilised in incentivising change of commuters' behaviour in the direction of sustainable mobility solutions. This gamification framework is further developed into a gamification platform described in (Kazhamiakin, Loria, et al., 2021), wherein the authors identified three application scenarios of sustainable mobility: home–school, commuting, and large-scale mobility, i.e. regular citizens' use of public transportation and their movement throughout the city. A similar approach, in the context of motivating users to utilise sustainable mobility solutions, is presented in (Zinke-Wehlmann and Friedrich, 2019). Another example is presented in (Di Dio et al., 2019) as a mobile activity-based game that engages three stakeholders in the domain of building and encouraging sustainable mobility: citizens, local businesses, and public authorities. A successful approach to motivating player behaviour through challenges is presented in (Khoshkangini et al., 2021), where authors describe a devised solution for automatically generating personalised challenges for individual players, based on their preferences, history, and past performance. Users who participated in using Play&Go game described in (Ferron et al., 2019), a "long-running open-field urban mobility game promoting voluntary travel behaviour change," demonstrated perceived change of behaviour in favour of more sustainable mobility patterns.

2.2 Gamification and Personal Vehicles

In the automotive domain, gamification approaches are found to be applied within at least six use cases - eco-driving ((Günther et al., 2020; Helvaci et al., 2018; Inbar et al., 2011; Magaña and Organero, 2014)), driving safety ((Helvaci et al., 2018; Shi et al., 2012; Steinberger et al., 2015; Wang et al., 2014)), marketing (Tillström, 2012), vehicle user interfaces training (Diewald et al., 2015), increasing social awareness (Wang et al., 2014), and increasing trust in autonomous driving (Häuslschmid et al., 2017). Eco-driving and driving safety scenarios could be integrated in the vehicles' infotainment systems, and as such could award points and badges for safe and ecological driving (Diewald et al., 2015).

Reducing average energy consumption and enhancing the driving range of battery electric vehicles (BEV) is discussed in (Günther et al., 2020). Authors included the influence of a gamification approach as one of the three persuasive strategies on the average BEV energy consumption used over 22 months on 108 participants on the field study in Germany. The results showed that the use of game design elements and financial rewards "significantly reduced energy consumption as compared to baseline or mere feedback regarding energy consumption".

A system that tries to identify driving style mistakes and rate the driving with a score is presented in (Helvaci et al., 2018). There are three components to the system, namely data capture from the vehicle CAN bus, mistake identification with multiple criteria, and score presentation in real time. With the usage of gaming elements such as achievements and leader boards, the driver is motivated to sustain a positive score while driving.

Similarly, (Magaña and Organero, 2014) analyses and validates the impact of using gamification techniques for improving eco-driving learning. Authors use game mechanisms such as the score and achievements systems to motivate driver for an efficient driving. The system warns the driver each time it detects "an inefficient action of the driver to a previously known situation such as a bad reaction to a detected traffic sign or a detected traffic accident". Authors have validated

their proposal on 14 different drivers on more than 300 rides with 5 different models of vehicles on 4 different regions of Spain and showed "a positive correlation in the use of gamification techniques and the application of the proposed of eco-driving tips, especially for aggressive drivers". What is interesting, the implemented techniques showed potential to prevent drivers from regressing to their previous driving habits. They used OBD2 diagnostic port in order to obtain the vehicle telemetry and the knowledge base containing rules on the driving efficiency.

Some cars are suffering from user experience problems. In a slightly different usage scenario, a paper is tackling gamification application on exploration and practicing on user interfaces and vehicle functions (Diewald et al., 2015). If a driver is using the car with interfaces and functions that are unfamiliar to her, it could quickly became a challenge to navigate through those interfaces as well as become a safety concern. Authors have thus created "a gamified automotive exploration and practicing framework" which would allow exploring vehicle functions and user interfaces as found in the real vehicles - the framework consists of a mobile application that "recreates the vehicle cockpit for allowing offline exploration and training, and an in-vehicle application that replaces the owner's manual and provides hints and tips for the driver".

Increasing trust in autonomous driving was tackled by (Häuslschmid et al., 2017). Authors have developed 3 different visualizations for a user study, which were overlaid to a driving scene - a chauffeur avatar, a world in miniature, and a display of the car's indicators as the baseline. Authors argued that trust "can be increased by means of a driver interface that visualizes the car's interpretation of the current situation and its corresponding actions." The results of the study showed that the world in miniature visualization increased trust the most, but they did not find a significant difference between the chauffeur and the baseline - although there might be a potential in this context.

Authors in (Steinberger et al., 2015) propose a combination of video game design theory and road safety psychology in order to engage safer driving, with 3D AR objects displayed on the windshield of the car.

Expressing the appreciation for other drivers' polite driving behavior is described in the concept presented in (Wang et al., 2014). Social awareness is the foundation for improving the driving behaviour - there is a possibility in this context to give a positive feedback about the driving behavior of other drivers. In a way, a driver thus collects "positive points" which encourages her for improving her driving behaviours.

3 Discussion

The introductory use case example in the last section is about fitness and step trackers. Even though the idea of gamification applied to the most basic mode of transportation and the most basic form of mobility of people seems like an interesting and possibly highly successful idea, the results vary depending on a number of other factors. Based on the research described in the previous section (Cardoso et al., 2019; S. R. Greysen et al., 2021), gamification alone is not enough to motivate people to move more, on average. Specific groups benefit from the added motivation and engagement mechanism they are exposed to and interact with, but the success of those mechanisms depends on the type of person, in either the physical, mental, or social context. The set of recommendations presented in (Cellina et al., 2019) aims at improving the effectiveness of persuasive applications by providing the necessary information just in time, providing goal setting opportunities, social engagement of players, and improved basic gamification concepts, e.g. feedback, and rewards.

The most prominent goal of using gamification in the domain of public transportation is geared towards promoting sustainable transportation options and available public transportation segments that are aligned with sustainability goals. The results, as shown in the previous section, are no all positive, although a positive trend can be observed, i.e. applied gamification methods have a positive impact on changing the public's behaviour towards sustainable. In the context of automotive industry, the list of domains of applied gamification methods features use cases from a broader set of domains. This discrepancy might be argued to have its source in relatively larger population that can take part in using the applications open to use to the whole population, regardless of their means of transportation, as opposed to the automotive industry being focused on the people who can drive.

Although the set of gamification methods is a rich one, the examples described in the previous section usually work with a limited subset of these methods. The most used methods are scoring, leaderboards, and badges. A small number of papers expanded the set of methods to interaction with other players, and cooperation or collaboration within groups of players, or between such groups. Some of the examples utilise the concept of quests, social interaction, and tournaments on various levels, with the goal of directing player behaviour. Some of the examples take a more graphic approach, and use the powers of narrative or graphical immersion, in order to convey the intended message. It should be noted here that it was decided that the previous section would not include applications that are not directly related to mobility, i.e. they might have change in mobility behaviour as a side-product, such as the mobile game PokemonGO (Althoff et al., 2016).

4 Roadmap

By analyzing existing body of research, as discussed in more detail within Literature Review section, we

have identified six application domains for gamification in automotive context: eco-driving, driving safety, marketing, vehicle user interfaces training, increasing social awareness, and increasing trust in autonomous driving. As we have not considered gamification in marketing as a critical use-case that can substantially contribute to the field of smart and sustainable mobility, we exclude it from our proposed roadmap.

Application domains from other smart mobility contexts include increasing overall health, promoting sustainable means of transportation, and targeting non-rush hours as means of a transport time optimization. Putting it all together, we propose a roadmap with hierarchically positioned elements for the further development of the field, with driving safety being the most critical application domain in this context.

Mastering the vehicle user interfaces can contribute to the overall driving safety, as the driver can focus on the driving activity without much interruptions from searching needed functions.

As the driving safety increases, gamification could help drivers with the social awareness and more respectful driving, showing appreciation and understanding for other fellow drivers.

On the other hand, gamification should promote more sustainable means of transportation - cycling, walking, public transportation, which leads to the research concerned with increasing the overall health of the individual, but also potentially, the health of the planet.

Gamification methods could reward individuals for taking the non-rush hours, contributing to the more effective use of smart and sustainable transportation, but also in some aspects to the driving safety.

As we can expect autonomous driving to become a reliable method of transportation in the near future, motivating people to use it will become a challenge – and gamification could tackle this issue.

Speaking of specific gamification methods that we think might be useful in the future of applying gamification to smart mobility, especially in the context of promoting sustainable mobility, the following should be noted:

- **Rewarding mechanisms** in cooperation with local businesses, wherein users might be able to be awarded certain real-world awards sourced in the local business community, especially by the local handicraft shops or various craftsmen, based on their achievements in the virtual world (solving quests, performing specific kinds of behaviour, and similar), measured through a developed mobility application.

- **Local currency** might be used a means of trading between the users who accrue currency by being active in the real world (quests, and achievements), thus creating another bond between the real world and its digital twin (albeit created on a highly abstract level).

- **Titles and appearance** are already shown to be useful and successful in motivating users (who might be called players at this level) to perform specific tasks within a game. In the domain of this paper, the titles feature should be expanded, when compared to that of Foursquare's (described in (Frith, 2013)), for example by awarding titles based on a local (neighbourhood) level, city or municipality level, etc. Some other key title elements might be found in awarding various titles based on players' achievements, recognised types of behaviour, most used means of mobility, etc. In addition to titles, another effective way of motivating users to take part in the activities provided by the game, i.e. the application promoting either sustainable mobility, driving safety, or eco-driving, are visual avatars – user's representation in the game world. This avatar might be completely virtual, or it might be based on the player's real-world experience. The beauty of this approach is that the avatars can be designed in a way that they can be customised, e.g. based on the achievements of their players, their performance, or successfully completed challenges. This way, players can clearly and easily showcase their success visually, which is one of the most popular forms of achieving prestige in massively multiplayer online games in general.

5 Conclusion

There are many application domains of gamification in the context of mobility, especially smart and sustainable mobility, even though the domain of gamification can still be considered as rather new. The currently available applications, or use cases that are available in the real world, use a set of gamification methods that encompasses the most-used methods, and this paper provides an opinion, and a suggestion, on the gamification methods that might be useful in further motivating users of mobility applications, towards achieving the application developers' planned goals.

It has been shown in a substantial body of research that gamification in general can be used to enhance motivation and incentivise a specific kind of behaviour of players. Applications in the domain of promoting e.g. sustainable mobility, driving safety, eco-driving, etc. are shown to benefit from use of gamification methods as well, even though the benefits are not always visible immediately. Naturally, in order to perceive the change in mobility, a specific amount of time must pass, and the change visible in an application alone is not enough.

Although gamification methods have varying effects on the ultimate player motivation and their engagement, their careful combination is what an application or a game benefits from, especially when combined with cultural, behavioural, and social features of its users. Thus, sustainable mobility, and the other domains mentioned in this paper, is a domain worth ex-

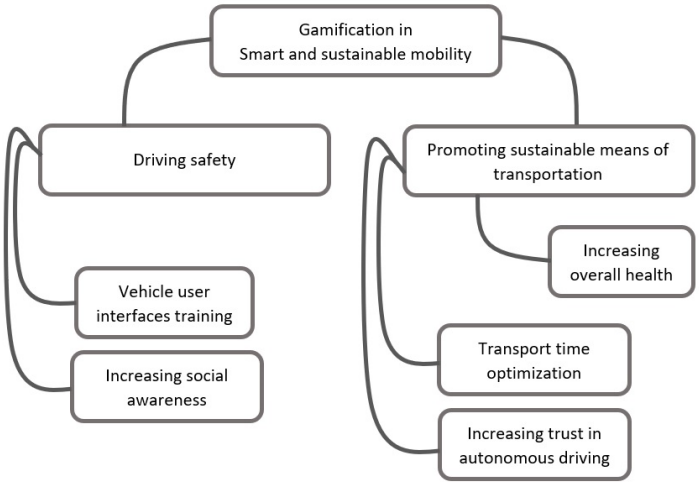


Figure 1: A proposed roadmap for gamification in smart and sustainable mobility (Helvaci et al., 2018)

ploring and enhancing with further and more rich and engaging gamification methods.

Acknowledgement

This work has been fully supported by the Croatian Science Foundation under the project number IP-2019-04-5824.

References

Althoff, T., White, R. W., & Horvitz, E. (2016). Influence of Pokémon Go on Physical Activity: Study and Implications. *Journal of Medical Internet Research*, *18*(12), e315. https://doi.org/10.2196/jmir.6759

Broll, G., Cao, H., Ebben, P., Holleis, P., Jacobs, K., Koolwaaij, J., Luther, M., & Souville, B. (2012). Tripzoom: An App to Improve Your Mobility Behavior. *Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia*, 1–4. https://doi.org/10.1145/2406367.2406436

Bucher, D., Cellina, F., Mangili, F., Raubal, M., Rudel, R., Rizzoli, A. E., & Elabed, O. (2016). Exploiting Fitness Apps for Sustainable Mobility - Challenges Deploying the GoEco! App. *Proceedings of ICT for Sustainability 2016*. https://doi.org/10.2991/ict4s-16.2016.11

Cardoso, B., Ribeiro, M., Prandi, C., & Nunes, N. (2019). When gamification meets sustainability: A pervasive approach to foster sustainable mobility in madeira. *Proceedings of the 1st ACM Workshop on Emerging Smart Technologies and Infrastructures for Smart Mobility and Sustainability*, 3–8.

Cellina, F., Bucher, D., Veiga Simão, J., Rudel, R., & Raubal, M. (2019). Beyond Limitations of Current Behaviour Change Apps for Sustainable Mobility: Insights from a User-Centered Design and Evaluation Process. *Sustainability*, *11*(8), 2281. https://doi.org/10.3390/su11082281

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining "Gamification". *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments*, 9–15. https://doi.org/10.1145/2181037.2181040

Di Dio, S., Lissandrello, E., Schillaci, D., Caroleo, B., Vesco, A., & D'Hespeel, I. (2019). MUV: A Game to Encourage Sustainable Mobility Habits. In M. Gentile, M. Allegra, & H. Söbke (Eds.), *Games and Learning Alliance* (pp. 60–70). Springer International Publishing. https://doi.org/10.1007/978-3-030-11548-7_6

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in Education: A Systematic Mapping Study. *Journal of Educational Technology & Society*, *18*(3), 75–88.

Diewald, S., Möller, A., Stockinger, T., Roalter, L., Koelle, M., Lindemann, P., & Kranz, M. (2015). Gamification-supported exploration and practicing for automotive user interfaces and vehicle functions. In *Gamification in education and business* (pp. 637–661). Springer.

Ferron, M., Loria, E., Marconi, A., & Massa, P. (2019). Play&Go, an Urban Game Promoting Behaviour Change for Sustainable Mobility., 22.

Frith, J. (2013). Turning life into a game: Foursquare, gamification, and personal mobility. *Mobile Media & Communication*, *1*(2), 248–262.

Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., & Landay, J. A. (2009). UbiGreen: Investigating a Mobile Tool for Tracking and Supporting Green Transportation Habits. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1043–1052. https://doi.org/10.1145/1518701.1518861

Greysen, S. R., Changolkar, S., Small, D. S., Reale, C., Rareshide, C. A., Mercede, A., Snider, C. K., Greysen, H. M., Trotta, R., Halpern, S. D., et al. (2021). Effect of behaviorally designed gamification with a social support partner to increase mobility after hospital discharge: A randomized clinical trial. *JAMA network open*, *4*(3), e210952–e210952.

Günther, M., Kacperski, C., & Krems, J. F. (2020). Can electric vehicle drivers be persuaded to eco-drive? a field study of feedback, gamification and financial rewards in germany. *Energy Research & Social Science*, *63*, 101407.

Häuslschmid, R., von Buelow, M., Pfleging, B., & Butz, A. (2017). Supportingtrust in autonomous driving. *Proceedings of the 22nd international conference on intelligent user interfaces*, 319–329.

Helvaci, S., Senova, A., Kar, G., & Gören, S. (2018). Improving driver behavior using gamification. *International Conference on Mobile Web and Intelligent Information Systems*, 193–204.

Inbar, O., Tractinsky, N., Tsimhoni, O., & Seder, T. (2011). Driving the scoreboard: Motivating eco-driving through in-car gaming. *Proceedings of the CHI 2011 Workshop Gamification: Using Game Design Elements in Non-Game Contexts*, 7–12.

Kalmpourtzis, G. (2018). *Educational Game Design Fundamentals: A Journey to Creating Intrinsically Motivating Learning Experiences*. CRC Press.

Kazhamiakin, R., Loria, E., Marconi, A., & Scanagatta, M. (2021). A gamification platform to analyze and influence citizens' daily transportation choices. *IEEE Transactions on Intelligent Transportation Systems*, *22*(4), 2153–2167.

Kazhamiakin, R., Marconi, A., Perillo, M., Pistore, M., Valetto, G., Piras, L., Avesani, F., & Perri, N. (2015). Using gamification to incentivize sustainable urban mobility. *2015 IEEE first international smart cities conference (ISC2)*, 1–6.

Khoshkangini, R., Valetto, G., Marconi, A., & Pistore, M. (2021). Automatic Generation and Recommendation of Personalized Challenges for Gamification. *User Modeling and User-Adapted Interaction*, *31*(1), 1–34. https://doi.org/10.1007/s11257-019-09255-2

Lee, J. J., Matamoros, E., Kern, R., Marks, J., de Luna, C., & Jordan-Cooley, W. (2013). Greenify: Fostering Sustainable Communities Via Gamification. *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, 1497–1502. https://doi.org/10.1145/2468356.2468623

Lombardi, P., & Giordano, S. (2015). Evaluating the Smart and Sustainable Built Environment in Urban Planning. In A. Vesco & F. Ferrero (Eds.). Y. Wang (Ed.), *Handbook of Research on Social, Economic, and Environmental Sustainability in the Development of Smart Cities* (pp. 44–59). IGI Global. https://doi.org/10.4018/978-1-4666-8282-5

Magaña, V. C., & Organero, M. M. (2014). The impact of using gamification on the eco-driving learning. In *Ambient intelligence-software and applications* (pp. 45–52). Springer.

Marconi, A., Schiavo, G., Zancanaro, M., Valetto, G., & Pistore, M. (2018). Exploring the World Through Small Green Steps: Improving Sustainable School Transportation with a Game-Based Learning Interface. *Proceedings of the 2018 International Conference on Advanced Visual Interfaces*, 1–9. https://doi.org/10.1145/3206505.3206521

Matallaoui, A., Hanner, N., & Zarnekow, R. (2017). Introduction to Gamification: Foundation and Underlying Theories. In S. Stieglitz, C. Lattemann, S. Robra-Bissantz, R. Zarnekow, & T. Brockmann (Eds.), *Gamification: Using Game Elements in Serious Contexts* (pp. 3–18). Springer International Publishing. https://doi.org/10.1007/978-3-319-45557-0

McCall, R., Koenig, V., & Kracheel, M. (2013). Using gamification and metaphor to design a mobility platform for commuters. *International Journal of Mobile Human Computer Interaction (IJMHCI)*, *5*(1), 1–15.

Montola, M. (2005). Exploring the Edge of the Magic Circle: Defining Pervasive Games. *Proceedings of Digital Arts and Culture*, 1–4.

Nakashima, R., Sato, T., & Maruyama, T. (2017). Gamification approach to smartphone-app-based mobility management. *Transportation Research Procedia*, *25*, 2344–2355.

Salat, S., Labbé, F., Nowacki, C., & Walker, G. (2011). *Cities and Forms: On Sustainable Urbanism*. CSTB Urban Morphology Laboratory : Hermann
OCLC: ocn770161250.

Schatten, M., Okreša Đurić, B., & Protrka, R. (2021). Conceptual Architecture of a Cognitive Agent for Telemedicine based on Gamification. In N. Vrček, E. Pergler, & P. Grd (Eds.), *Proceedings of the Central European Conference on Information and Intelligent Systems* (pp. 3–10). Faculty of Organization and Informatics.

Shi, C., Lee, H. J., Kurczal, J., & Lee, A. (2012). Routine driving infotainment app: Gamification of performance driving. *Adjunct Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 181–183.

Smith, G. G., Okreša Đurić, B., & Ricijaš, M. (2021). Automated Educational Game Design from EBooks. In N. Vrček, E. Pergler, & P. Grd (Eds.), *Proceedings of the Central European Conference on Information and Intelligent Systems* (pp. 17–24). Faculty of Organization and Informatics.

Steinberger, F., Schroeter, R., Lindner, V., Fitz-Walter, Z., Hall, J., & Johnson, D. (2015). Zombies on the road: A holistic design approach to balancing gamification and safe driving. *Proceedings of the 7th international conference on automotive user interfaces and interactive vehicular applications*, 320–327.

Tillström, J. (2012). Gamification in automotive marketing: A conceptual framework for implementation.

Wang, C., Terken, J., & Hu, J. (2014). " liking" other drivers' behavior while driving. *Adjunct Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 1–6.

Werbach, K., & Hunter, D. (2012). *For the Win: How Game Thinking Can Revolutionize Your Business*. Wharton Digital Press.

Zinke-Wehlmann, C., & Friedrich, J. (2019). Commute Green! The Potential of Enterprise Social Networks for Ecological Mobility Concepts. In L. M. Camarinha-Matos, H. Afsarmanesh, & D. Antonelli (Eds.), *Collaborative Networks and Digital Transformation* (pp. 128–139). Springer International Publishing. https://doi.org/10.1007/978-3-030-28464-0_12

Towards a Streamed Holographic 3D Game Engine

Markus Schatten, Bogdan Okreša Đurić, Tomislav Peharda
Faculty of Organization and Informatics, Artificial Intelligence Laboratory
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{markus.schatten,dokresa,tomislav.peharda}@foi.unizg.hr

Abstract. *In this work-in-progress article we present our initial attempt in building a streamed 3D holographic game engine. Game streaming technology provides the opportunity to use cloud-based resources like advanced graphic processors to play computer games on low-cost end-user devices. Holographic games are a relatively new technology which provides a new way of experiencing multimedia and games but are currently a high-cost. Herein we present a prototype game streaming engine we have built that allows us to stream holographic games to consoles based on Peppers's Ghost pyramids which emulates a holographic display.*

Keywords. game engine, game streaming, holographic, pepper's ghost, computer games

1 Introduction

A game engine is a computer program that enables the development and management of all aspects and computing resources required to run a game (e.g., graphics, sound, user interface, scripts, events, data storage, networking, artificial intelligence, physics, etc.). The user of a computer game engine is a computer game programmer. In this work-in-progress paper we are on our journey to develop a game engine that would allow games to be (1) streamed over a network and (2) displayed on a holographic display device.

Game engines are complex systems consisting of numerous subsystems (see figure 1). Our objective herein is not to *reinvent the wheel* and reimplement all parts of a game engine from scratch, but to focus on two particular aspects of game engines which will allow us to achieve the above stated objectives: game streaming and holographic display of games.

Thus, we have chosen to use an existing game engine (namely UPBGE - a community based continuation of the Blender Game Engine[1] and extend it with: (1) a game streaming architecture, (2) a display transformation to allow display on a holographic display.

The rest of this paper is organized as follows: firstly in section 2 we provide an overview of game streaming technology. Then in section 3 we familiarize the reader with holographic display technology. In section 4 we describe the prototype system that we have developed. In the end in section 5 we draw our conclusions and provide guidelines for future research.

[1] For further information, see https://upbge.org/

2 Game Streaming

The advancement of technology allows most households to own at least one personal computer (PC). Such PCs can usually suffice when playing video games is considered, most modern games included. The key and the most observable element of enjoying a game is observing rich and smooth graphics provided by the game designers, especially when combined with smooth and fluid storytelling and game mechanics, provided by the games' developers and other artists. The most important building block of a PC, in the context of providing enjoyable graphics, is graphics hardware, i.e. external graphics cards (as opposed to on-chip integrated graphics hardware). The main problem in the domain of graphics cards is their price, which soared during the last decade, and made high-tier PCs almost unavailable to average targeted consumers in e.g. the EU, North America, or Far East. This obstacle, in turn, makes it harder for gamers, i.e. variously skilled players of games, to enjoy the full potential of games and their developed graphics.

With the advent of access to the Internet, and the increasing speeds available in increasingly large areas and portions of the world, e.g. access to high-speed optical fibre across the EU, an average user can rely more on the constant availability of information, and the speed of recovering it whenever needed. Constant availability of the Internet rendered music and video storage, such as a CD-ROM, a DVD-ROM, a BluRay disks, or a USB memory drive, or even a portion of a disk with stored music files, almost obsolete. Reliable access to the sources that provide multimedia files as streams of data shifted the focus of an average user, shifted the behaviour towards legal use of various types of multimedia, and created new habits when multimedia consumption is considered. Even though they are dispersed amongst many streaming services, a rich collection of music, movies, TV shows, cartoon series,

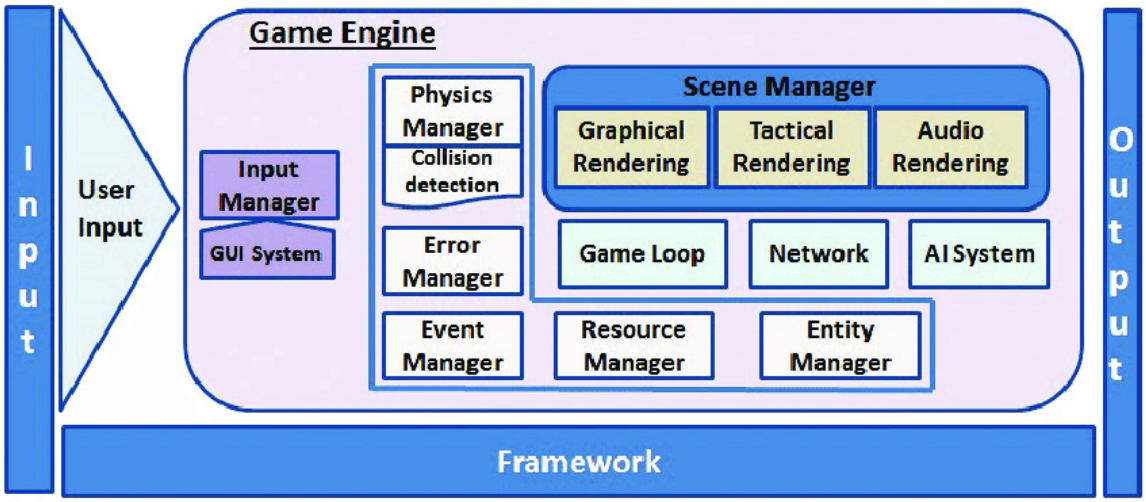


Figure 1: A game engine's architecture (Zarrad, 2018)

etc. is available in a matter of seconds to a great number of users throughout the world.

The trouble with streaming video games is that video games are an interactive type of multimedia entertainment. Videos and music can be streamed safely, but they are being streamed one way only: from the service provider to the user, and not much interaction is expected except for maybe starting a stream, pausing, moving the time line and other simple instructions. In streaming video games, the game is expected to behave according to the input of its player in real time. For example, if the player moves by pressing a key on the keyboard, the in-game character is expected to move almost instantly; if the player clicks on a menu item, the interface is expected to change accordingly; if the player's avatar is being attacked, or is found in a situation where speed is of the essence, the game is expected to react to the player's input almost instantly, etc. The described scenarios are only a few, a needle in a haystack of playing video games, that require almost instant reaction of the game – moreover, they require that the player receives an almost instant feedback that the game instantly reacted to the received player's input. Usually, establishing such a quick feedback loop is not a problem in PCs, but it does present a challenge in video game streaming services.

There is a caveat that should be addressed though. Not all games require feedback loops that are (as much as possible) instantaneous. Some games (usually single-player games) can be played with an occasional slight delay, because such a game, for example, does not rely on the speed of its player's reaction, or that particular feature is of low significance.

The benefit of combining the ideas of video games and multimedia streaming services is manifested in the opportunity to play video games without having necessarily to own the hardware that can run those video games. Usually, the time spent on gaming, i.e. playing video games, is only a fraction of the time spent on using the PC for other purposes. Therefore, owning hardware that will provide the resources for playing a video game can be expensive, especially when time, utility, and use are considered. On the other hand, a video game streaming service can be used whenever needed, and (in some cases) on virtually any device available, if the appropriate Internet speed and game-controlling hardware (e.g. a game controller, a keyboard, a mouse, a wheel, etc.) are at the player's disposal.

There are technologies and approaches that are used for further reducing the expected feedback loop duration, and providing smoother playing experience, such as predicting a player's move and rendering possible actions before they actually happen, or rendering only specific sets of pixels that change in consequent frames, etc. (Schatten et al., 2020)

Game streaming systems are usually complex and consist on numerous components that deal with various aspects of networking, game interaction as well as graphics rendering (see figure 2 for an example architecture based on virtual machines proposed by (Hossain et al., 2015)).

Many video game streaming platforms exist today (Greenwald, 2022), each of which uses a different method of providing games to be streamed, some of which are: NVIDIA GeForce NOW[2], Amazon Luna[3], Sony PlayStation Now[4].

The initial test case scenario for the holographic platform described in this paper will consist of a game that does not rely on the speed of its player, rather it will be used to showcase the possibilities of playing a game developed for, and played on, such a platform.

[2]For further information, see https://play.geforcenow.com
[3]For further information, see https://www.amazon.com/Luna/
[4]For further intofmation, see https://www.playstation.com/en-us/ps-now/

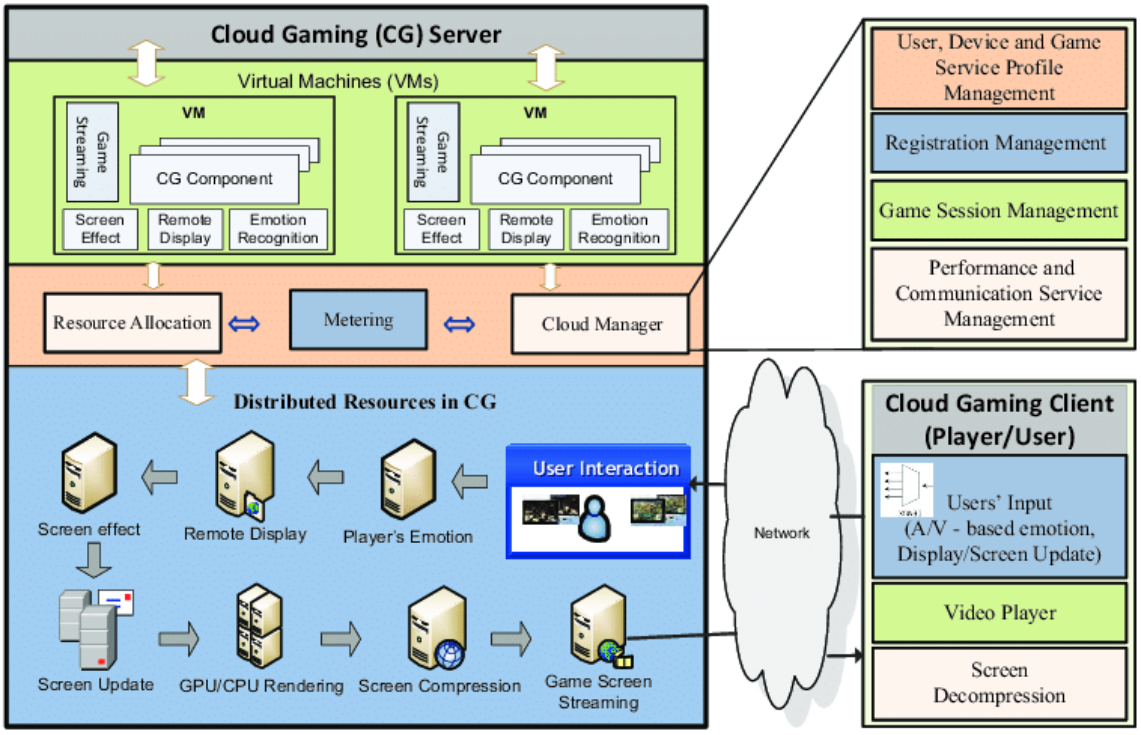


Figure 2: A game streaming architecture (Hossain et al., 2015)

3 Holographic Technology

A hologram is 3D visual display that is visible from all angles. At the time of writing there are a number of promising holographic technologies available (Blanche, 2022; Chang et al., 2018; Fadzli et al., 2022; Haleem et al., 2022; Li et al., 2022; Piliar, n.d.; Wang et al., 2022). From a commercial perspective, there are various technologies that will produce a similar effect.

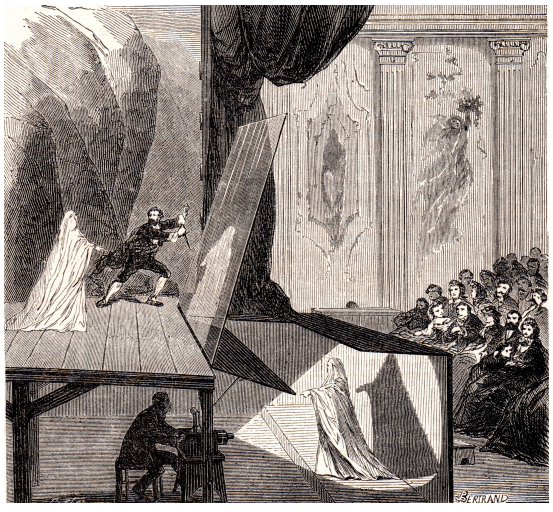


Figure 3: By Le Monde Illustré, Public Domain

One such technology are so called Pepper's Ghost images, after the English scientist John Henry Pepper (1821–1900) shown on figure 3. The basic idea behind this technology is to use a light display that emits an image on a reflective (usually transparent) surface which reflects the image and provides a holographic experience.

A common enhancement of the original system is to use a glass or acrylic four sided pyramid instead of just one surface, and place a display above the pyramid so that the holographic image is visible from all four sides of the pyramid.

On the other hand, voxel technology is based on producing an information that consists of volume and pixel. Pixels are physically created in space as floating objects. Displays that support such technology are called volumetric displays and stand for actual true holograms (Yang et al., 2015).

Holography based on fans are also drawing attention. The idea behind the technology is to transmit a light towards a running fan. Unless on a close proximity, one will not notice visually that a fan is spinning, however, transmitting a light on it will produce an effect of having floating objects from any distance (Prado Ortega et al., 2020).

4 Prototype Implementation

In order to provide a working proof-of-concept holographic we have designed and partially implemented a game engine for streaming holographic content. During design and implementation we have used multiagent systems (MASs) as a natural way of developing intelligent distributed systems. The architec-

ture of the system is shown on figure 4. Additionally, we have used a gaming console as a metaphor for the implementation. The system consists of a pool of existing game agents which comprise containers called cartridges that represent the games that are supported by the engine (implemented by actual game developers). Each cartridge is a microservice container (in our case Docker[5] container) which consists of a minimalized Linux operating system that includes a graphical user interface and a game engine (in our case UPBGE), the actual game implementation as well as a camera transformation that allows us to display the game on a holographic display based on Pepper's Ghost technology.

We have implemented this transformation in UPBGE by using a simple technique. In order to get the state of the 3D world for each of the four sides of the pyramid we have used four cameras which display the world in real time from four perspectives. Additionally, we have created a custom viewport that shows all four camera outputs on the same display at once, which allows us to use one display above the pyramid to show all four perspectives. The implementation of this script is shown in the following listing:

```
from bge import logic, render
camList = logic.getCurrentScene().cameras

cont = logic.getCurrentController()
own = cont.owner
camn = camList[ own[ 'camn' ] ]
cams = camList[ own[ 'cams' ] ]
camw = camList[ own[ 'camw' ] ]
came = camList[ own[ 'came' ] ]

a = render.getWindowWidth()
b = render.getWindowHeight()

x = int( ( a - b ) / 2 )
y = int( b / 4 )
z = int( b / 2 )

# setViewport( left, bottom, right, top )
camn.setViewport( x+y, y+z, x+y+z, b )
cams.setViewport( x+y, 0, x+y+z, y )
camw.setViewport( x+y+z, y, x+2*y+z, y+z)
came.setViewport( x, y, x+y, y+z )

camn.useViewport = True
cams.useViewport = True
camw.useViewport = True
came.useViewport = True
```

For example figure 5 shows the UPBGE interface for a simple world in which we have put an interactive 3D model of a dragon[6] on a simple plane. The dragon can be manipulated using the usual keyboard and mouse inputs in real time.

By using the script shown above, the world is rendered as shown on figure 6. As one can see, four cameras show the dragon from four sides (north, east, south, west) at the same time.

By using a transparent pyramid and placing the center of it in the center of the display by aligning the sides with the sides of the display, one can get a hologram-like appearance as shown on figure 7. In this way, any game implemented in UPBGE with the given transformation can be transformed in a holographic game. Of course, this transformation is not suitable for all existing games, but only for those for which a holographic display can or should be used. We envision that various new types of games, which make use of the holographic nature of the display in their core game mechanics can be implemented that go beyond games based on traditional non-holographic displays.

In addition to game agents, there are two other important types of agents: (1) game streaming agents as well as (2) videoconferencing agents. The gane streaming agents are envisioned as orchestrators. They get requests from clients that would like to play a game from the pool of available games, allocate resources in the could (which could be any orchestration platform like Kubernetes of Docker Swarm) and instruct a game agent to start it's cartridge as a container instance on the allocated end system. This allocation is considered a game instance in which an instance of a game agent directly communicates with an end-user's web based interface to which a game is streamed. The streaming is implemented by using XPRA[7] and a X11VNC[8] server which are contained in each game cartridge.

In addition to game streaming agents, videoconferencing agents are implemented as well that allow audiovisual and text communication between players (in multiplayer games for example) or live streaming of a players video and audio feeds (for live events for example). The game sgtraming agents instructs a videoconferencing agent to open an new room for each game that is being streamed and this new videoconference instance then directly communicates with the player's web based user interface. Videoconferencing is implemented using Janus WebRTC server[9].

In the end, the web based user interface features a VNC client that displays the streamed game and allows user interaction as well as a WebRTC client to allow for video and audio streaming.

All agents are implemented using Python, especially the SPADE platform (Palanca et al., 2020). The web based user interface is an adaptation of noVNC[10] with additional WebRTC client implementation in JavaScript.

[5]For further information, see https://www.docker.com/

[6]Available at https://3dmdb.com/en/3d-model/bge-dragon-20/1005315/?free=True&q=bge+dragon

[7]For further information, see https://xpra.org/

[8]For further information, see https://launchpad.net/ubuntu/+source/x11vnc

[9]For further information, see https://janus.conf.meetecho.com/

[10]For further information, see https://novnc.com/

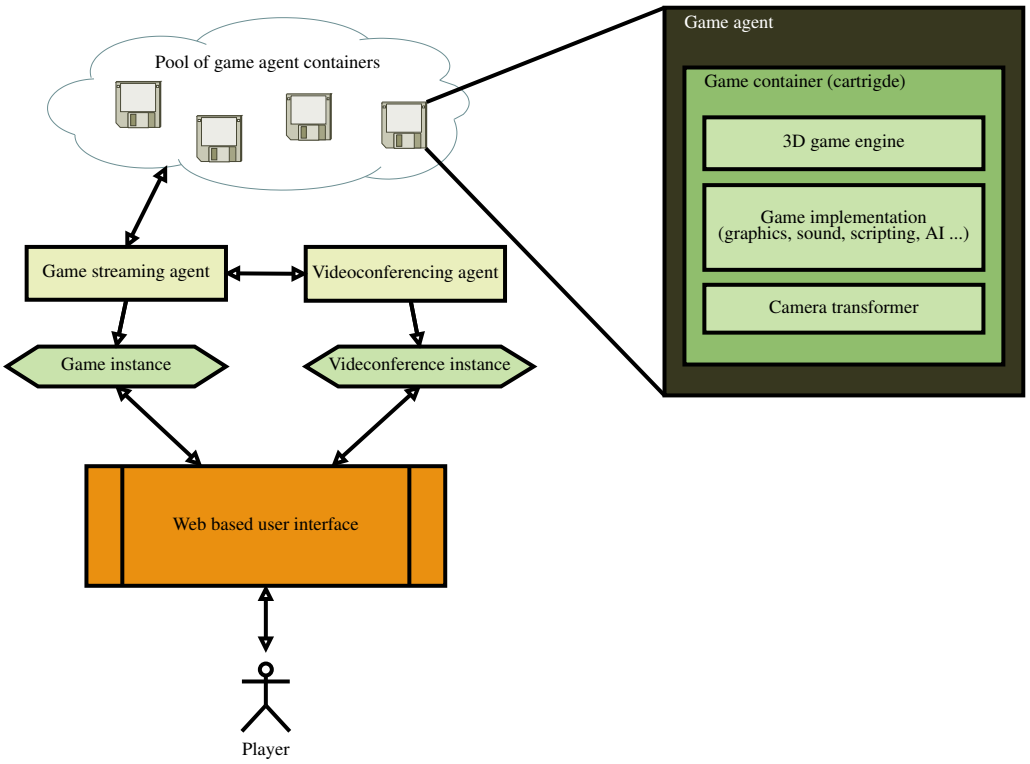


Figure 4: Prototype Architecture

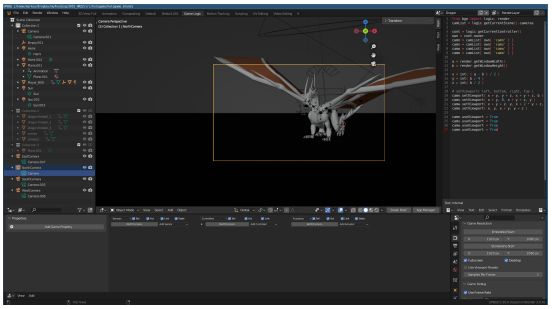


Figure 5: UPBGE interface

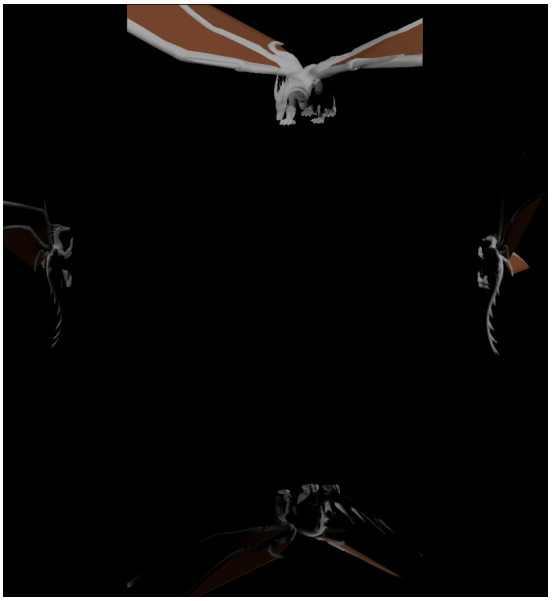


Figure 6: Display with four parallel cameras

5 Conclusion & Future Reasearch

In this paper we have provided a first step towards the implementation of a streamed 3D holographic game engine. We have developed a proof-of-concept game streaming platform that allows us to stream almost any game developed for the Linux X Window System to a web based VNC client. Additionally, we have extended the UPBGE game engine with a output transformation script that allows us to display games developed in the engine on a pyramidal holographic display device based on Pepper's Ghost technology.

Our future research is aimed towards enhancing this technology with possibly other 3D game engines that allow for multiple camera views that could be embedded into game cartridges as described above. Additionally, we are working on the design and assembly of a 3D holographic game console based on the technology described herein.

Acknowledgement

This work has been fully supported by the Croatian Science Foundation under the project number IP-2019-

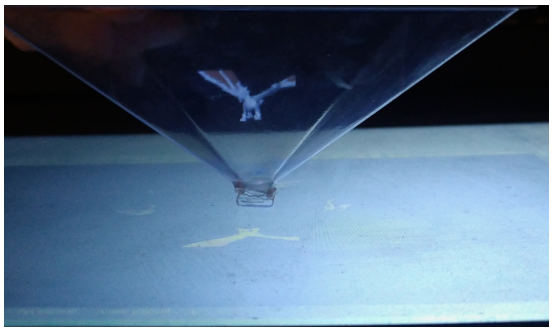


Figure 7: Hologram-like appearance

04-5824.

References

Blanche, P.-A. (2022). Towards a modular and scalable holographic display. *Light: Science & Applications*, *11*(1), 100. https://doi.org/10.1038/s41377-022-00786-9

Chang, E.-Y., Choi, J., Lee, S., Kwon, S., Yoo, J., Park, M., & Kim, J. (2018). 360-Degree Color Hologram Generation for Real 3d Objects. *Applied Optics*, *57*(1), A91. https://doi.org/10.1364/AO.57.000A91

Fadzli, F. E., Ismail, A. W., Abd Karim Ishigaki, S., Nor'a, M. N. A., & Aladin, M. Y. F. (2022). Real-Time 3D Reconstruction Method for Holographic Telepresence. *Applied Sciences*, *12*(8), 4009. https://doi.org/10.3390/app12084009

Greenwald, W. (2022). *The Best Game Streaming Services for 2022*. PCMag. Retrieved June 18, 2022, from https://www.pcmag.com/picks/the-best-game-streaming-services

Haleem, A., Javaid, M., Singh, R. P., Suman, R., & Rab, S. (2022). Holography and Its Applications for Industry 4.0: An Overview. *Internet of Things and Cyber-Physical Systems*, *2*, 42–48. https://doi.org/10.1016/j.iotcps.2022.05.004

Hossain, M. S., Muhammad, G., Song, B., Hassan, M. M., Alelaiwi, A., & Alamri, A. (2015). Audio–visual emotion-aware cloud gaming framework. *IEEE Transactions on Circuits and Systems for Video Technology*, *25*(12), 2105–2118.

Li, Z., Gao, C., Li, H., Wu, R., & Liu, X. (2022). Portable Autostereoscopic Display Based on Multi-Directional Backlight. *Optics Express*, *30*(12), 21478. https://doi.org/10.1364/OE.460889

Palanca, J., Terrasa, A., Julian, V., & Carrascosa, C. (2020). Spade 3: Supporting the new generation of multi-agent systems. *IEEE Access*, *8*, 182537–182549.

Piliar, M. (n.d.). Collaborative use of virtual and augmented reality in selected domains, 8.

Prado Ortega, M. X., Delgado Ramírez, J. C., Valarezo Castro, J. W., Armijos Carrión, J. L., Ávila Carvajal, A. A., & González Segarra, A. N. (2020). Application of the technical - pedagogical resource 3d holographic led-fan display in the classroom. *Smart Learning Environments*, (7).

Schatten, M., Tomičić, I., & Okreša Đurić, B. (2020). Towards application programming interfaces for cloud services orchestration platforms in computer games. *2020 Central European Conference on Information and Intelligent Systems*, 9–14.

Wang, J., Wu, Y., Lei, X., & Wang, Q. (2022). Bidirectional Phase Compensation for Curved Hologram Generation in Holographic Display. *IEEE Photonics Technology Letters*, *34*(11), 579–582. https://doi.org/10.1109/LPT.2022.3173439

Yang, L., Dong, H., & El Saddik, A. (2015). See in 3d: State of the art of 3d display technologies. *Multimedia Tools and Applications*, (75).

Zarrad, A. (2018). Game engine solutions. *Simulation and Gaming*, 75–87.

Data and Knowledge Bases

Andrea Toic, Patrizia Poščić, Danijela Jaksic
Analysis of Selected Business Intelligence Data Visualization Tools

Anita Ciunova-Shuleska, Nikolina Palamidovska-Sterjadovska, Violeta Cvetkoska, Irena Bogoevska-Gavrilova
Ten Years of Customer Engagement Research in Tourism and Hospitality: A Bibliometric Analysis

Martina Asenbrener Katic, Sanja Candrlic, Marina Rauker Koch
Adverb and Preposition Representation in Croatian and English using the Node of Knowledge Method

Jakob Jelenčič, Jože M. Rožanec, Dunja Mladenić
KL-ADWIN: enhanced concept drift detection over multiple time windows

Jelena Petrovački, Dejana Gladić, Sara Havzi, Darko Stefanović, Sonja Ristić
NoSQL Databases Deployment in Various Application Domains: A Systematic Literature Review

Ines Vlahović
Global Repeat Map algorithm as genomic technology for Higher Order Repeat identification

Analysis of Selected Business Intelligence Data Visualization Tools

Andrea Toic, Patrizia Poscic, Danijela Jaksic
Faculty of informatics and digital technologies
University of Rijeka
Radmile Matejcic 2, Rijeka, Croatia
{atoic, patrizia, dsubotic}@inf.uniri.hr

Abstract. *This paper presents a detailed overview of business intelligence and analysis tools that enable data-based decision-making. There are many self-service tools on the market today, and the paper compares three that were named leaders in Gartner's 2022 report: Power BI, Tableau, and Qlik. The research aims to compare these tools based on defined criteria and show their advantages and disadvantages, with a special focus on the use for educational purposes.*

Keywords. Business Intelligence, BI tools, data Analytics, Data Visualization, Data Reports, Qlik, Microsoft Power BI, Tableau

1 Introduction

In 2021, 79 billion terabytes (TB) of data were created in the world, which is 60 times more than the amount of data produced in 2010 and 6 times more than in 2017 (Statista Research Department 2021). To discover those relevant to their business in this sea of information, companies have begun a business analytics-oriented business transformation to meet the demands for faster and more comprehensive access to data, related analytics, and insights (Vesset et al. 2020). Unlike many other IT areas, business intelligence and analytics services continued to grow exponentially despite the pandemic. Faced with the new situation, organizations are turning to insights and intelligent automation solutions to maintain positive business continuity despite the COVID-19 crisis (Hamel 2021).

Nearly 50% of McKinsey study respondents agreed that analytics and Big Data have fundamentally changed the way things work in sales and marketing across all organizations (McKinsey Analytics 2018). Despite the increase in the number of tools designed to facilitate data collection, transformation, and processing, many companies are still unsure how best to handle them and whether they get feedback from data and analytics (Cardoso et al. 2020). Managers require information in the decision-making process, but research shows that, despite the information available, business decisions are still often made based on feelings and experience (Delen, Moscato, and Toma 2018). Companies maintain market competitiveness by using predictive analytics to predict maintenance and performance problems before they become larger and riskier (KPMG 2019).

The "rule of thumb" is gradually giving way to decision-making based on insights gathered from predictive analytics. Self-service is a group of BI and analytical tools that allow users to perform data processing and analysis without having a statistical and technological background. Data visualizations provide a clearer insight than the classic, tabular presentation from which it is difficult to identify relevant, hidden information in the data and improve the decision-making process and strategic planning. This lowers operating costs and enables higher profits thanks to accurate analysis and forecasting (Zhao Yifan 2021). Over the next five years, Cloud computing, Internet of Things (IoT), and artificial intelligence / machine learning (AI / ML) are projected to have the largest impact on enterprise analytics initiatives, with 65% planning to increase their analytics spending (MicroStrategy 2020). The result is a departure from traditional BI platform markets and a focus on data science, machine learning software, and cloud data analysis tools. Business intelligence in the cloud provides an agile way to access BI applications. Most tools on the market also offer their service in the cloud, which allows availability on multiple devices and a web browser, reduces costs and allows better accessibility and faster implementation so the need to access data anytime and anywhere is becoming a top priority for business leaders. According to Gartner's 2020 study, by 2025, data stories will be the most widespread way of consuming analytics, and 75% of stories will be automatically generated using extended analytics techniques (Richardson et al. 2020).

Based on an analysis of business software solutions for 2021, Gartner (Richardson et al. 2021) and Forrester (Evelson et al. 2021) ranked business intelligence tools by customer ratings and unbiased methodology.

The latest report for 2022 shows that for the second-straight year, there are again only three tools in the Leaders quadrant - Microsoft Power BI (Power BI documentation n.d.), Tableau which is now a part of Salesforce (Tableau n.d.), and Qlik (Qlik Help n.d.).

This paper seeks to facilitate the selection of the most appropriate of these tools for business intelligence and analysis based on selected criteria, with a special focus on educational usage.

The paper is organized as follows. Section 2 describes Business Intelligence and Analytics. Section 3 contains Related work. Section 4 is about the Power BI tool in general, section 5 is about Tableau, and section 6 is about Qlik tool. Section 7 explains the evaluation criteria used in this study and presents a comparison of tools. Section 8 explains the Future work and Conclusion.

2 Business Intelligence and Analytics

Business intelligence (BI) refers to a set of methods, technologies, and related tools that help gather essential information from a wide range of unstructured data and transform it into actionable information to improve business decision making (Niu, 2021).

Business analytics (BA) uses methods and tools that are faster and more technical than business intelligence, such as data mining, regression analysis, correlation analysis, segmentation, text mining, modeling, and machine learning (Zhao Yifan 2021). In other words, BA is the process of using quantitative methods to extract meaning from data to make more informed business decisions. Advanced analytics through data preparation using machine learning (ML) and artificial intelligence (AI) simplify the creation and explanation of the visualizations obtained, especially thanks to the ability to create queries in natural language (Richardson et al. 2021). For meaningful interpretation and discovery of values from their temporally, spatially, and hierarchically variable forms, Big Data requires real-time analytics (Özdemir and Hekim 2018). Big Data analytics that can be visually presented and interpreted is a management tool for gaining a competitive advantage when making strategic decisions that increase the company's value and revenue (Akter et al. 2016). This change has led to more efficient, accurate, and objective decision-making.

Business intelligence and Analytics (BI&A) is a subfield of decision support systems (DSS) that seeks to turn more data into deeper insight (Phillips-Wren, 2021). DSS are based on artificial intelligence techniques and use data and models to help analyze the problem through visual representation. Data analysis provides an understanding of specific data handling approaches and techniques that include data visualization, research, discovery, or a combination of these approaches. Techniques that support data analysis (Orlovskyi and Kopp 2021) include Reporting, Dashboards, ad-hoc queries, OLAP, Data mining and machine learning, and Scorecards.

Data visualization is preceded by the process of data access, collection, and transformation. The Extract Transform Load (ETL) pipeline is a vital procedure in the Business Intelligence (BI) workflow. It is the process of structuring data for querying or analysis (Machado, 2019).

Working with data is becoming more accessible to a wider audience thanks to the natural language queries (NLQ) implemented in the tools. Self-service tools for analytics and business intelligence mostly have similar features and are no longer compared only based on the ability to create interactive dashboards and key performance indicators (KPIs) but also on how well they support extended analytics.

3 Related work

The Gartner Magic Quadrant is a report on qualitative market research aimed at ranking and evaluating technology suppliers and facilitating the selection of the right tool. The 2022 Magic Quadrant for Analytics and Business Intelligence Platforms report by Gartner assessed the strengths and weaknesses of the 20 most important service providers based on their ability to execute and completeness of vision. The graph is divided into four quadrants: niche players, challengers, visionaries, and leaders. Microsoft Power BI, Tableau (Salesforce), and Qlik took the leading positions in it (Richardson et al. 2021).

Fifteen BI tools were compared with a brief analysis of technical details, the advantages of using business strategy, and testing the features of these BI tools was conducted based on data from twelve different review rating websites so that business users could choose the best tool for the organization (Srivastava et al. 2021). The comparison concluded that Kyubit is the best and most adaptable tool for organizations.

In (Oliveira and Bernardino 2020) the authors compare 5 free solutions including the three most popular free but Closed source solutions (Power BI Free, Qlik View, and Tableau Public) and two open-source solutions (Metabase and Pentaho). The comparison criteria were selected based on 15 critical Self-Service functions according to Gartner and the recommendations of the Predictive Analytics Today Research website. After comparing a set of free features, it turned out that all three Closed source solutions offer more functionality at no cost to SMEs that analyse their data.

A survey on business intelligence tools for marketing, transportation, and financial services presents a comparison of the tools Tableau, Power BI, Pentaho, QlikView, and Micro Strategy Analytical Express (Chavva, Sangam, and Rao 2019) concluded

that Pentaho is the best tool in terms of data processing and services.

Tableau, Power BI, Sisense, and Qlikview (Lousa, Pedrosa, and Bernardino 2019) were compared based on the criteria from (Caldarola and Rinaldi 2017).

Based on the analysis, it was concluded that the choice of the most appropriate tool depends on the type of organization, but that Tableau and Power BI are the leading tools in the field of visualization.

The authors reached different conclusions, which shows that all tools have many commonalities and that nuances play a role in choosing the best tool for a particular type of business. In the fifth chapter, comparisons are made according to the most important criteria for use in education.

4. BI tools

The following subsections provide a brief overview of the history, versions, and functionality of the BI tools that are ranked best in Gartner's Magic Quadrant - Power BI, Tableau, and Qlik.

4.1 Power BI

Power BI (Power BI Documentation, n.d.) was conceptualized by Ruler and Dhers Netz of the SQL server coverage services team at Microsoft, and the first version was released in 2015. It was originally designed as an add-on to Excel, but due to a significant increase in functions, the tool was launched as a separate application based on Power Query, Power Pivot, and Power View concepts from Excel.

Power BI offers the following products: Power BI Desktop (free application) and Power BI Service (SaaS segment used to collaborate and distribute Power BI reports) for creating visualizations and reports. Power BI Report Builder is a tool for creating paginated reports published to Power BI services, and Power BI Embedded is a service to embed visuals, reports, and dashboards into an application. Users can view on-premises Power BI reports on Power BI Report Server and all these reports and dashboards are also available in the Power BI mobile apps.

Microsoft does not offer customers a choice, so the tool is only available as a SaaS service running in the Azure cloud. Power BI is currently available as part of the Office 365 E5 Cloud suite of applications, which has given it a huge market reach. By integrating tools into Microsoft Teams, users can access data and dashboards from a centralized location and make business decisions together with co-workers.

4.2 Tableau

Tableau (Tableau n.d.) is one of the oldest self-service platforms for visual analytics, founded in 2003 in Mountain View, California. In 2019 the company was bought by Salesforce. Co-founders Chris Stolte, Pat Hanrahan, and Christian Chabot developed and patented the VizQL, a visual query language that allows users to create and customize visualizations more than any other BI platform. Through a graphical interface, the tool converts user drag-and-drop actions into database queries, giving non-technical users the ability to create custom dashboards quickly and easily.

The Tableau platform supports the entire cycle from access and data preparation to analysis and sharing. Tableau data analytics can be divided into two categories, developer tools, and sharing tools. The developer tools include Tableau Desktop and Public, which are used to develop dashboards and create reports, visualizations, and charts. Sharing tools include Tableau Online, Server, and Reader whose function is to share existing visualizations and reports. In addition to computer versions, there are also mobile Tableau iOS and Android applications that allow users to monitor dashboards from mobile devices.

4.3 Qlik

Qlik (“Qlik”, n.d.) was founded by Staffan Gestreliu and Bjorn Berg in 1993 in Sweden and offers two BI tools on the market, traditional Qlik View and modern data exploration platform Qlik Sense. QlikView is used for providing analytical solutions, data visualization, delivering insights, and for business growth and monitoring. It is also used to generate models, applications, and dashboards for data analytics, and for understanding the data trends. On the other hand, Qlik Sense is the next version of QlikView launched in 2014. Qlik Sense has some advanced features specifically used for visualizing and analyzing the data. It helps build interactive dashboards and reports, extract data from a variety of data sources, and thanks to expanded analytics enable better and more efficient data visualizations.

The Qlik Analytics Modernization Program encourages and helps existing Qlik View users to switch to Qlik Sense and use its next-generation BI capabilities.

According to Gartner’s research, Qlik’s market momentum remains the lowest compared to the other two tools. Although they provide customer support to existing customers when switching to Qlik Sense, many users choose to modernize their business by choosing competing tools (Richardson et al. 2021).

5 Comparison

To assess the potential of visualization tools for educational purposes and student projects, they were compared based on specific characteristics important to students and professors - prices and types of licenses, supported operating systems, storage capacity, existing training, certification, and community and user support. The most important thing for new users is the existence of an academic or trial

version of the tool and the availability of many learning materials and customer support.

A critical review of available tools for real-time information processing and their application in production was also conducted. The study of related works also identified several criteria for comparing the most sought-after functionalities based on the needs and requirements of the industry, which are compared in subchapter 5.6.

5.1 Prices and types of licenses

Tools were compared based on price and license duration.

The price for Power BI depends on the selected product. Power BI Desktop comes free if you are already licensed with Office 365. Power BI Pro (for collaboration, included in Microsoft 365 E5) and Power BI Premium (use for business purposes, includes auto scale and all the features available with Power BI Pro) are paid and most of the widely used functionalities are only available in the Premium version. Licenses are paid monthly per dedicated cloud compute and storage resource and the free trial period lasts 60 days.

While Microsoft offers a lot of features at a relatively low price, Tableau has a reputation for being expensive. Tableau Public is the free version of the tool with limited capabilities. Tableau license prices are not based on a standard model for the entire organization but the role of the user. Each role has a different level of access and rights. There are three types of users – Creator (includes Tableau Desktop, Tableau Prep Builder, and one Creator license of Tableau Server or Tableau Online), Explorer (includes One Explorer license of Tableau Online), and Viewer (includes One Viewer license of Tableau Online). Tableau offers to students a free one-year license, while start-ups get a 20% discount on the license price. Tableau Desktop trial lasts 14 days.

Qlik offers a choice of SaaS solutions for use in groups and teams or a multi-cloud solution to make the most of an organization's data analytics. Prices are fixed per individual and per month, and for users of Qlik View and Qlik Sense tools there is a reduced dual-use offer. The limited resource version is free, and the free business trial lasts 30 days.

Table 1. Price and license comparison

VERSION	PRICE	FOR
Power BI Pro	$9.99 per user/ month	For individuals
Power BI Premium	$20 per user/ month	For individuals
	$4,995 Per capacity/month	For Teams & Organizations
Tableau Creator	€70 user/month billed annually	For individuals, Teams & Organizations
Tableau Explorer	€42 user/month billed annually	For Teams & Organizations
Tableau Viewer	€15 user/month billed annually	For Teams & Organizations
Qlik Sense Business	$30/user/month, billed annually	For individuals, Teams & Organizations
Qlik Sense Enterprise SaaS	Price is individually negotiated	For Teams & Organizations

5.2 Operating systems

Table 2 lists the requirements that must be fulfilled by the target system to successfully install and run desktop versions of Power BI, Tableau, and Qlik Sense. Power BI and Qlik don't run on Mac while Tableau supports Windows and Mac OS.

Table 2. Hardware/software requirements

	Power BI Desktop	Tableau Desktop	Qlik Sense
OS	Windows 8.1 or later	Windows 8/8.1 or later, Mac	Windows 10 (x64)
Server	Windows Server 2012 R2, or later	Microsoft Windows Server 2016/ 2019	Microsoft Windows Server 2012 R2 /2016 /2019
RAM	2 GB min, 4 GB or more recommended	2 GB	8 GB min
CPU	1 gigahertz (GHz) 64-bit (x64) processor or better	Must support SSE4.2 and POPCNT instruction sets	Multi-core x64 compatible processors with AVX support

Table 3. Supported web browsers and mobile OS

	Power BI	Tableau	Qlik Sense
Web browser	Internet Explorer 11 or later	Chrome, Microsoft Edge, Mozilla Firefox & Firefox ESR, Apple Safari	Microsoft Edge, Google Chrome, Mozilla Firefox, Apple Safari
Mobile	iOS 12.2 or later or Android 6.0 or later	iOS 13 or later and Android 8 or later	iOS 14/, iPadOS 14 or later, Android 10 or later

All three tools can create data visualizations directly in a browser with data connectors and view them in the mobile app. Table 3 shows supported web browsers and mobile apps for Android and iOS.

5.3 Storage capacity

Capacity includes storing your own and shared reports to which the user is connected, dashboards, uploaded images, etc. Power BI is compatible with Microsoft Azure, Tableau is compatible with modern cloud platforms including Azure and AWS, and Qlik Sense offers a SaaS cloud product.

Table 4 shows the Cloud storage capacity of each of the tools depending on the license.

Table 4. Data capacity

	Capacity Type	Capacity
Power BI	Pro	10 GB
	Premium	100 TB
Tableau	Online	100 GB
Qlik Sense	Business	250 GB
	Enterprise SaaS	500 GB

5.4 Training

Microsoft Power BI, Tableau, and Qlik provide great learning opportunities officially for free to get started with the tools. Their websites provide links to documentation that includes a variety of tutorials, webinar recordings, and blogs with helpful tips on new features. All these programs offer academic programs and discounts.

While Power BI has one exam that covers all the main functionalities of the platform, Tableau and Qlik offer a wide range of certification options designed for specific professions.

Table 5. Available types of training

TRAINING	Power BI	Tabelau	Qlik
In-Person	No	Yes	Yes
Live Online	No	Yes	Yes
Webinars	Yes	Yes	Yes
Documentation	Yes	Yes	Yes
Videos	Yes	Yes	Yes

5.5 Community and Customer Support

Power BI is limited in terms of customer support and has a smaller community that is active on forums. Each year they organize a Power BI Summit with more than 100 sessions and regularly explain new features on the blog and Twitter. Training in the form of short videos on Youtube and personal lab training are also available.

Tableau has the most inclusive community and dedicated Customer support for questions related to the tool. Its inclusive community forum for discussions with over 220,000 users provides many resources and posts on forums and over 500 global user groups. Tableau organizes numerous conferences as well as virtual events such as the Tableau Live discussion panel.

The Qlik community consists of about 150,000 users and has an active forum. In addition to the Knowledge base, a Q&A chat with experts is organized every Tuesday, and once a month there is a Techspert Talks, a free webinar on important topics where experts answer questions. Qlik Fix is a series of short videos on various functionalities, and Support Chat, webinars, blogs, and articles provide answers in the way that best suits users.

Table 6. Customer support available

SUPPORT	Power BI	Tabelau	Qlik
Email/Help Desk	Yes	Yes	Yes
FAQs/Forum	Yes	Yes	Yes
Knowledge Base	Yes	Yes	Yes
Phone Support	No	Yes	Yes
24/7 (Live Rep)	No	Yes	Yes
Chat	Yes	Yes	Yes

5.6 Functions and functionalities

For a data analysis visualization tool to be at the forefront of an analyst's choice, the following criteria can be usually considered. Table 7 shows a comparison of tools and the leader of some category.

Techniques that support data analysis are Reporting, Dashboards, ad-hoc queries, OLAP, Data mining and machine learning, and Scorecards.

Table 7. Comparison

	Power BI	Tableau	Qlik
Ad-hoc analysis	+	+	+
Advanced analytics capabilities	Lead	+	-
Analytics	+	Lead	+
Augmented Analytics	Lead	+	+
Associative engine	-	-	+
Connectivity	+	+	+
Dashboards	+	Lead	+
Data management	+	+	+
Data preparation tools	Lead	+	+
Data visualization	+	Lead	+
Document Management	+	+	+
Embedded analytics	+	Lead	+
OLAP	+	+	+

Ad-hoc analysis
Ad-hoc analyses are intended for business analysts, to determine the measures and dimensions of data with the help of SQL queries but without the use of OLAP servers, ie without OLAP cubes, with which we create arbitrary derived tables in the next phase. In this category, all three tools proved equally successful.
Advanced analytics capabilities
In terms of software integration, Tableau is the best solution for visual analytics and reporting that connects directly to R, Python, and other statistical or data mining environments. Power BI supports R-based visualization while Qlik does not support R or Python-based projects.
Analytics
Power BI becomes slow when handling large amounts of data but proves to be the best tool for live data analysis (performing database queries and retrieving results in real-time). Power BI Desktop provides the following out-of-the-box analytics functionalities: Statistical Summary that includes Histogram, Top N, Outliers using Scatter plot, Grouping and Binning, Clustering, and Time series analysis.

Thanks to its VizQL engine, Tableau is suitable for working with large amounts of data and offers extensive data visualization features. Despite the slower ability to analyze real-time data, the tool has proven to be good in terms of live query capabilities and extracts. Tableau has the following analytics functions: Summary Statistics, Trend analysis, Basic Forecasting, and Clustering.

Qlik has the best analytical features thanks to a flexible platform that can be combined without the need for initial preparation. Guided analytics and an associative engine help user discover information hidden among data.
Augmented Analytics
Augmented Analytics is a class of artificial intelligence and machine learning that help users to prepare data, generate insights and explain insights to increase the way people research and analyze data in analytics and BI platforms.

With Cognitive Services in Power BI, it is possible to apply various algorithms from Azure Cognitive Services for data enrichment. Power BI Premium also supports emotion analysis, key phrase extraction, language detection, and image tagging services.

As part of its expanded intelligence, Tableau offers Ask Data and Explain Data functions. Ask Data is the functionality of writing questions in natural language and creating visualizations in response to the query. Explain Data is based on advanced statistical models and suggests possible explanations for the selected label. Einstein Discovery is machine learning-driven and delivers predictions and recommendations of prediction models without the need to write algorithms.

Qlik also allows users to ask questions in natural language and automatically suggests insights and new links for research. Qlik's Cognitive Engine uses artificial intelligence and ML to automatically offer the user the best visualizations for selected data and minimize cognitive bias by learning from data.
Connectivity
Gartner defines connectivity as "Capabilities that enable users to connect to, query and zing for performance". Data source support is broadly categorized into the following: File, Database (on-premises and cloud), Online data sources (apps), and Others (web services).

Both Tableau and Power BI easily integrate with third-party data sources. As Power BI was developed by Microsoft it ensures easy integration with Azure and the PowerApps ecosystem. On the other hand, Tableau was recently acquired by Salesforce, so their integration runs more smoothly than with other tools. Qlik also has a wide range of connectors and other data connection types but lags behind the other two tools.
Dashboard
The dashboard is a set of several views that allows the user to compare different data at the same time instead of viewing separate sheets. Due to the built-in animations and dashboard templates in this category, Tableau has an advantage.
Data engine
Most data engines are based on linear search patterns – the user makes a query, and the data engine searches predetermined data sets for the solution. Power BI and Tableau run on an SQL database, which means that queries follow a predefined linear path, giving solutions from query-based searches.

Qlik Sense, on the other hand, uses a non-query-based associative engine. The associative engine combines all the data, maps the relationships between the data, and creates a compressed binary index optimized for interactive analytics so users can see connections or patterns that a typical data engine would probably have missed.
Data management
The BI tool should allow users the freedom to research data while preventing user error and data contamination. Working on a local computer allows more flexibility for the individual while working on a central server allows for collaboration and validation of changes, but users cannot work offline.

Power BI allows offline work on a local computer and sending finished books to a single server to pass a pre-release check. Tableau also allows you to work on the user's hard drives and check through a central server. Qlik Sense, unlike previous tools, offers only cloud work. All users work on the same server, which prevents the circulation of old content or its pollution.
Data preparation tools
All three compared tools contain ETL functionalities. Using these applications, data analysts can easily create data flows using well-known self-service tools, which reduces the time required and the cost of data transformation.

Power BI has an inbuild data preparation and cleaning functions tool Power Query that offers data flow creation, permission management, scheduled

reporting, etc. Also provides options natively for users to model their data which includes defining relationships between different data sources and writing code (or using drag-drop functionalities) to clean data. Tableau for data modification opens a separate program called Tableau Prep Builder which allows operations such as filtering, sharing, renaming, rotating, merging, and more to clean and format data. Qlik Compose is an agile platform for data warehouse automation and transformation. Integration with Qlik Replicate enables real-time ETL by combining real-time source data integration with automated ETL generation and supports the broadest ecosystem of heterogeneous data sources, including relational, legacy, and NoSQL data warehouses.

Data visualization

All tools offer traditional business visualizations out of the box, such as bar, line, area, histograms, bullet charts, treemaps, scatterplots, tables, and map-based visuals.

In the data visualization category, Tableau leads with an intuitive interface tailored to non-technical users that enables data aggregation, grouping, segmentation, and real-time analytics.

Document Management

All programs can convert analytical reports to a variety of formats including PDFs, spreadsheets, images, and crosstabs.

Embedded analytics

BI tools enable the implementation of built-in analytics into existing workflows. In this way, it is possible to improve insights without losing efficiency by implementing them in existing applications, portals, and workflows.

Power BI does not give priority to the API, which initially makes it difficult for users to implement, but the tool offers several objects, not just built-in dashboards. Tableau users can only install dashboards into their applications and processes, while other objects such as metrics or individual values are not available for deployment. Qlik Sense allows users to embed metrics, individual values, numbers, and full or partial dashboards in edge devices, portals, and workflows. The Qlik API is quite intuitive for users to use.

OLAP

OLAP (Online Analytical Processing) enables faster queries and calculations and different measures/dimensions according to the needs of the end-user. All compared programs share standard OLAP functionalities.

6 Conclusion

Data analysis and visualization are two key skills that can be applied in almost any field of work. An analysis of existing research on the best data visualization tools has shown that there is no comparison of tools based on criteria relevant to students and professors for use in educational purposes. Using open-source tools in teaching instead of trial versions of the most popular tools used in industry is a major drawback when learning this important analytical skill.

In this paper, three tools that occupy the position of a leader according to Gartner's magic quadrant are analyzed and compared based on criteria important for application for educational purposes.

Comparing Power BI, Tableau, and Qlik software, Tableau proved to be the most suitable tool for beginners. Through the Tableau for Students program, the tool offers students a one-year free academic license at accredited academic institutions, including the Tableau Desktop, Tableau Prep, and Tableau Public tools. The license can be renewed every year as long as the student status is active. In addition to free software, they also provide students with a 20% discount on the Tableau certificate. Student Resources Page is available for each product as well as on-demand training videos. Tableau Desktop and Tableau Prep are available for Windows and Mac operating systems and Android and iOS mobile devices. The Tableau program for students does not include access to Tableau Online or Tableau Server but visualizations can be shared with colleagues using the free Tableau Public platform. Of all the tools being compared, Tableau has the largest community of users, and groups such as Student Resource are of great help to first-time data visualization students. Thanks to the simple drag-and-drop interface, the tool has an initial gentle learning curve and is not difficult for beginners to understand to use the basic functionalities. All the compared tools have similar functionalities, and it is difficult to give preference to one of them without an exact specification of the area for which they will be used. Most of the tools have focused their vision of development on ML, AI, and NLP. Some of Tableau's functionalities, such as Querying in natural language with Ask Data, Explain Data, and Einstein Discovery are also innovations that can greatly help new users. Data visualization tools continue to evolve in features that enable user-friendliness, simplified workflows, and predictive analytics and every year they become more accessible to ordinary users in everyday work.

Thanks to the growing intuitiveness of these tools, the skills of business analysts are becoming more and more accessible, and the need for data-based decision-making is crucial for successful business in any sector. All students should be familiar with the basics of using these tools that employers value highly because of their contribution to profitability, data-based decision-making, and business performance. Educational institutions should recognize this and include this software in their curricula.

Acknowledgments

This paper is based upon work supported by the University of Rijeka under project "uniri-drustv-18-182".

References

Akter, S., Wamba, S. F., Gunasekaran, A., Dubey, R., & Childe, S. J. (2016). How to improve firm performance using big data analytics capability and business strategy alignment?. *International Journal of Production Economics*, 182, 113-131.

Caldarola, E. G., & Rinaldi, A. M. (2017). Big data visualization tools: a survey. *Research Gate*.

Cardoso, M., Guimarães, T., Portela, C. F., & Santos, M. F. (2020). Data extraction and exploration tools for business intelligence. In *Fourth International Congress on Information and Communication Technology* (pp. 489-497). Springer, Singapore.

Delen, D., Moscato, G., & Toma, I. L. (2018, January). The impact of real-time business intelligence and advanced analytics on the behaviour of business decision makers. In *2018 International Conference on Information Management and Processing (ICIMP)* (pp. 49-53). IEEE.

Evelson, Boris, Srividya Sridharan, Danielle Jessee, and Fayzan Sabri. 2021. *The 15 Providers That Matter Most and How They Stack Up.*

Hamel, Jennifer. 2021. "Worldwide Business Analytics Services Forecast, 2021–2025." https://www.idc.com/getdoc.jsp?containerId=US47133321&utm_medium=rss_feed&utm_source=alert&utm_campaign=rss_syndication (April 2, 2022).

KPMG. 2019. *Emerging Trends in Infrastructure*. McKinsey Analytics 2018. Analytics Comes of Age.

Lousa, A., Pedrosa, I., & Bernardino, J. (2019, June). Evaluation and Analysis of Business Intelligence Data Visualization Tools. In *2019 14th Iberian Conference on Information Systems and Technologies (CISTI)* (pp. 1-6). IEEE.

Machado, G. V., Cunha, Í., Pereira, A., & Oliveira, L. B. (2019). DOD-ETL: distributed on-demand ETL for near real-time business intelligence. *Journal of Internet Services and Applications, 10*(1), 1-15.

MicroStrategy. 2020. "Global State of Enterprise Analytics."

Niu, Y., Ying, L., Yang, J., Bao, M., & Sivaparthipan, C. B. (2021). Organizational business intelligence and decision making using big data analytics. *Information Processing & Management*, 58(6), 102725.

Oliveira, A., & Bernardino, J. (2020). Evaluating Self-Service BI and Analytics Tools for SMEs. In *ICETE (3)* (pp. 89-97).

Orlovskyi, D. L., & Kopp, A. M. (2021). *An overview of business intelligence core capabilities and software platforms* (Doctoral dissertation, Черкаський національний університет ім. Б. Хмельницького).

Özdemir, V., & Hekim, N. (2018). Birth of industry 5.0: Making sense of big data with artificial intelligence,"the internet of things" and next-generation technology policy. *Omics: a journal of integrative biology, 22*(1), 65-76.

Phillips-Wren, G., Daly, M., & Burstein, F. (2021). Reconciling business intelligence, analytics and decision support systems: More data, deeper insight. *Decision Support Systems, 146*, 113560.

"Power BI Documentation." https://docs.microsoft.com/en-us/power-bi/ (April 18, 2022).

"Qlik Help." https://help.qlik.com/ (April 19, 2022).

Reddy, C. S., Sangam, R. S., & Srinivasa Rao, B. (2019). A survey on business intelligence tools for marketing, financial, and transportation services. In *Smart intelligent computing and applications* (pp. 495-504). Springer, Singapore.

Richardson, James et al. 2020. *Magic Quadrant for Analytics and Business Intelligence Platforms*. https://www.gartner.com/doc/reprints?id=1-1Y7VEZB3&ct=200128&st=sb. (March 24, 2022).

Richardson, James et al. 2021. *Magic Quadrant for Analytics and Business Intelligence Platforms*.

Srivastava, G., Muneeswari, S., Venkataraman, R., Kavitha, V., & Parthiban, N. (2021). A review of the state of the art in business intelligence software. *Enterp. Inform. Syst.*

Statista Research Department. 2021. "Total Data Volume Worldwide 2010-2025." https://www.statista.com/statistics/871513/worldwide-data-created/ (April 16, 2022).

"Tableau." https://www.tableau.com/ (April 21, 2022).

Vesset, Dan et al. 2020. *IDC FutureScape: Worldwide Data, Integration, and Analytics 2020 Predictions*

Zhao, Y. (2021). Transformation of Business Analytics from Business Intelligence. In *E3S Web of Conferences* (Vol. 253, p. 03013). EDP Sciences

Ten Years of Customer Engagement Research in Tourism and Hospitality: A Bibliometric Analysis

Anita Ciunova-Shuleska, Nikolina Palamidovska-Sterjadovska, Violeta Cvetkoska, Irena Bogoevska-Gavrilova

Ss. Cyril and Methodius University in Skopje,

Faculty of Economics – Skopje,

Boul. Goce Delchev 9V, 1000 Skopje, Republic of North Macedonia

{anita, nikolina, vcvetkoska, irena.bogoevska.gavrilova}@eccf.ukim.edu.mk

Abstract. *This research study aims to review and synthesize customer engagement (CE) research in tourism and hospitality published between 2012 and 2021. A total of 134 articles published in this period were collected from the SCOPUS database. A systematic review through bibliometric analysis was conducted and the results of the descriptive analysis along with the results from keywords analysis based on text mining in the abstracts were presented. This study represents an attempt at providing a comprehensive bibliometric analysis of CE research in the tourism and hospitality field following the PRISMA protocol and using* a *VOSviewer software.*

Keywords. Bibliography, customer engagement, tourism, hospitality

1 Introduction

Customer engagement (CE) is the new buzz word in the marketing literature. The inclusion of customer engagement as a research priority topic of the Marketing Science Institute (2010, 2020) is a verification of the prominence of this concept (So, Li & Kim, 2020). The rise of social media enhances customer-brand interactions which increase the initial practical and research interest in CE (Tsiotsou, 2019). Despite the ample academic discussion of CE, there is no consensus on the conceptualization, dimensionality, and operationalization of this concept (Syrdal & Briggs, 2018). While some researchers focus on behavioral aspects of CE (Van Doorn et al., 2010, p. 253), other researchers conceptualize it as a "psychological state" (Brodie et al., 2011; Hollebeek, 2011).

Founded on the service-dominant (S-D) logic (Hollebeek, Srivastava & Chen, 2019), CE is gaining recognition in high-involvement and high-value industries that are intangible and experiential in nature, such as hospitality and tourism (Rather et al., 2022; Chen et al., 2021; Hao, 2020; So, Li & Kim, 2020). Along with the practitioners' interest, the scientific acknowledgment of CE is also evident. Recently, three systematic literature reviews on CE in the hospitality and tourism sector (i.e., Hao, 2020; Chen et al., 2021; So, Li & Kim, 2020) provided an in-depth overview of the studies on CE. Besides elucidating the present position and paving the future research on CE, they call for further development of a research outline in the hospitality and tourism sector (So, Li & Kim, 2020). Based on this and having in mind the limitations of the previous literature review studies, the present paper applies a bibliometric and network analysis for quantitative mapping of the academic intellectual structure of CE in the context of hospitality and tourism (2012-2021). Aside from systematic literature review, the bibliometric analysis applies statistical tools for revealing "trends and citations and/or co-citations of a particular theme" (Paul & Criado, 2020, p. 2). Additionally, this paper extends the review to more than a single keyword (customer engagement), including several related keywords (i.e., consumer engagement, brand engagement, etc.), thus comprising a broader review scope. Therefore, this study strives to reveal the research stream of CE in the hospitality and tourism setting, by identifying the authors, journals, and countries that have the highest contribution to this research setting. Further, keyword network co-occurrence analysis is performed, and leading research topics are revealed. By applying these methods, our study attempts to address the following research questions:

RQ1: Which are the most important keywords existing in the most influential articles in this field?

RQ2: Which are the most prominent research themes on CE in hospitality and tourism?

The remainder of the paper is as follows. First, the theoretical background of the CE concept is presented, followed by the presentation of the methodology and data. Further, results and analysis are elaborated, and future research streams are discussed.

2 Theoretical background

The concept of customer engagement appeared in the marketing literature around 2005 (Sawhney, Verona & Prandelli, 2005), and since then the concept has attracted growing research, especially at the end of the first decade of the 21st century (Bowden, 2009; Van Doorn et al., 2010; Brodie et al., 2011).

Indeed, there is an increasing number of published articles on the CE topic (Vivek, Beatty & Morgan, 2012; Wang & Lee, 2020; Ciunova-Shuleska, Palamidovska-Sterjadovska & Bogoevska-Gavrilova, 2022) due to the increased interest in the concept of engagement among both researchers and practitioners. However, there is a lack of consensus about the definition and operationalization of the CE concept, and therefore different conceptualizations of the construct (as a psychological and/or behavioral) (Van Doorn et al., 2010; Brodie et al., 2011; Hollebeek, 2011; Bowden, 2009) exist. Bowden (2009) defined CE as a sequential psychological process that individuals move through to become loyal to the brand. According to Brodie et al. (2011, p. 260), CE is a psychological state that "occurs by virtue of interactive, co-creative customer experiences with a focal agent/object (e.g., a brand) in focal service relationships". According to Van Doorn et al. (2010, p. 254), CE is defined as "a customer's behavioral manifestations that have a brand or firm focus, beyond purchase, resulting from motivational drivers", whereas Hollebeek (2011, p. 790) defined CE as "the level of an individual customer's motivational, brand-related, and context-dependent state of mind characterized by specific levels of cognitive, emotional, and behavioral activity in direct brand interactions."

The review of the literature has shown that the most of the published CE articles mainly study CE with brands no matter whether the brand as the central engagement object is implicitly or explicitly stated in the article, while the customer is the central engagement subject (Hollebeek, 2011; Moliner, Monferrer-Tirado & Estrada-Guillén, 2018; Hollebeek, Glynn & Brodie, 2014; Brodie et al., 2011). Moreover, besides the CE studies that focus on customers as individuals paying for the brand, there are CE studies that focus on consumers as individuals who do not purchase the brand (Sharma et al. 2022; Hollebeek, Glynn & Brodie, 2014; Rather, 2019). However, according to Hollebeek et al. (2022) customer engagement and consumer engagement are highly related concepts. Moreover, some studies focused on CE in a brand community (Sharma et al., 2022; Kumar & Kumar, 2020; Gummerus et al., 2012) linking the CE concept to a virtual (online) brand community.

In the last decade, context-specific exploration of CE has received growing attention from marketing researchers focusing on specific industries such as tourism and hospitality (Rather et al., 2022; Harrigan et al., 2017; So et al., 2016; Dijkmans, Kerkhof & Beukeboom, 2015; So, King & Sparks, 2014; Wei, Miao & Huang, 2013). This led to increased research interest in review studies on CE in the tourism and hospitality field (So, Li & Kim, 2020; Hao, 2020; Opute, Irene & Iwu, 2020; Chen et al., 2021).

So, Li & Kim (2020) reviewed and synthesized CE research in tourism and hospitality published in seven leading hospitality and tourism journals between 2009 and 2019, whereas Hao (2020) provided a systematic review of a total of 173 peer-reviewed articles collected from seven databases and published in the period 2007-2020. Additionally, Opute, Irene & Iwu (2020) used the qualitative approach in systematic reviewing existing literature on CE in tourism, while the research of Chen et al. (2021) employed a systematic bibliometric analysis and content analysis and focused on top-ranked journals from Google Scholar, Science Direct, EBSCOhost, and CNKI databases.

Based on all the above mentioned, in our review study, we analyzed several related terms to CE. Namely, by encompassing a collection of nine CE-related keywords and by conducting a bibliometric analysis on CE research from the tourism and hospitality field generated from the SCOPUS database only, our study differs from others that focused on a single keyword, focused on a review of CE research in general, and relied on conducting qualitative research, and thus enriching the previous CE review literature.

3 Methodology and data

In this study, we followed the PRISMA protocol for systematic reviews (Moher et al., 2009) which consists of four stages: identification, screening, eligibility, and inclusion for a transparent systematic literature review (Lim & Rasul, 2022). In the phase one we found 157 articles by searching the Scopus database on October 16 2021 within the article title (TITLE), abstract (ABS), and keywords (KEY): "customer engagement" OR "consumer engagement" OR "brand engagement" OR "customer brand engagement" OR "consumer brand engagement" OR "customer engagement behavior" OR "customer engagement behaviour" OR "consumer engagement behavior" OR "consumer engagement behaviour" OR "brand community engagement" and using only articles as the document type and the entire period (except 2022). The articles contain the words tourism and/or hospitality (in title, abstract, or keywords). The screening phase cleared the database from duplicates and non-English articles (one duplicate and one non-English article were removed). In the eligibility phase, we checked for editorials, book reviews, and systematic literature reviews, and thus excluded 8 literature review articles to avoid repetition in the results (Lim & Rasul, 2022). In the last phase, we applied content analysis of the abstracts to make the final decision for inclusion and

removed 13 articles. A total set of 23 articles are inadequate after following the PRISMA protocol, and the sample for analysis thus consists of 134 articles. Despite the data visualization on article distribution by year, top journals based on the number of published articles, and the 5 most cited articles, the research hotspots are discovered in the whole analyzed period and in the last five years by using text mining with the most recent version 1.6.18 of the software VOSviewer. Details of the software can be found in Van Eck and Waltman (2010).

4 Results and analysis

4.1 Descriptive analytics

We found that majority of the articles (37) are published in 2021 within the studied period of 2012 to 2021. Furthermore, 88.06 % of articles (118 articles) were published in the recent five years, indicating that researchers' interest in this area is booming (Figure 1).

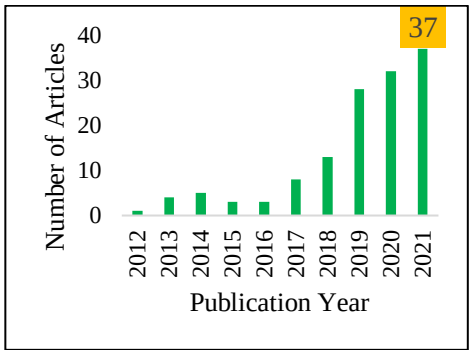


Figure 1. Distribution of articles by year

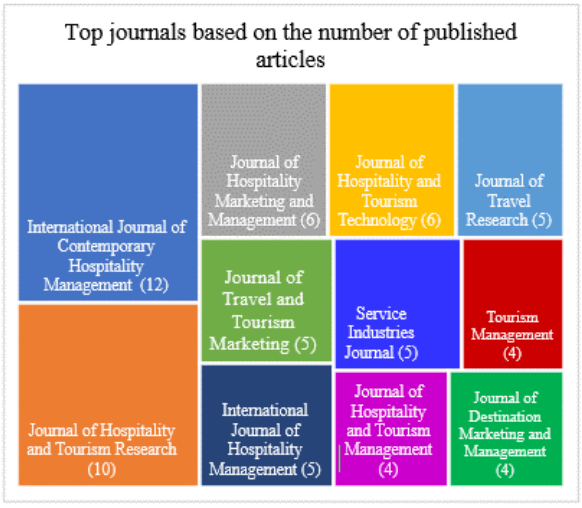


Figure 2. Top journals based on the number of published articles

In Figure 2, we present a tree map with the top journals based on the number of published articles. The International Journal of Contemporary Hospitality Management is the leader with 12 published articles, followed by the Journal of Hospitality and Tourism Research, etc.

In Table 1, the 5 most cited articles are presented. The most cited article is by Harrigan et al. (2017), with 270 citations, followed by the article by So et al. (2016), with 248 citations, etc.

Table 1. The top five most cited articles

Title	Source title	Authors	Year	Cited by
Customer engagement with tourism social media brands	*Tourism Management*	Harrigan P., Evers U., Miles M., Daly T.	2017	270
The Role of Customer Engagement in Building Consumer Loyalty to Tourism Brands	*Journal of Travel Research*	So K.K.F., King C., Sparks B.A., Wang Y.	2016	248
Customer Engagement With Tourism Brands: Scale Development and Validation	*Journal of Hospitality and Tourism Research*	So K.K.F., King C., Sparks B.	2014	247
A stage to engage: Social media use and corporate reputation	*Tourism Management*	Dijkmans C., Kerkhof P., Beukeboom C.J.	2015	231
Customer engagement behaviors and hotel responses	*International Journal of Hospitality Management*	Wei W., Miao L., Huang Z.	2013	149

We created a country co-authorship map based on the bibliometric data. The counting method is full-

counting, the analysis unit is countries, the maximum number of countries per document is set to 10, the minimum number of documents per country is 1, and the number of citations has no condition. Based on the findings, 46 countries have been found, although the largest group of connected countries is 33, which is used for the density visualization map (Figure 3). The weight of a country is determined by the number of documents it has, and we used rainbow colors the countries in the red area have more documents than the others. The countries in the red zone have a greater number of papers compared to others. According to Figure 3, the United States has the most documents (38), followed by China and India (both 14), Australia (13), the United Kingdom (11), etc.

The countries are divided into seven clusters, with the US collaborating closely with Switzerland, Belgium, and Colombia. India is in the cluster with Estonia, Malta, France, and Saudi Arabia, whereas China is in the cluster with Hong Kong, Macau, Pakistan, South Korea, and Taiwan. In addition, the United States has the highest number of citations (1266), followed by Australia (966), the United Arab Emirates (495), the United Kingdom (315), India (310) etc.

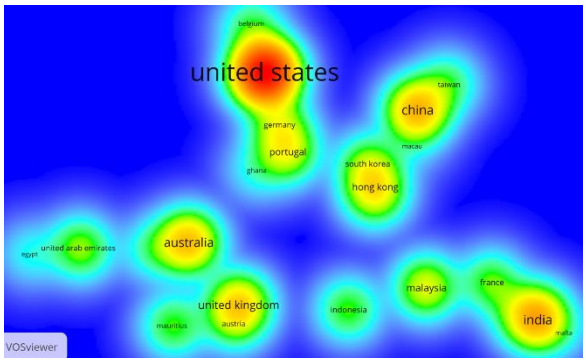


Figure 3. Country density visualization for articles on customer engagement in tourism and hospitality

4.2 Keywords analysis based on text mining in the abstracts

The keyword distribution visualizations built in VOSviewer software based on the keywords extracted from the 134 abstracts are shown and analyzed in this section. First, the focus is on the keywords that appear at least twice in the observed period of 2012-2021, in order to discover research hotspots over the entire period by presenting a network visualization of keyword co-occurrence. The keywords are then analyzed throughout the last three years (2019-2021), with the minimum number of occurrences of the keywords set at two. We show a cluster density visualization of keywords in the last three years for this period.

4.3 Keywords analysis based on text mining in the abstracts over the entire analyzed period

Based on text data from the Scopus database, we develop a keyword co-occurrence network visualization. The terms (keywords) are mined from abstracts, and the abstract structure labels and copyright assertions are not included in the abstract fields. We found 3048 keywords, and we set the minimum number of occurrences at two, therefore 612 fulfilled the conditions. The software's default choice is to select the 60% most relevant keywords based on the relevance score, leaving 367 terms to be verified. We have removed the keywords that have no meaning for the analysis, such as basis, conceptual framework, consequence, current research, direct effect, and so on, from the list of 367 keywords (Appendix 1a). Items and lines constitute the visualization network. Labels and circles are used to display the items. The item's weight is dependent on the size of the label and circle. The label may not be visible for some items, which is done to avoid the problem of overlapping labels (Van Eck & Waltman, 2020). A link is created by drawing a line between two items. The stronger the relationship between two items, the thicker the line between them and the shorter the distance between them (Cvetkoska & Savic, 2021). The quantity of distinct colors indicates how many different clusters there are. Items that are the same color belong to the same cluster. Figure 4 shows a visualization of the keyword co-occurrence network from 2012 to 2021. The total number of items in this network is 290, with 2,645 links and a link strength of 2,805. We can observe that the items are grouped into 11 clusters based on the different colors. The largest label and circle are displayed for the keyword hospitality management, which belongs to the orange cluster, indicating that it has the highest weight (based on occurrences), followed by engagement behavior, which also belongs to the orange cluster. The keyword hospitality management has 18 occurrences with 95 links, giving it a total link strength of 119. The keyword engagement behavior has ten occurrences with 61 links, and the total link strength is 79.

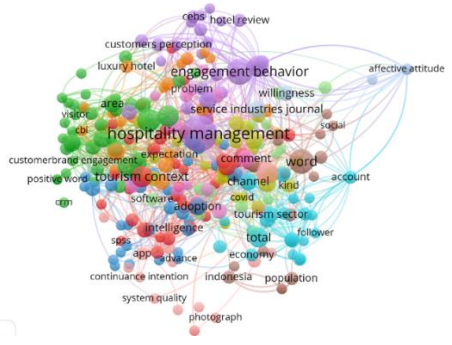


Figure 4. Keyword co-occurrence network visualization over the entire analyzed period

4.4 Keywords analysis based on text mining in the abstracts in the last three years (2019-2021)

There have been 97 papers published in the last three years whose abstracts have been used to identify research hotspots using text mining. We detected 2200 keywords using the VOSviewer software, and we set the minimum number of occurrences of a keyword to 2, so 434 meet this criterion. The software then chooses the 60% most relevant of these by default, leaving 260 keywords. We also removed 53 non-relevant keywords, such as article, base, contrary, direct effect, direct influence, direct relationship, effective implementation, empirical proof, empirical outcome, empirical study, end, and so on (Appendix 1b), leaving 207 keywords for the cluster density visualization (Figure 5). The item satisfaction that belongs to the pink cluster has the highest weight (11 occurrences, 54 links, and the total link strength is 75). In addition, we have found that research hotspots in 2019 are linked with: satisfaction, customer trust, customer loyalty, customer engagement strategy, emotional connection, relationship market strategy, confirmatory factor analysis, etc., while in 2020 the researchers most focus on the hospitality industry, customer brand engagement, engagement behavior, customer experience, destination brand engagement, behavioral engagement. In the last analyzed year, researchers' focus was on COVID, value cocreation, Instagram, celebrity effect, servicescape, and customer retention, and the used software is AMOS for structural equation modeling.

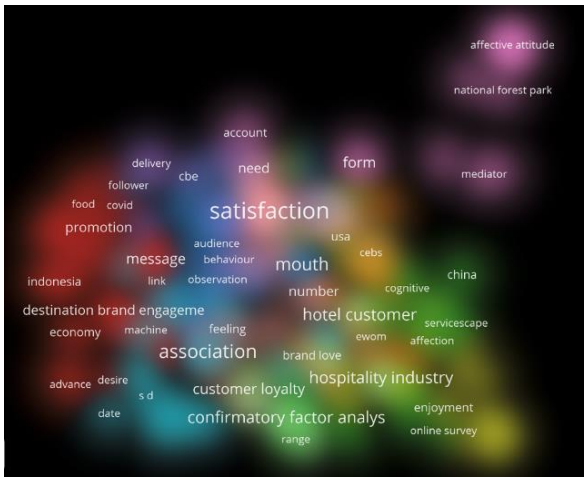


Figure 5. Overlay density cluster visualization in the last 3 years (2019-2021)

5 Conclusion

The article's review of CE research in tourism and hospitality outlined several takeaways that provide important implications for future CE research in this field. *First*, the interest in researching the CE in the tourism and hospitality field has been continually growing from 1 article in 2012 to 37 published articles in 2021. The highest increase in the number of published articles can be noticed in 2019 compared to 2018, with 88.06% of all articles being published in the period from 2017 to 2021 which indicates that SCOPUS journals have been highly responsive to CE research studies in tourism and hospitality in the last 5 years. Second, the top 5 most influencing authors Harrigan et al. (2017), So et al., (2016), So, King & Sparks (2014), Dijkmans, Kerkhof & Beukeboom (2015), and Wei, Miao & Huang (2013) published their research studies in Tourism Management, Journal of Travel Research, Journal of Hospitality and Tourism Research and International Journal of Hospitality, with Tourism and Management being the journal where two of the top five articles were published. These five most cited studies amounted to a total of 1145 citations with an average of 229 citations per article indicating that prospective authors should be aware of these articles and use them as a basis when making the theoretical foundation of their future CE research. *Third*, authors who worked on the CE topic in tourism and hospitality come from 46 countries with the United States being the country with the most published documents. Regarding the collaboration, and based on the identified clusters, it can be noticed that there is a strong collaboration among researchers from several Asian countries, whereas the other researchers collaborate with colleagues from countries that do not belong to the same region and are located far from each other. *Fourth*, the most often occurred keywords in the research studies published in the period 2012-2021 are hospitality management and engagement behavior indicating that the researchers predominantly analyzed the CE topic from a behavioral perspective. *Fifth*, the keyword satisfaction is the most often occurred keyword in the research studies from the last three years (2019-2021) meaning that the latest CE research mostly linked CE to satisfaction as an important marketing indicator, leading to the conclusion that the satisfaction will be the major research hotspots in CE studies in tourism and hospitality.

Customer satisfaction is strongly related to CE and both could be improved by enhancing customer experience and motivating customer co-creation in hospitality management. Providing a personalized customer experience in tourism is expected to enrich the customer value, as well as the value for tourist providers (Mohammadi et al., 2021). Nowadays, the active role of customers in creating their experience is enabled by using the Internet and especially through social media.

So, based on the extracted theoretical highlights, future academic and managerial efforts in tourism and hospitality should be put into stimulating customer co-creation as a superior form of CE (Chathoth et al., 2014). This in turn will provide enhanced customer

experience and customer satisfaction, as well as customer retention and loyalty (Rather et al., 2022).

Though this study contributes to paving the CE research in tourism and hospitality, it is limited in several ways. First, Scopus is used for retrieving the 134 English journal articles, so this study cannot exclude the possibility of omitting some articles published in journals that are indexed in other databases. Future studies may also consider other databases, as well as non-English articles and conference papers, book reviews, doctoral and master thesis, etc. Second, although this study applies text mining in terms of keywords analysis, the future study may use topic modeling to reveal the hidden research themes. At last, although the present study thoroughly reviewed the past and present CE studies in the context of tourism and hospitality, continuous mapping of the scientific knowledge will be needed in the following years, due to the dynamic nature of this field.

References

Bowden, J.L.H., (2009). The Process of Customer Engagement: A Conceptual Framework. *Journal of Marketing Theory and Practice*, 17(1), 63-74.

Brodie, R.J., Hollebeek, L.D., Jurić, B. & Ilić, A., (2011). Customer Engagement: Conceptual Domain, Fundamental Propositions, And Implications for Research. *Journal of service research*, 14(3), 252-271.

Chathoth, P. K., Ungson, G. R., Altinay, L., Chan, E. S., Harrington, R., & Okumus, F. (2014). Barriers Affecting Organisational Adoption of Higher Order Customer Engagement in Tourism Service Interactions. *Tourism Management*, 42, 181-193.

Chen, S., Han, X., Bilgihan, A. & Okumus, F., (2021). Customer Engagement Research in Hospitality And Tourism: A Systematic Review. *Journal of Hospitality Marketing & Management*, 30(7), 871-904.

Ciunova-Shuleska, A., Palamidovska-Sterjadovska, N. & Bogoevska-Gavrilova, I. (2022). What Drives Liking Different Brand-Related Social Media Content?. *Marketing Intelligence & Planning*, 40(4), 542-556.

Cvetkoska, V. & Savic, G. (2021). DEA In Banking: Analysis and Visualization of Bibliometric Data. *Data Envelopment Analysis Journal*, 5(2), 455–485.

Dijkmans, C., Kerkhof, P. & Beukeboom, C.J., (2015). A Stage to Engage: Social Media Use and Corporate Reputation. *Tourism management*, 47, 58-67.

Gummerus, J., Liljander, V., Weman, E. & Pihlström, M. (2012). Customer Engagement in A Facebook Brand Community. *Management Research Review*, 35(9), 857-877.

Hao, F. (2020). The Landscape of Customer Engagement In Hospitality And Tourism: A Systematic Review. *International Journal of Contemporary Hospitality Management*, 32(5), 1837-1860.

Harrigan, P., Evers, U., Miles, M. & Daly, T. (2017). Customer Engagement with Tourism Social Media Brands. *Tourism Management*, 59, 597-609.

Hollebeek, L., (2011). Exploring Customer Brand Engagement: Definition and Themes. *Journal of Strategic Marketing*, 19(7), 555-573.

Hollebeek, L.D., Glynn, M.S. & Brodie, R.J., (2014). Consumer Brand Engagement in Social Media: Conceptualization, Scale Development And Validation. *Journal of Interactive Marketing*, 28(2), 149-165.

Hollebeek, L.D., Sharma, T.G., Pandey, R., Sanyal, P. & Clark, M.K. (2022). Fifteen Years of Customer Engagement Research: A Bibliometric and Network Analysis. *Journal of Product & Brand Management*, 31(2), 293-309.

Hollebeek, L.D., Srivastava, R.K. & Chen, T. (2019). SD Logic–Informed Customer Engagement: Integrative Framework, Revised Fundamental Propositions, And Application To CRM. *Journal of the Academy of Marketing Science*, 47(1), 161-185.

Kumar, J. & Kumar, V. (2020). Drivers Of Brand Community Engagement. *Journal of Retailing and Consumer Services*, 54, 101949.

Lim, W.M. & Rasul, T. (2022). Customer Engagement and Social Media: Revisiting The Past To Inform The Future. *Journal of Business Research*, 148, 325-342.

Mohammadi, F., Yazdani, H.R., Jami Pour, M. & Soltani, M. (2021). Co-creation in tourism: a systematic mapping study. *Tourism Review*, 76 (2), 305-343.

Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine*, 151(4), 264–269.

Moliner, M.Á., Monferrer-Tirado, D. & Estrada-Guillén, M. (2018). Consequences Of Customer Engagement and Customer Self-Brand Connection. *Journal of Services Marketing*, 32(4), 387-399.

MSI - Marketing Science Institute (2010). *Research Priorities 2010-2012*. Retrieved from: http://image.sciencenet.cn/olddata/kexue.com.cn/upload/blog/file/2010/9/201091515178616316.pdf.

MSI - Marketing Science Institute (2020). *Research Priorities 2020-2022*. Retrieved from: https://www.msi.org/wp-content/uploads/2020/06/MSI_RP20-22.pdf.

Opute, A.P., Irene, B.N. & Iwu, C.G. (2020). Tourism Service and Digital Technologies: A Value Creation Perspective. *African Journal of Hospitality, Tourism and Leisure*, 9(2), 1-18.

Paul, J., & Criado, A. R. (2020). The Art Of Writing Literature Review: What Do We Know And What Do We Need To Know?. *International Business Review*, 29(4), 101717.

Rather, R.A. (2019). Consequences Of Consumer Engagement in Service Marketing: An Empirical Exploration. *Journal of Global Marketing*, 32(2), 116-135.

Sawhney, M., Verona, G. & Prandelli, E. (2005). Collaborating To Create: The Internet as A Platform for Customer Engagement In Product Innovation. *Journal of interactive marketing*, 19(4), 4-17.

Sharma, P., Sadh, A., Billore, A. & Motiani, M. (2022). Investigating Brand Community Engagement and Evangelistic Tendencies on Social Media. *Journal of Product & Brand Management*, 31(1), 16-28.

So, K., King, C., & Sparks, B. (2014). Customer Engagement with Tourism Brands: Scale Development and Validation. *Journal of Hospitality & Tourism Research*, 38(3), 304-329.

So, K., Li, X. & Kim, H. (2020). A Decade of Customer Engagement Research in Hospitality and Tourism: A Systematic Review and Research Agenda. *Journal of Hospitality & Tourism Research*, 44(2), 178-200.

So, K.K.F., King, C., Sparks, B.A. & Wang, Y. (2016). The Role of Customer Engagement in Building Consumer Loyalty to Tourism Brands. *Journal of Travel Research*, 55(1), 64-78.

Syrdal, H.A. & Briggs, E. (2018). Engagement With Social Media Content: A Qualitative Exploration. *Journal of Marketing Theory and Practice*, 26(1-2), 4-22.

Tsiotsou, R.H. (2020). Social Media and Customer Engagement. In Bridges, E. & Fowler, K. (Eds.), *The Routledge Handbook of Service Research Insights and Ideas* (pp. 373-387). Routledge, Bisbaine.

Van Doorn, J., Lemon, K., Mittal, V., Nass, S., Pick, D., Pirner, P. & Verhoef, P. (2010). Customer Engagement Behavior: Theoretical Foundations and Research Directions. *Journal of Service Research*, 13(3), 253–266.

Van Eck, N.J. & Waltman, L. (2010). Software Survey: VOSviewer, a Computer Program for Bibliometric Mapping. *Scientometrics*, 84(2), 523-538.

Van Eck, N.J. & Waltman, L. (2020). *Manual for VOSviewer Version 1.6.15*. Leiden: Centre for Science and Technology Studies (CWTS) of Leiden University (2020).

Vivek, S.D., Beatty, S.E. & Morgan, R.M. (2012). Customer Engagement: Exploring Customer Relationships Beyond Purchase. *Journal of marketing theory and practice*, 20(2), 122-146.

Wang, T. & Lee, F.Y. (2020). Examining Customer Engagement and Brand Intimacy In Social Media Context. *Journal of Retailing and Consumer Services*, 54, 102035.

Rather, R. A., Hollebeek, L. D., & Rasoolimanesh, S. M. (2022). First-time versus Repeat Tourism Customer Engagement, Experience, and Value Cocreation: An Empirical Investigation. *Journal of Travel Research*, 61(3), 549-564.

Wei, W., Miao, L. & Huang, Z.J. (2013). Customer Engagement Behaviors and Hotel Responses. *International Journal of Hospitality Management*, 33, 316-330.

Appendix 1

a) Excluded terms for the keyword co-occurrence network visualization over the entire analyzed period

2013international journal, 2015international journal, 2017international journal, 2018international journal, 2019international journal, 2020current issue, 2020international journal, academic research, basis, conceptual framework, consequence, current research, direct effect, direct influence, direct relationship, ease, effective implementation, empirical analysis, empirical evidence, empirical result, empirical study, end, essential element, establishment, evidence, example, extant literature, findings highlight, first study, first time, flow, further research, further understanding, future research direction, higher level, hospitality literature, identification, iii, important driver, important implication, indirect effect, influential impact, item, key implication, large number, limited attention, limited research, literature review, main purpose, mediating effect, moderating effect, new insight, novelty, object theme, positive attitude, positive relationship, present research, present study, previous research approach, prior research, recent year, referral, research finding, research framework, research gap, share, significance, significant antecedent, significant difference,

significant impact, significant positive effect, significant role, subject, theoretical, theoretical framework, third

b) Excluded terms for the keyword co-occurrence network visualization for the last three years (2019-2021)

article, basis, contrary, direct effect, direct influence, direct relationship, effective implementation, empirical evidence, empirical result, empirical study, end, evidence, example, extent, first study, first time, flow, further research, higher level, iii, important implication, important role, indirect effect, influential, impact, item, key implication, limitation, limited attention, mediating effect, medium, moderating effect, new insight, novelty, object, positive impact, positive influence, positive relationship, present study, previous research approach, recent year, referral, research finding, research gap, sense, significance, significant antecedent, significant difference, significant impact, significant positive effect

Adverb and Preposition Representation in Croatian and English using the Node of Knowledge Method

Martina Asenbrener Katic, Sanja Candrlic
Faculty of Informatics and Digital Technologies
University of Rijeka
Radmile Matejcic 2, 51000 Rijeka
{masenbrener, sanjac}@inf.uniri.hr

Marina Rauker Koch
Polytechnic of Rijeka
Vukovarska 58, 51000 Rijeka
mrauker@veleri.hr

Abstract. *A conceptual framework Node of knowledge enables, among others, presentation of knowledge in a formalized record suitable for insert into a relational database. To correctly transform all words from natural language sentences to formalized records, it is necessary to design a metamodel of a language and define rules for transformation of each sentence part. This paper focuses on adverbs and prepositions and their representation in Croatian and English language. It presents rules and examples for their transformation into formalized records in both languages and discusses results of preliminary testing of the system with a small set of sentences and questions.*

Keywords. Knowledge representation, Node of Knowledge, NOK, Adverb, Preposition

1 Introduction

In the field of artificial intelligence and knowledge representation, many methods have been developed to record knowledge. Some of these methods are semantic networks (Brachman & Levesque, 2004), (Gomez-Perez & Benjamins, 2002), frames (Liao et al., 2012), ontologies (Cali et al., 2012), Fuzzy Petri nets (H.-C. Liu et al., 2017), neural networks (Neelakantan, 2017), the Node of Knowledge method (NOK) (Jakupovic et al., 2014), (Pavlic et al., 2013) or other graphical methods for knowledge representation (Jakupovic et al., 2013).

The Node of Knowledge method (NOK) is an integral part of the conceptual framework "Node of Knowledge (NOK)", which, in addition to the NOK method, consists of a formalism for graphical representation (Diagram Node of Knowledge, DNOK), a formalism for representing knowledge in textual form (Formalized Node of Knowledge, FNOK), and a formalism for representing questions in textual form (Question Formalized Node of Knowledge, QFNOK) (Candrlic et al., 2019).

The NOK conceptual framework enables the entry of knowledge, texts expressed in sentences of natural written language, into the Question answering (QA) system in such a way that the knowledge expressed in sentences of natural language is formatted into a record suitable for input into a relational database. This database is an integral part of the QA system. Once the knowledge is written into the QA system, it enables the retrieval and questioning of that knowledge, i.e., searching for and answering questions expressed in natural language sentences.

Previous research by the authors (Rauker Koch et al., 2017), (Rauker Koch et al., 2015), (Rauker Koch et al., 2014), (Candrlic et al., 2020) has shown that using the NOK conceptual framework, it is possible to write different texts in different languages in a formalized format suitable for input to a relational database. In parallel with the human transformation of natural language into FNOK record, an application for automatic transformation of natural language sentences into FNOK record was developed (Dovedan Han, 2021) and the results were compared with the transformation performed manually (Rauker Koch et al., 2022). To convert natural language sentences into a formalized record suitable for storage in a relational database, it is necessary to analyse each word type and define the rules for formalization or create a metamodel of the language. Previous research has analysed nouns (Asenbrener Katic et al., 2021), verbs (Asenbrener Katic et al., 2018), and adjectives (Pavlic et al., 2017). The question is whether and how adverbs and prepositions can be written in a formalized record.

This paper is primarily focused on rules for manual transformation of adverbs and prepositions from natural language sentences in Croatian and English into FNOK record. The tests were performed with sentences containing these word types.

The paper is organized as follows. After the introduction in Section 1 and the presentation of related work in Section 2, Section 3 presents the motivation for the research. Section 4 describes the research methodology. In Section 5, authors analyse adverbs and prepositions in Croatian and English and

give examples and rules for their transformation from natural language to formalized records.

2 Related work

In (Timko, 2018) adverbs are described as a specific and important word type, in terms of their structure, morphological, syntactic and semantic properties. Many works deal with the analysis of adverbs and prepositions from different aspects.

Socher (Socher et al., 2012) presented a recursive neural network (RNN) model that learns compositional vector representations for phrases and sentences of arbitrary syntactic type and length. This model assigns a vector and a matrix to every node in a parse tree: the vector captures the inherent meaning of the constituent, while the matrix captures how it changes the meaning of neighbouring words or phrases. Each word and phrase is represented by a vector and a matrix. Adverbs are studied within adverb-adjective pairs, while NOK method allows representation of all combinations of adverbs and other types of words.

Pollin (Pollin et al., 2018) describes the annotation process and model of the Open Access Database 'Adjective-Adverb Interfaces in Romance' (AAIF) project. Their approach is to create a domain-specific ontology and the data is annotated in TEI/XML, formalized and enriched with RDF, and its conceptual data model is stored in and published via the GAMS digital repository. In (Gerhalter, 2021), only adjective - adverbs - i.e. adjectives with adverbial functions - are tagged and annotated, while adjectives with adjectival syntax are not tagged in the AAIF-database.

Prepositions play an important role in syntax and semantics of sentences in Croatian and English. This leads to difficulties in automatic processing of sentences. According to (Pawar et al., 2021), prepositions have been explored using various classical machine learning techniques, for example: they have used heavy linguistic tools and resources such as part-of-speech taggers, chunkers, dependency parsers, named entity extractors, Word-Net based super-sense taggers, and semantic role labelers for feature engineering. Implementations of such approaches can be evaluated with varying degrees of success.

Recent studies aimed at solving prepositional attachment and prepositional selection problems rely heavily on external linguistic resources and use dataset-specific word representations (Gong et al., 2018). In (Gong et al., 2018), word-triples are counted (where one of the triples is a preposition) to capture the interaction of a preposition with its attachment and complement.

Prepositions have been studied, among others, as part of the BERT language representation model (Devlin et al., 2019) and its derivatives such as RoBERTA (Y. Liu et al., 2019) and Big Bird (Zaheer et al., 2020), which were developed to pre-train contextualized deep representations of words based on large unlabeled corpora (Wiedemann et al., 2019).

In our approach using FNOK, each word in a sentence is examined and its type is determined using a dictionary. The rules for translation into a formalized notation apply to all words in a sentence. First, the type of each word in the sentence is determined. Then, depending on the word type of the word we associate it with, a particular rule is applied. The author's previous research has shown that the system is very accurate in answering questions. For example, in (Asenbrener Katic et al., 2021) it is shown that the prototype provided 93% of correct answers in English and 94% of correct answers in Croatian. The prototype was tested with 30 sentences and 55 questions in English and 30 sentences and 54 questions in Croatian. These sentences contain nouns and noun phrase structures.

The rules for adverbs and prepositions, as well as the results of the preliminary research, are presented in this paper.

3 Research Motivation

Knowledge is described in (Hrvatska enciklopedija, 2021) as a system and logical outline of facts and generalizations about objective reality that are accepted and permanently retained in consciousness; a set of facts, information, and skills acquired through training or experience for theoretical or practical understanding and problem solving. People express their knowledge in natural languages in various ways, for example, in writing, orally, using a figure, etc.

This research aims to store knowledge expressed in texts in the knowledge base so that it can give answers to the questions asked, based on stored knowledge. In other words, the main motivation is to develop a question-answering (QA) system over a relational database that allows to input any text or knowledge expressed in text in any language. To achieve this, it is necessary to study how to write text expressed in natural language into a relational database. In addition, it is necessary to perform an analysis of each type of word in a given language and to define the rules by which it is possible to enter text or sentences in natural language into a relational database. It is also necessary to define a metamodel for each word type, which will later form the basis of the relational knowledge base metamodel. We are currently limited to Croatian and English, but it is possible to do the same for other languages.

Since analyses of nouns, adjectives and verbs have been made so far, further analysis is made for the closed word classes adverbs and prepositions. It included the study on how they could be written in a formalized form (FNOK record) which has so far proved suitable for entry into a relational knowledge base on which interrogation could be conducted.

4 Methodology

The basic concepts of the NOK method are nodes (which can be ordinary or process) and links with a single role. Two ordinary nodes, two process nodes or one ordinary and one process node are connected with a link. Each link between the nodes has only one role. The link between two nodes is accomplished by asking a question related to the first node, and the response results in another node (which is on the other side of the link (Asenbrener Katic et al., 2021). By using the NOK method, FNOK is created.

According to the general rules described in (Pavlic et al., 2013), (Asenbrener Katic et al., 2021), (Asenbrener Katic et al., 2018), can be briefly summarised:

- A word in a sentence in natural language corresponds to a node in the FNOK record. More precisely, the verb in a sentence in natural language corresponds to a process node in FNOK. Every other word type corresponds to an ordinary node (or simply: node).
- Transformation of a sentence in natural language into its formalized FNOK record is preceded by analysis of that sentence. The first step is to find verbs in it, because verbs are most common process nodes and they belong to the highest level of hierarchy. Other nodes belong to lower levels of hierarchy.
- Open bracket marks descending to lower levels of hierarchy, while closed bracket marks return to the previous level.
- Question or a question word is in front of each node in the hierarchy and indicates a role that links a node on a lower level with a node on a higher level, i.e. node on a lower level is an answer to the asked question.
- Question (Role) is marked by quotation marks and a question mark ("?"). Some of the questions that can be asked are: "who?", "what?", "when?", "how?", "where?", "how much?", "how many?" etc., depending on the word type. Questions asked with roles are usually from the group of wh-questions.
- General FNOK representation: *process node ("role 1?" node 1, "role 2?" node 2 ("role 3?" node 3 ("role 4?" node 4, ...), ...), "role n?" node n)* where *role 1?* to *role n?* are wh-questions, *node 1* to *node n* are nodes, and process node is the process node.
- Each node can have its own hierarchy.

5 Research Results

The rules for converting adverbs and prepositions into FNOK records are presented in this chapter.

5.1. Adverbs

In Croatian, adverbs are a closed word classes that are most often added to verbs to express the circumstances of an action (place, time, manner, cause, purpose, and quantity or amount). Adverbs can also be added to adjectives (e.g. pretty good), adverbs (e.g. a little earlier) and nouns (e.g. many leaflets) (Tezak & Babic, 2005).

Table 1. gives a description of the types of adverbs (Tezak & Babic, 2005), (Silic & Pranjkovic, 2005) in Croatian with an explanation in English.

Table 1. Description of the types of adverbs in Croatian with an explanation in English.

Types of adverbs in Croatian		Croatian	English
Vremenski prilozi (Adverbs of time)	Kada? (When?)	danas, večeras, noćas, jučer, sinoć, preksinoć, danas, sutra, preksutra, ljetos, proljetos, jesenas, zimus, proljeti, ljeti, jeseni, zimi, lani, preklani, obdan, obnoć, odmah, smjesta, sada, tada, onda, ikada, bilo kada, nikada, nekada, ponekad, katkad, uvijek, svagda, često, rijetko, rano, kasno, prije, poslije, potom, nedavno, skoro, uskoro, napokon, ...)	today, tonight, yesterday, last night, the day before yesterday, today, tomorrow, the day after tomorrow, last summer, last spring, last autumn, last winter, in spring, in summer, in autumn, in winter, last year, two years ago, during the day, during the night, immediately, promptly, now, then, ever, anytime, never, once, sometimes, always, often, rarely, early, late, before, after, recently, almost, soon, finally, …
	Otkada? (Since when?)	odsad, otad, oduvijek, odavna, odmalena…	from now on, since then, always, long ago, from an early age, …
	Dokada? (Until when?)	dosad, dotad, dogodine, dovečer ...	so far, until then, until next year, until tonight …
Mjesni prilozi (Adverbs of place)	Gdje? (Where?)	ovdje, tu, ondje, negdje, igdje, nigdje, ponegdje, gore, dolje, unutra, vani, sprijeda, straga, ...	here, there, somewhere, anywhere, nowhere, up, down, inside, outside, in front, behind, ...
	Kamo? (Where to?)	ovamo, onamo, tamo, nekamo, nikamo, ikamo, naprijed, natrag …	here, there, somewhere, nowhere, forward, backward
	Kuda? (Where?, Which way?)	ovuda, onuda, tuda, nikuda, nekuda, ikuda, kojekuda ...	This way, that way, nowhere, somewhere, anywhere, …

	Otkuda? Odakle? (Where from?)	odavde, otud, odatle, odonud, niotkuda, odozgo, odozdo, odostraga, izdaleka, izvana, izbliza ...	from here, from there, from nowhere, from above, from below, from behind, from afar, from outside, up close ...
	Dokle? Dokud? (Until where?)	donekle, dovle, dotle, donle	to some extent, to that extent
Načinski prilozi (Adverbs of manner)	Kako?, Kojim načinom? (How?)	ovako, tako, onako, nikako, nekako, ikako, kojekako, svakojako, kriomice, potajice, poimence, iznenada, odjednom, jedva, svejedno, ...	like this, like that, no way, somehow, secretly, namely, suddenly, barely, anyway
	Adverbs of manner in Croatian are also derived from adjectives	From descriptive adjectives in any degree (three states) of comparison: positive, comparative and superlative: blago, blaže, najblaže, brzo, daleko, dalje, jako, lako, lijepo ,teško, teže, najteže, tiho, veselo, zlo, žestoko, … (for example, Dijete se igra (kako?) veselo.)	mild, milder, mildest, fast, far, further, strong, easy, beautiful, hard, harder, hardest, quiet, merry, evil, fierce, … (for example, Child plays (how?) merry).
		From relative adjectives (finishing with –ski in Croatian): bratski, gospodarski, hrvatski, ljudski, slavenski, školski, vojnički, etc. (for example, Govorimo (kako?) hrvatski.)	fraternal, economic, Croatian, human, Slavic, etc. (for example, We speak (how?) Croatian.)
Uzročni prilozi (Adverbs of cause)	Zašto?, Zbog čega? (Why?)	zato, stoga	because
Količinski prilozi (Adverbs of quantity)	Koliko? (How much? How many?)	ovoliko, toliko, onoliko, nekoliko, malo, premalo, više, previše, prekoviše, najviše, ponajviše, manje, najmanje, ponajmanje, dosta, odveć, opet, još, sasvim, potpuno, previše	so much, so many, a few, a little, too little, more, too much, too many, less, least, enough, again, still, quite, completely
	Koliko puta? (How many times?)	jedanput, dvaput, triput, stoput	Once, twice, three times, a hundred times
Posljedični prilozi (Adverbs of consequence)	S kojom posljedicom? S kojim ishodom? (With what consequence? With what outcome?)	uzalud, uzaman, utaman	in vain

There are a number of adverbs that do not answer the above questions. They are attached not only to verbs, but also to other types of words and even to the whole sentence, and they state:

- Reliability, truthfulness of what the article refers to: *naravno, pouzdano, neosporno, neprijeporno, nesumnjivo (of course, reliably, indisputably, unquestionably, undoubtedly)*
- More or less likely, the probability of what it refers to: *sigurno, zasigurno, jamačno, vjerojatno (certainly, surely, guaranteed, probably)*
- Doubt of what the article refers to: *navodno, tobože, naizgled (allegedly, supposedly, apparently)*
- Unexpectedness of what the adverb refers to: *ipak (however)*
- Limitation of what the adverb refers to: *jedan, jedino (one, only)*
- …

In English, adverbs are the words most often attached to verbs, indicating the circumstances of the verb's action (place, time, manner, quantity, and cause). The adverb comes after the verb it modifies (Bozic, 2006).

Types of adverbs are (Bozic, 2006):

1. adverbs of place: *here, there, down,…*
2. adverbs of time: *yesterday, today, now, tonight, often, always…*
3. adverbs of manner: *softly, slowly, quickly, carefully, …*
4. adverbs of frequency/amount: *enough, quite, very, too much, too little, …*
5. adverbs of purpose: *so, since, thus, because, …*

Sometimes adverbs can modify adjectives to strengthen or weaken them. Also, some types of adverbs can modify other adverbs by changing their degree or accuracy. Adverbs are converted to FNOK in two ways.

5.1.1. Adverbs attached to verbs

Since when analysing sentences, the first step is to look for verbs that have the role of a process node in FNOK and are at the top of the hierarchy, the question of the relationship between adverbs and verbs arises. The following rule is defined.

Rule 1. If adverbs are attached to verbs, then in FNOK they are hierarchically dependent on verbs, i.e. they are subordinate to verbs in the hierarchy. It can briefly be written like this:

V, Adv → V ("role?" Adv), where *Adv* stands for adverb, *V* stands for verb and *role?* stands for a wh-question.

Note: When more than one adverb is associated with the verb in a sentence, they are analyzed individually, each in relation to the verb. The order of adverbs and corresponding questions in FNOK is taken from the sentence:

V, Adv 1,… , Adv n → V ("role 1?" Adv 1, …, "role n?" Adv n) where *Adv 1, …, Adv n* stand for different adverbs, *V* stands for verb and *role 1?, … role n?* stand for different wh-questions.

Examples are given first in Croatian, then the same sentence in English:

- Medvjedi jedu *pohlepno*. → jedu ("tko?" medvjedi, *"kako?" pohlepno*)
- The bears eat *greedily*. → eat ("who?" bears ("art?" the), *"how?" greedily*)
- Marko *danas brzo* vozi auto. → vozi ("tko?" Marko, "što?" auto, *"kada?" danas*, *"kako?" brzo*)
- Mark drives a car *quickly today*. → drives ("who?" Mark, "what?" car ("art?" a), *"when?" today*, *"how?" quickly*)

5.1.2. Adverbs attached to other word types (other than verbs)

When adverbs are attached to adjectives or other adverbs, they are translated in FNOK as hierarchically dependent on the adjectives, nouns, or adverbs they modify. The rule is:

Rule 2. Adverbs attached to other word types (adjectives or other adverbs, but not verbs), are translated in FNOK as hierarchically dependent on the adjectives, nouns, or adverbs they modify. It can be written in this form:

Adv, X → X ("role?" Adv), where *Adv* stands for adverb, *X* stands for adjective, noun or adverb modified by the adverb, and *role?* stands for a wh-question.

More complex examples can be reduced to this basic rule.

Example:

- Ti izgledaš *jako dobro*. → izgledaš ("tko?" ti, *"kako?" dobro* (*"kako?" jako*))
- You look absolutely fabulous! → look ("who?" you, "how?" fabulous ("how?" absolutely))
- Marija svira violinu *izuzetno dobro*. → ("tko?" Marija, "što?" violinu, *"kako?" dobro ("kako?" izuzetno)*)
- Mary plays the violin *extremely well*. → plays ("who?" Mary, "what?" violin ("art?" the), *"how?" well ("how?" extremely)*)
- Marko danas *vrlo brzo* vozi auto. → vozi ("tko?" Marko, "što?" auto, "kada?" danas, *"kako?" brzo ("koliko?" vrlo)*)
- Mark drives a car *very quickly* today. → drives ("who?" Mark, "what?" car ("art?" a), "when?" today, *"how?" quickly ("how?" very)*)

5.2. Prepositions

In Croatian, prepositions are a closed word classes used to indicate the relations between beings, things, and phenomena, and they affect the case of the word they stand next to. They are most often associated with nouns and noun pronouns but can also refer to open word classes (adjectives, adjective pronouns and numbers) (Tezak & Babic, 2005), (Silic & Pranjkovic, 2005). Prepositions cannot stand alone in sentences.

In English, prepositions are a type of words that indicate the relationships between objects, beings, and phenomena. They cannot stand alone and are always connected with nouns (Biblioteka Strani Jezici, 2009). Prepositions are combined with other types of words to form prepositional phrases, which usually have a similar form: a preposition followed by determiners, and one or two adjectives followed by nouns or pronouns. This whole phrase has a modifying role, acting as an adjective or adverb that locates something in time and space, modifies a noun, or tells when, where, or under what conditions something happened (Greenbaum & Quirk, 2006).

The following rule is defined for the conversion of prepositions into FNOK.

Rule 3. In FNOK prepositions are hierarchically positioned below nouns, i.e. below the part of speech with which they are associated (noun, noun pronoun or open word classes (adjectives, adjective pronouns and numbers)). Each preposition belongs to the noun to which it refers. The noun is mentioned first, followed by the preposition. The question (role) for the preposition is taken from the noun, which is hierarchically superior to it. Briefly, the rule is:

X, Prep, N → X ("role1?" N ("role1?" Prep), where *Prep* stands for preposition, *X* stands for any kind of word and *role1?* stands for a wh-question.

Examples are given first in Croatian, then the same sentence in English:

- Knjiga je *na* stolu. → je ("što?" knjiga, *"gdje?"* stolu *("gdje?" na)*)

- The book is *on* the table. → is ("what?" book ("art?" the), *"where?"* table (*"where?" on*, "art?" the))
- Ja radim svoju domaću zadaću svaki dan *u* pet sati. → radim ("tko?" ja, "što?" zadaću ("čiju?" svoju, "kakvu?" domaću), "kada?" dan ("koji?" svaki), *"koliko?"* sati *("koliko?" u*, "koliko?" pet))
- I do my homework every day *at* five o'clock. → do ("who?" I, "what?" homework ("whose?" my), "when?" day ("what?" every), "when?" o'clock (*"when?"* five *("when?" at)*))
- Jack je *u* kući. → je ("tko?" Jack, *"gdje?"* kući (*"gdje?" u*))
- Jack is *in* the house. → is ("who?" Jack, *"where?"* house (*"where?" in*, "art?" the))

5.3. ER diagrams for Adverbs and Prepositions

Figure 1. and Figure 2. show entity-relationship (ER) diagrams for adverbs and prepositions, respectively. The ER diagrams are presented using the entity-relationship method in accordance with the Methodology for development of information systems (MIRIS).

In creating the ER diagram for adverbs, rules 1 and 2 were used, as well as general knowledge about adverbs in Croatian and English.

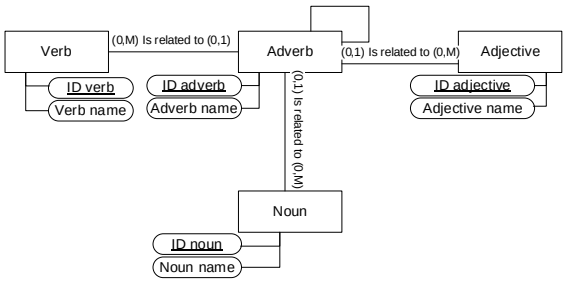


Figure 1. ER diagram for adverbs

In building the ER diagram for prepositions, rule 3 and general knowledge about prepositions in Croatian and English were used.

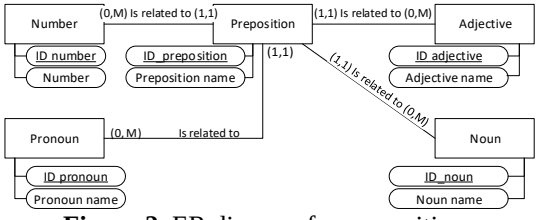


Figure 2. ER diagram for prepositions

Note: Not necessarily all, but at least one relationship between Preposition and other word types (Noun, Pronoun, Adjective, Number) must be fulfilled.

Both models will be part of a metamodel of the language, that will show the relationships between all types of words and can be used to store natural language sentences to relational database. According to the ER diagrams, adverbs are related to verbs, adjectives, nouns and other adverbs, while prepositions are related to nouns, pronouns, adjectives and numbers.

5.4. Prototype testing

The rules presented in this paper were used in the development of the metamodel of the language and in the development of the prototype of a QA system over a relational database. The first tests, which focused on verbs, nouns, and noun phrases, were extended by new sentences that focused on adverbs and prepositions. The prototype was tested on 54 sentences and 120 questions in English and 53 sentences and 118 questions in Croatian. The sentences consisted of nouns, verbs, adjectives, prepositions and adverbs. The rules described in previous work (Asenbrener Katic et al., 2021), (Asenbrener Katic et al., 2018), (Pavlic et al., 2017) and in this paper were used to transform sentences into a FNOK record suitable for input to a relational database. For the given questions, the system offered 90,3% of correct answers in English and 90,4% of correct answers in Croatian language.

6 Conclusion

In this paper, an analysis of adverbs and prepositions in Croatian and English is presented, and rules are defined on how to write these two word types in a formalized FNOK record suitable for input into the knowledge database. The Node of knowledge method was used, as well as its formalisms FNOK, which have already shown that it is possible to enter nouns, verbs and adjectives of Croatian and English into the QA system.

The paper defines two rules for adverbs and one for prepositions that allow natural language sentences containing these word types to be written in formalized form and entered into a relational database. Each rule is accompanied by several examples. ER diagram of adverbs and ER diagram of prepositions, which will be part of the language metamodel required for the development of the QA system, have been defined. The defined rules were implemented in a prototype QA system tested on 54 sentences and 120 questions in English and 53 sentences and 118 questions in Croatian. We achieved very good results (the system provided 90,3% correct answers in English and 90,4% correct answers in Croatian).

However, there are some limitations. This research is part of a larger research that will be expanded step by step. Accordingly, future research will include rules for pronouns and conjunctions, and the QA system tests will be expanded with more complex sentences and questions. Problems encountered in these tests will be the subject of further research. The language metamodel will need to be improved and developed for other word types, such as pronouns and conjunctions. It will be necessary to revise the algorithms used in the prototype and implement additional rules to improve the prototype. It is also planned to extend the research to other languages.

Acknowledgments

This work has been fully supported by the University of Rijeka under projects uniri-drustv-18-73 and uniri-drustv-18-140.

References

Asenbrener Katic, M., Candrlic, S., & Pavlic, M. (2018). Modeling of verbs using the node of knowledge conceptual framework. 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2018 - Proceedings, 1022–1027. https://doi.org/10.23919/MIPRO.2018.8400187

Asenbrener Katic, M., Candrlic, S., & Pavlic, M. (2021). Nouns in the Conceptual Framework "Node of Knowledge." Tehnicki Vjesnik - Technical Gazette, 28(6), 2088–2093. https://doi.org/10.17559/TV-20200731095329

Biblioteka Strani Jezici. (2009). Engleska gramatika, English Grammar. Knjigotisak d.o.o.

Bozic, D. (2006). Rječnik englesko-hrvatski, hrvatsko-engleski: s gramatikom. Marijan tisak.

Brachman, R., & Levesque, H. (2004). Knowledge representation and reasoning. The Morgan Kaufmann Series in Artificial Intelligence, Morgan Kaufmann.

Cali, A., Gottlob, G., & Pieris, A. (2012). Towards more expressive ontology languages: The query answering problem. Artificial Intelligence, 87–128.

Candrlic, S., Asenbrener Katic, M., & Jakupovic, A. (2020). Preliminary Multi-lingual Evaluation of a Question Answering System Based on the Node of Knowledge Method. Lecture Notes in Networks and Systems, San Francisco, SAD: Springer, 998–1009. https://doi.org/doi:10.1007/978-3-030-12388-8_69

Candrlic, S., Asenbrener Katic, M., & Pavlic, M. (2019). A system for transformation of sentences from the enriched formalized Node of Knowledge record into relational database. Expert Systems with Applications, 115, 442–464. https://doi.org/10.1016/j.eswa.2018.07.021

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Proceedings of the 2019 Conference of the North, 4171–4186.

Dovedan Han, Z. (2021). Progovorimo pythonski. *Vlastita naklada*

Gerhalter, K. (2021). Annotation Model: Criteria for Linguistic Categorization in the Database "Adjective-Adverb Interfaces in Romance."

Gomez-Perez, A., & Benjamins, V. R. (2002). Knowledge Engineering and Knowledge Management, Ontologies and the SemanticWeb. Proceedings of 13th International Conference, EKAW 2002.

Gong, H., Bhat, S., & Viswanath, P. (2018). Embedding syntax and semantics of prepositions via tensor decomposition. ArXiv Preprint ArXiv:1805.09389.

Greenbaum, S., & Quirk, R. (2006). A Student's Grammar of the English Language. Longman.

Hrvatska enciklopedija, mrežno izdanje. (2021). knowledge. Leksikografski Zavod Miroslav Krleža, Pristupljeno 15. 6. 2022. <http://Www.Enciklopedija.Hr/Natuknica.Aspx?ID=67357>.

Jakupovic, A., Pavlic, M., & Dovedan Han, Z. (2014). Formalisation method for the text expressed knowledge. Expert Systems with Applications, 41(11), 5308–5322. https://doi.org/10.1016/j.eswa.2014.03.006

Jakupovic, A., Pavlic, M., Mestrovic, A., & Jovanović, V. (2013). Comparison of the Nodes of Knowledge method with other graphical methods for knowledge representation. Proceedings of the 36th International Convention /CIS/, MIPRO 2013, 1276–1280.

Liao, S.-H., Pei-Hui, C., & Pei-Yuan, H. (2012). Data mining techniques and applications – A decade review from 2000 to 2011. Expert Systems with Apllications (39), 11303–11311.

Liu, H.-C., You, J.-X., Li, Z., & Tian, G. (2017). Fuzzy Petri nets for knowledge representation and reasoning: A literature review. Engineering Applications of Artificial Intelligence, 60, 45–56. https://doi.org/10.1016/J.ENGAPPAI.2017.01.012

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., &

Stoyanov, V. (2019). RoBERTa: A robustly optimized BERT pretraining approach. ArXiv Preprint ArXiv:1907.11692.

Neelakantan, A. R. (2017). Knowledge Representation and Reasoning with Deep Neural Networks. Doctoral Dissertations. 1114.

Pavlic, M., Dovedan Han, Z., Jakupovic, A., Asenbrener Katic, M., & Candrlic, S. (2017). Adjective representation with the method Nodes of Knowledge. 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), 1221–1226. https://doi.org/10.23919/MIPRO.2017.7973610

Pavlic, M., Jakupovic, A., & Mestrovic, A. (2013). Nodes of knowledge method for knowledge representation. Informatologia, 46(3), 206–214.

Pawar, S., Thombre, S., Mittal, A., Ponkiya, G., & Bhattacharyya, P. (2021). Tapping BERT for Preposition Sense Disambiguation. ArXiv Preprint ArXiv:2111.13972.

Pollin, C., Schneider, G., Gerhalter, K., & Hummel, M. (2018). Semantic Annotation in the Project "Open Access Database 'Adjective-Adverb Interfaces' in Romance." Proceedings of the Workshop on Annotation in Digital Humanities. CEUR Workshop Proceedings, 41–46.

Rauker Koch, M., Asenbrener Katic, M., & Pavlic, M. (2017). Fable representation in fnok and dnok formalisms using the nok conceptual framework. Annals of DAAAM and Proceedings of the International DAAAM Symposium. https://doi.org/10.2507/28th.daaam.proceedings.061

Rauker Koch, M., Candrlic, S., & Asenbrener Katic, M. (2022). Automation of the conversion of natural language to formalized node of knowledge record. *Zbornik Veleučilišta u Rijeci*, 10(1), 57–71. https://doi.org/10.31784/zvr.10.1.4

Rauker Koch, M., Pavlic, M., & Asenbrener Katic, M. (2015). Homonyms and Synonyms in NOK Method. Procedia Engineering, 100, 1055–1061. https://doi.org/10.1016/j.proeng.2015.01.466

Rauker Koch, M., Pavlic, M., & Jakupovic, A. (2014). Application of the NOK method in sentence modelling. Proceedings of the 37th Internation Convention MIPRO 2014, 1426–1431.

Silic, J., & Pranjkovic, I. (2005). Gramatika hrvatskoga jezika za gimnazije i visoka učilišta. Školska knjiga.

Socher, R., Huval, B., Manning, C. D., & Ng, A. Y. (2012). Semantic compositionality through recursive matrix-vector spaces. Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, 1201–1211.

Tezak, S., & Babic, S. (2005). Gramatika hrvatskoga jezika, Priručnik za osnovno jezično obrazovanje. Školska knjiga.

Timko, Đ. O. (2018). Najproduktivniji modeli tvorbe priloga u hrvatskom i ukrajinskom jeziku. Visnyk of Lviv University. Series Philology., 69, 102–110.

Wiedemann, G., Remus, S., Chawla, A., & Biemann, C. (2019). Does BERT make anysense? interpretable word sense disambiguation with contextualized embeddings. ArXiv Preprint ArXiv:1909.10430.

Zaheer, M., Guruganesh, G., Dubey, A., Ainslie, J., Alberti, C., Ontanon, S., Pham, P., Ravula, A., Wang, Q., Yang, L., & et al. (2020). Big bird: Transformers for longer sequences. ArXiv Preprint ArXiv:2007.14062.

KL-ADWIN: enhanced concept drift detection over multiple time windows.

Jakob Jelenčič, Jože M. Rožanec
Jožef Stefan International Postgraduate School
Jožef Stefan Institute
Jamova cesta 39, 1000 Ljubljana, Slovenia
{jakob.jelencic, joze.rozanec}@ijs.si

Dunja Mladenić
Jožef Stefan Institute
Jamova cesta 39, 1000 Ljubljana, Slovenia
dunja.mladenic@ijs.si

Abstract. *The main contribution of the paper is method that combines adaptive windowing and the drift detection method for detecting concept drift and therefore determine when a machine learning model should be retrained. We evaluated the method on four real-world datasets with concept drift. Our results show that the proposed method improves the models' performance (in three of four datasets) on unseen data when compared to the baseline.*

Keywords. Concept Drift, Kullback–Leibler Divergence, Drift Detection Method

1 Introduction

Unforseeable changes that occur in the underlying distribution of data over time is known as concept drift. Concept drift affects machine learning models, degrading their performance. While this phenomena can be observed in many use cases, it frequently affects time-series related models. Learning and running the best model in a given point in time turns out to be a non-optimal decision when the target distributions shift. Constantly re-learning a new model to address the concept shift can be an unfeasible strategy when deploying models into production, given the re-learning cost and time needed to do so.

When drastic changes take place in the data distribution, some of the historic data may no longer provide adequate learning ground for the algorithm, most probably increasing the cost of learning and hurting performance. In this paper, we propose combining two concept drift methods (Drift Detection Method (DDM) and Kullback–Leibler (KL) divergence) to determine the optimal time to relearn a machine learning model based on monitoring their performance across different time windows. The size of the re-learning window is considered a parameter.

We evaluate the method on four real datasets that present concept drift. As a baseline, we built a model that was able to see all historical data and was retrained for each step. In our research, we aimed to develop a concept drift method that would allow us to outperform the baseline or reduce the amount of retrains required to match the baselines' performance. Our results show that the proposed method improves the machine learning models' performance on unseen data.

The remainder of the paper is organised as follows. Section 2 is a short overview of related work and current state of the art methods. Section 3 describes the data we use, Section 4 reviews existing concept drift methods and we describe the proposed methodology. In Section 5, we present the results we obtained. Finally, in Section 6, we discuss the results and provide directions for future work.

2 Related Work

2.1 Concept drift

Gama et al. (2014) defines concept drift as a change in relation between in supervised learning scenario of input data and the target variable over time. Souza et al. (2020) define concept drift as a data stream where changes in distribution happened, governed by the dynamics of real-world problems and application domains that evolve. In the context of machine learning, these changes in data distribution are named concept drifts.

In practice dealing with concept drift is a wide researched problem Gama et al. (2014). Usually once you detect that a real-world data stream on which you relay to perform a activity has changed, major costs could occur. Imagine a fine tuned trading algorithm that predict change in underlying price, equity for example. Systems like that are often very sensitive to small changes in relations between input data and target variable, meaning that a small change can trick the algorithm to make a series of bad decisions, resulting in potentially enormous costs.

There exist a couple of methods how to detect and handle the concept drift problem. One way is to measure the difference in the distribution of target variable in two separate point of time Tsymbal (2004). In some case one can detect a change in distribution before there is a significant effect in the target metric, such as accu-

racy. A good example would be a Kullback-Leibler divergence. Another possible approach is to measure an error rate in target metric, and once there is a sufficient change there is a good chance that a concept drift occurred. Example of such method would be drift detection method Baena-Garcıa et al. (2006).

2.2 Kullback-Leibler divergence

The Kullback-Leibler (Van Erven and Harremos (2014)) divergence is a method of measuring a statistical distance between two different distributions. It is a distance, not a metric, and is often used to detect concept drift (Goldenberg and Webb (2019)). Let Q and P be two different probability distributions defined on the same probability space X, then the Kullback-Leibler divergence is defined as:

$$D_{KL}(P||Q) = \sum_{x \in X} P(x) \log \frac{P(x)}{Q(x)}. \quad (1)$$

2.3 Drift detection method

The drift detection method (Baena-Garcıa et al. (2006)) is a data stream error monitoring method where one knows the realized truth in a relatively short feedback loop. DDM compares the current error and its variance with the minimum error and its variance in a given time window (Gama et al. (2004)). DDM is divided into two zones: Alarm zone (2) and Detection zone (3).

$$p_{min} + 2 \cdot s_{min} \leq p_i + s_i \quad (2)$$

$$p_{min} + 3 \cdot s_{min} \leq p_i + s_i \quad (3)$$

where p_i is the error rate at time i, s_i is the standard deviation at time i, and p_{min}, s_{min} are the minimum values in the selected window. One of the disadvantages of DDM is that it provides only three types of signals, either no displacement, alarm or displacement detection.

3 Datasets

Our research focuses on real-world use cases that are prone to concept drift. We address four datasets (Amazon stock prices [1], credit card fraud detection [2], National Oceanic and Atmospheric Administration (NOAA) (NOAA) weather measurements Souza et al. (2020), and Electricity Market prices Souza et al. (2020)). All of them provide time series describing different phenomena, and a Boolean target value. We therefore address forecasting the target value as a binary classification problem. The last two datasets (NOAA and Electricity) were considered given they are part of a benchmark for streaming algorithms (Souza et al. (2020)). When selecting a subset of datasets from the benchmark, we opted for the two above-mentioned ones, given they were not strongly imbalanced.

3.1 Equity Dataset: Amazon Stock Price

Stocks are assumed to follow some form of stochastic process, either the Black-Scholes process (Merton (1976)) or more complex processes with an unknown formulation. Such processes evolve over time, and the distribution can change dramatically with extreme events (known as Black Swan events (Taylor and Williams (2009)) that alter the macroeconomic environment. We collected freely available daily stock prices from 2007 and derived an outcome variable as the change in the simple moving average over the last ten days.

3.2 Credit Card Fraud Detection Dataset

The credit card fraud dataset contains transactions made by European cardholders with credit cards in September 2013. The dataset is highly imbalanced and can be found on Kaggle. For the purpose of this research, we under-sampled the original dataset, to decrease the imbalance with respect to the minority class and thus enhance the learning of our machine learning model for the experiments we performed. We randomly under-sampled the majority class until the percentage of minority class was around ten percent.

3.3 NOAA Weather Measurements Dataset

Souza et al. (2020) published the NOAA dataset in 2020. The dataset consists of weather measurements collected over 50 years at Bellevue, Nebraska by the National Oceanic and Atmospheric Administration. This dataset contains eight features: temperature, dew point, sea-level pressure, visibility, average wind speed, max sustained wind-speed, minimum temperature, and maximum temperature. The learning task is to determine whether it will rain or not. The dataset contains 18,159 daily readings of which 5,698 are rain and the remaining 12,461 are no rain.

3.4 Electricity Market Prices Dataset

Electricity market prices dataset probably is one of the most used for the tasks of stream classification and drift detection. The data are from the Australian New South Wales Electricity Market. Prices are affected by the demand and supply. The learning task is to predict a rise or a fall in electricity prices, given recent consumption and prices in the same and neighboring regions. The dataset contains 45,312 instances, eight attributes, and

[1] The data was retrieved from the following URL (last access May 17th 2022): `https://finance.yahoo.com/`

[2] The dataset can be found at the following URL (last access May 17th 2022): `https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud`

two class labels. The dataset was published in Souza et al. (2020).

4 Methodology

4.1 Data preprocessing

All three datasets-credit fraud, NOAA, and electricity-were partitioned to match the size of the Amazon stock dataset to speed the calculation. With respect to features, the two datasets from real-world streaming dataset collection and credit card dataset were taken as they are. For the Amazon stock price dataset we have taken the most common indicators (various crossovers from moving averages, Bollinger bands, volume analysis etc.) from technical analysis Edwards et al. (2018) and transformed the one publicly available on Yahoo Finance. All four dataset were normalized using standard Z normalization. Since nothing especially complex was done and preprocessing adds nothing to the actual contribution of this work we will not go into further details.

4.2 KL-ADWIN: enhacing concept drift detection

Concept deviations can manifest themselves with varying speed, severity, and patterns (Souza et al. (2020)). There is no universal method for detection that would work perfectly in every scenario. A practical way to detect shifts is to monitor the distribution with sliding windows of different sizes or to monitor errors over time and their variance (Tsymbal (2004)). Failure to detect a shift can have a high cost, depending on the target area. For example, if a forecasting model used for automated trading in the stock markets fails to detect a shift in the distribution and starts making incorrect trading decisions. On the other hand, constant relearning may not be feasible given the frequency of new data or the hardware capacity required for complex models.

In our research, we propose KL-ADWIN, a method similar to the ADaptive WINdowing (ADWIN) algorithm (Bifet and Gavalda (2009)), where we monitor different time windows to detect possible distribution shifts in the underlying data. In conjunction, we use a slightly modified version of DDM, defined as:

$$mDDM(t) = p_i + s_i - (p_{min} + 2 \cdot s_{min}). \quad (4)$$

In this way, we obtain a real number for each moment in time. The difference between classic DDM is that by doing so we can assign a weight to it and create a continuous indicator. Ideally mDDM is a function that would output something close to 0 and when the output is higher than 0 that would indicate alarm zone or possible change.

Similarly, we calculate KL divergence for two time windows that follow each other without overlap. For better understanding, imagine test folds $[1, 2, 3, 4, 5]$. For windows of size two and one, we would first compare the distributions from folds one and two with the distribution in three. Then the distributions from folds two and three with the one in fold four, and so on. To make the signal comparable to that of $mDDM$, we divide by the mean and subtract one. Let m and k be the indices of the windows, then we define Kl criteria as:

$$mKL(t) = \frac{KL(X_m, X_k)}{\frac{1}{t}\sum_{i=1}^{t} mKL(i)}. \quad (5)$$

To illustrate the idea of KL divergence we have plotted distribution of target variable of Amazon stock prices dataset after concept drift. Both distributions can be seen on Figure 1. With the help of KL divergence one can measure how much is the difference between two. The bigger the difference the higher the integral will be. Now since we cannot understand the meaning of a single number, we have divided the current one with the historic average. To elaborate on example, if the difference plotted on Figure 1 would happen all the time, then the (5) will not sound the alarm since the similar change happened in historic data and the model should be able to handle it. If not, then the (2) will be higher than one, and possibly signaled that relearn is needed.

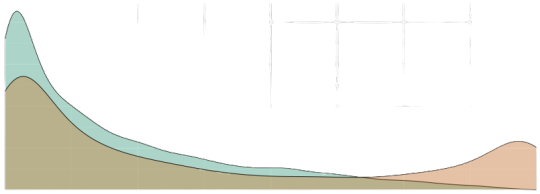


Figure 1: Example of distribution shift from Amazon stock prices dataset.

Finally, we can define the proposed method for determining the optimal time for relearning the model. Let w_1 and w_2 be the weights that sum to 1; $w_1 + w_2 = 1$. Let be the th scalar threshold to be optimized. Then we can define the criteria as

$$\textit{KL-ADWIN} = w_1 \cdot mDDM(t) + w_2 \cdot (mKL(t) - 1) > th. \quad (6)$$

4.3 Relearning window

In machine learning, it is widely assumed that more data usually means a better model. Nevertheless, more data does not always mean more information (a phenomena exploited by e.g., active learning (Kumar and Gupta (2020)), and more recently, the datacentric artificial intelligence approaches (Marcu and Prügel-Bennett (2021); Paiva et al. (2021))). Furthermore, if the distribution of the data changes over time (concept drift), past data may no longer provide cues for forecasting current target values. For example, imagine

a world where stock prices only go up and the market suddenly reverses. Adding additional data from the historical distribution that is no longer accurate could only hurt performance. While determining the optimal relearning window remains a subject of future work, for this research work we have fixed the time window on 25% of the existing dataset.

5 Results

In this section we present the results we obtained with the proposed method. We evaluated and analyzed both $mDDM$ and mKL over time for each test dataset to determine the optimal relearning point. We used the random forest model with 200 trees due to its robustness in terms of fine-tuning.

We used the following parameters: $w_1 = 0.5$, $w_2 = 0.5$ and $th = 0$. Relearning window was fixed at $p = 0.25$ (25% of total dataset). For further analysis, we also track the two combinations where one of w_i is set to zero. In Figures 2 through 5, the green dots represent mKL and the red dots represent $mDDM$. Relearning occurred only where both dots appeared together. Where single dots appear, either red or green, this indicates that one of the methods detected a shift (w_i of the other was set to zero), but the other was not strong enough to warrant relearning.

We evaluated the method on all four datasets and compared it to the model that would be relearned for each new dataset (following a streaming fashion, despite being a batch machine learning model). It is important to emphasize that the proposed method is much more economical, since the baseline requires constant retraining of the model. When the proposed method signals relearning, we only take the last 25% of historical data, making the retrainig cheaper and faster than the baseline, which requires all of the historical data. The ROC AUCs of the streaming models are shown as a dashed line, while the proposed method has a solid line.

5.1 Equity dataset: Amazon Stock Price

In the stock dataset, we forecast the change in the moving average of the Amazon stock price (trend change). The results are shown in Figure 2. We can see that the proposed approach is more stable than the streaming approach with only three learnings. Also, we can see a green dot indicating a possible shift just before the single peak in the ROC AUC of the streaming model, which can lead to another relearn with a better parameter choice and possibly outperform the ROC AUC of the streaming model in this part of the test set as well.

5.2 Credit Card Fraud Detection dataset

Figure 3 shows the results of the credit card fraud detection dataset. At first glance, it is clear that the results

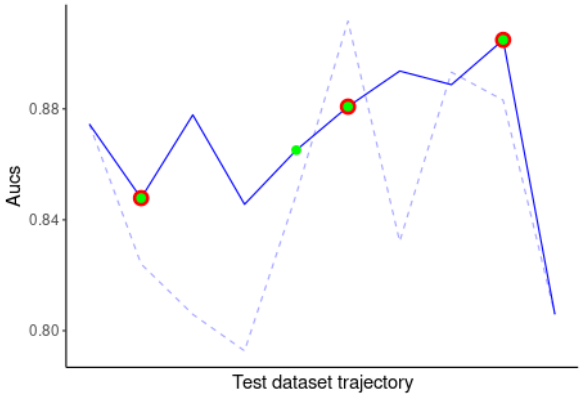


Figure 2: Results of proposed method on equity dataset.

are not as good as the previous dataset. Here, only two new learnings were required, and twice the signal indicated possible shifts but did not exceed the new learning threshold.

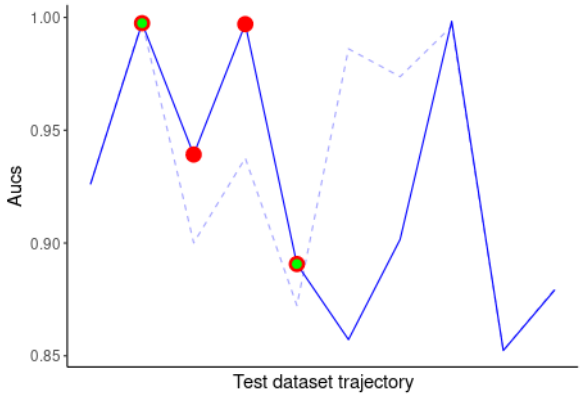


Figure 3: Results of proposed method on credit card fraud detection dataset.

5.3 NOAA Weather Measurements Dataset

Figure 4 shows the results for the NOAA dataset. The NOAA and Equity datasets best fit the type of problems that this method is designed to solve. We can see that the proposed method performs significantly better with only 2 learning operations.

5.4 Electricity market prices dataset

Figure 5 shows the results for the current dataset. This is the only dataset where we can say that the streaming model is better. The method detected only a single shift point, briefly outperformed the ROC AUC of the streaming model, and then performed worse all the time. We believe the reason for this is the nature of the dataset. Unlike the stock and NOAA data, where the distributions change and are somewhat dependent on the near past, the streaming dataset is seasonal and

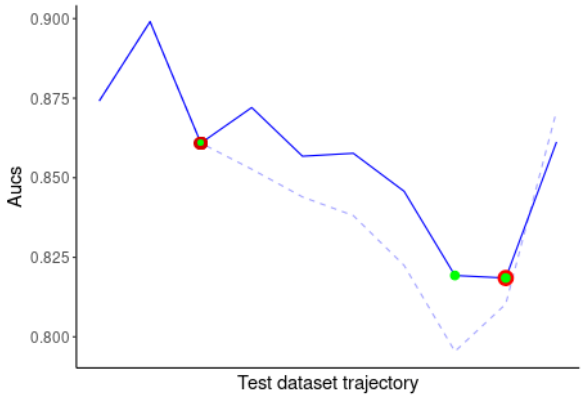


Figure 4: Results of the proposed method on NOAA weather measurements dataset.

cyclical. To illustrate the example, fitting the model to data from the spring does not help predict demand or prices in the summer. Then it makes sense that the model that saw the previous summer would perform better.

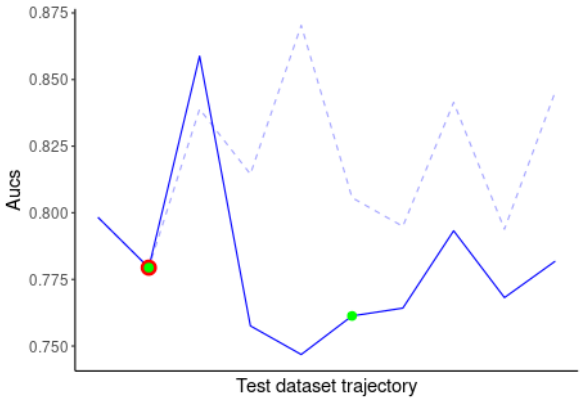


Figure 5: Results of proposed method on the electricity market prices dataset.

6 Conclusions, and Future Work

In this research, we have described how we combined two concept drift methods into a new one, which shows promising results. We evaluated the method on four real-world datasets and achieved comparable or better results than the baseline model, except for the Electricity Market Prices dataset. However, we consider that the performance on that dataset was degraded given that the concept drift detector could not distinguish between concept drift and seasonality changes, therefore discarding valuable data. On the other hand, we were pleasantly surprised with the results obtained from the NOAA dataset: our method significantly reduced the relearning cost and notably increased the models' performance.

Future work will focus on three directions of research. First, we will explore better ways to determine optimal parameters given a concept drift detection (e.g., size of time window considered to retrain and test the model). Second, the concept drift on feature values can be empirically monitored; therefore, such information is used to negate the shift, leading to better performance Turnbull (1976). Finally, we will explore new mechanisms to distinguish between concept drift and seasonality changes to enhance the proposed method further and make it more robust in such scenarios.

Acknowledgments

This work was supported by the Slovenian Research Agency.

References

Baena-Garcıa, M., del Campo-Ávila, J., Fidalgo, R., Bifet, A., Gavalda, R., and Morales-Bueno, R. (2006). Early drift detection method. In *Fourth international workshop on knowledge discovery from data streams*, volume 6, pages 77–86.

Bifet, A. and Gavalda, R. (2009). Adaptive learning from evolving data streams. In *International Symposium on Intelligent Data Analysis*, pages 249–260. Springer.

Edwards, R. D., Magee, J., and Bassetti, W. C. (2018). *Technical analysis of stock trends*. CRC press.

Gama, J., Medas, P., Castillo, G., and Rodrigues, P. (2004). Learning with drift detection. In *Brazilian symposium on artificial intelligence*, pages 286–295. Springer.

Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., and Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM computing surveys (CSUR)*, 46(4):1–37.

Goldenberg, I. and Webb, G. I. (2019). Survey of distance measures for quantifying concept drift and shift in numeric data. *Knowledge and Information Systems*, 60(2):591–615.

Kumar, P. and Gupta, A. (2020). Active learning query strategies for classification, regression, and clustering: a survey. *Journal of Computer Science and Technology*, 35(4):913–945.

Marcu, A. and Prügel-Bennett, A. (2021). On data-centric myths. *arXiv preprint arXiv:2111.11514*.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2):125–144.

Paiva, P. Y. A., Smith-Miles, K., Valeriano, M. G., and Lorena, A. C. (2021). Pyhard: a novel tool for generating hardness embeddings to support data-centric analysis. *arXiv preprint arXiv:2109.14430*.

Souza, V. M. A., Reis, D. M., Maletzke, A. G., and Batista, G. E. A. P. A. (2020). Challenges in benchmarking stream learning algorithms with real-world data. *Data Mining and Knowledge Discovery*, 34:1805–1858.

Taylor, J. B. and Williams, J. C. (2009). A black swan in the money market. *American Economic Journal: Macroeconomics*, 1(1):58–83.

Tsymbal, A. (2004). The problem of concept drift: definitions and related work. *Computer Science Department, Trinity College Dublin*, 106(2):58.

Turnbull, B. W. (1976). The empirical distribution function with arbitrarily grouped, censored and truncated data. *Journal of the Royal Statistical Society: Series B (Methodological)*, 38(3):290–295.

Van Erven, T. and Harremos, P. (2014). Rényi divergence and kullback-leibler divergence. *IEEE Transactions on Information Theory*, 60(7):3797–3820.

NoSQL Databases Deployment in Various Application Domains: A Systematic Literature Review

Jelena Petrovački, Dejana Gladić, Sara Havzi, Darko Stefanović, Sonja Ristić

Faculty of Technical Sciences, Department of Industrial Engineering and Engineering Management

University of Novi Sad

Trg Dositeja Obradovića 6, 21000 Novi Sad

{jelenapetrovacki, dejanagladic, havzisara, darko.stefanovic, sdristic}@uns.ac.rs

Abstract. *The continuous growth of data volume to be stored and processed in the shortest possible time brings new requirements in terms of data storage, processing, analysis and visualization. Relational databases are not capable to store and process big data efficiently and NoSQL databases have emerged as an appropriate alternative. The range of available NoSQL systems is wide and the suitability of a given NoSQL system depends on the problem domain. In this paper, a systematic literature review of various NoSQL systems by application domains is performed in order to ease the selection of appropriate NoSQL system for a specific purpose.*

Keywords. NoSQL, Big Data, application domain, usage domain, document database, graph database, key-value database, column database, data analysis, Systematic Literature Review

1 Introduction

Big data as a concept is increasingly used in the business and educational spheres of modern life. The term big data encompasses a huge set of data collected from a variety of sources, from business and educational processes, through administration and governance to the social aspects of human life, such as social networks and various electronic communication channels.

Several different authors describe the big data concept through five essential characteristics designated as 5V:

- Volume – represents the amount of stored and processed data that can reach up to several terabytes and petabytes;
- Variety – means the format of data that is stored and divided into three main categories: unstructured, semi-structured and structured data;
- Velocity – represents the speed of data generation and processing and big data is suitable for processes that must be processed in limited real-time;
- Veracity – represents the quality and accuracy of stored data; and
- Value – the value of data that is stored and analyzed because it makes no sense to create and store data that will later serve no purpose (Faridoon & Imran, 2021).

Adequately processed and analysed data should influence more efficient and effective decision-making that leads to progress.

Big data sets raise various challenges, including methods to collect, store and process different types and structures of data. Huge volumes of data recorded in heterogeneous formats that have to be processed in limited real-time, are difficult to be handled by traditional relational database management systems (RDBMSs). Structured Query Language (SQL) is standardized data manipulation and query language of RDBMSs. Therefore, relational databases managed by RDBMSs are also known as SQL databases. For decades, SQL databases were considered as standard technology for persisting and managing large volumes of data. Big data sets bring the need for levels of availability beyond those supported by SQL databases and the challenges of scaling such databases horizontally led to the emergence of a new generation of databases grouped under the umbrella term NoSQL databases (Strauch & Kriha, 2011). NoSQL databases are a current approach to collecting and storing such large and distributed datasets (Philip Chen & Zhang, 2014), that allow the volume, variety and velocity characteristics of big data sets (Naik & Joshi, 2017) to be adequately managed.

There is a plethora of various NoSQL systems. They are based on different paradigms and aimed at storage of various data formats. Consequently, the particular suitability of a given NoSQL database system depends on the problem it is aimed to solve. In the paper we perform systematic literature review (SLR) on the deployment of NoSQL systems in

various application domains. The main goals of presented SLR are: i) to understand the applicability of various NoSQL systems in different application domains, in order to ease the selection of appropriate NoSQL system for a specific purpose; ii) to identify challenges of various NoSQL systems and issues of their deployment in different domains; and iii) to discuss the potential research gaps. In order to accurately answer the given topic, the application of NoSQL databases in different domains, the research of the concept of big data and its prevalence in different sectors of the economy was first performed due to the close connection between these two concepts. Different authors mention the same economic activities as the domains of application, primarily big data, and in this paper, the ones that are most often mentioned in most papers are singled out in order to make a search query adequately. Economic activities that are analyzed are as follows: communications, media and entertainment, healthcare, education, industry, government, insurance, trade (retail and wholesale), transportation, and agriculture. Based on the listed economic activities, the search query described in Section 3 will be generated.

The reviewed papers most often mention the application of big data sets within the Internet of Things (IoT) system, which is implemented in various sectors, and it is very important to search for the application of NoSQL databases within such a created system.

In the second chapter, a theoretical overview of the concepts on which SLR on a given topic is based is described in detail. The third chapter includes a review of SLR planning, implementation, as well as analysis of selected papers through descriptive statistics analysis. The fourth chapter provides answers to research questions that are defined in the initial phase of the literature review. In the fifth chapter, the concluding opinions on this topic are given, as well as the directions for further research.

2 Theoretical foundations

The fact that most people think of tables and SQL as the first thought of the term "*database*" is not surprising. The reason for this is the relatively late emergence and beginning of the use of NoSQL databases that began to be used in the late 2000s. Since then, they have become more and more popular, and it seems that their usage will become more and more important in the future. The most significant difference with well-known relational databases is in the format of the stored data.

The creators of NoSQL databases emphasize their flexibility as a major advantage over traditional databases. Flexibility is reflected in a database schema that is not strictly defined as in a relational database, and data models may differ from each other. In addition to this advantage of NoSQL databases, many others stand out, among others:

- structured, semi-structured and unstructured data can be stored,
- the costs are much lower,
- ability to distribute large amounts of data stored in the cloud, via multiple servers,
- fast and efficient horizontal scaling,
- faster query execution,
- ease of integration of the persistence layer with the business logic layer of application and
- ease of usage.

Relational databases offer a very strict model of transaction control called ACID (Atomicity, Consistency, Isolation and Durability) that is not needed for many types of applications. A strict transactional model can reduce the performance of many systems (Stonebraker, 2010). Just as ACID is core feature of SQL databases, BASE (Basically Available, Soft state and Eventual consistency) is feature of NoSQL databases. NoSQL databases are designed with horizontal scalability and partition tolerance delivering increased availability and fault tolerance at a cost of temporary data inconsistency.

Four groups of NoSQL databases stand out: key-value databases, document oriented databases, column oriented databases and graph databases. They are quite different from each other and in the Table 1. are given an overview and general characteristics of each type of NoSQL databases.

3 Methodology

In the process of conducting the SLR, presented here, the approach of Barbara Kitchenham is followed, according to which all activities can be divided into three phases: planning the review, conducting the review and reporting the review (Kitchenham, 2004).

3.1 Planning the literature review

At the beginning of this phase of the research, a search of already existing literature reviews dealing with this topic is approached (Kitchenham, 2004). An SLR on a similar topic was found, but the way to review the literature is different. The query set in (Faridoon & Imran, 2021) differs from the query set in this paper and different index databases are searched. The focus of the search in (Faridoon & Imran, 2021) was on big data data storage tools based on NoSQL databases, their advantages and disadvantages.

Table 1. Overview of NoSQL database types

DOCUMENT DATABASE	Data model	JSON, BSON or XML document
	About DB	The data is stored in the form of documents. Each document contains fields and their values. Documents can be nested with each other.
	Use cases	Commerce platforms, trading platforms, content management, web and mobile application development across industries
	Advantages	Flexibility - documents within a collection do not have to be structured the same (consistency is not required). The format in which the data within the database document is written is similar to object notation, so the use of such structured data is easier to create applications.
	Disadvantages	Nesting documents can lead to a complex system and compromise performance.
	Popular databases	MongoDB, SimpleDB, CouchDB, Couchbase
KEY-VALUE DATABASE	Data model	Table with two columns: key and value
	About DB	The simplest form of NoSQL database that resembles a relational table with two columns - key and value. The keys are unique identifiers for the values. The values can be another key-value pair in which case the structure of the database grows more complex.
	Use cases	Application logs, shopping carts, user preferences, user profiles
	Advantages	Easy to use.
	Disadvantages	There is no defined query language, so search is only possible based on the key.
	Popular databases	Aerospike, Redis, Riak
WIDE-COLUMN DATABASE	Data model	A set of columns
	About DB	The table is organized as a set of columns - emphasis on columns, unlike relational databases where the emphasis is on rows. Multiple columns form a column family.
	Use cases	Analytics, internet search, large-scale web applications
	Advantages	They are suitable for applying aggregate functions to data stored within the same column. Also, they are very useful for data warehouses, or when there is a need for handling intensive querying.
	Disadvantages	These databases are not very efficient with online transactional processing.
	Popular databases	BigTable, HBase, Cassandra, Hypertable
GRAPH DATABASE	Data model	Graph structure (nodes and egdes)
	About DB	Data is stored in nodes and edges. Elements are represented as nodes, and connections between elements as edges. They are often used in addition to existing relational databases.
	Use cases	Fraud detection, social networks, knowledge graphs, geospatial application, recommendation engines
	Advantages	Suitable for highlighting the relationship between elements (without the need to join tables as in SQL)
	Disadvantages	Scalability -they are hard to be scaled across a number of servers.
	Popular databases	Neo4j, InfiniteGraph, HyperGraphDB

The aim of presented SLR is precisely determined, and it includes the identification of a data storage model that will in the best way that is in terms of the highest number of performances, meet the needs of the application for which the database is being created.

In order to meet the aim of this research, adequate research questions have been defined. Research questions include the following:

RQ1: Which NoSQL systems are suitable for storing data for a particular application domain?

RQ1.1: What are the properties and data types of data stored in these NoSQL databases?

RQ2: What are the advantages of NoSQL systems?

RQ2.1: What are the advantages of a particular NoSQL system compared to other NoSQL systems?

RQ2.2: What are the advantages of NoSQL systems compared to RDBMSs?

RQ3: What are the challenges of using the NoSQL systems?

In order to answer the previously defined questions, the index databases Scopus and Web of Science (WoS) were searched.

When searching for the listed index databases, the keywords *NoSQL database*, *usage*, *utilization*, *implementation* and *domain* were used firstly. A query that represents a combination of these keywords was generated as follows:

NoSQL AND domain AND (usage OR utilization OR implementation).

After analyzing the papers obtained by applying the presented query, it was concluded that the found papers will not provide an answer to the defined research questions, and that the research goal will not be fulfilled.

The main shortcoming of the previously defined query was noticed, which are imprecisely defined desired domains for the usage of NoSQL databases. Therefore, it is decided to explicitly enumerate the most important economic activities, as it is stated in Section 1. The modified query is:

NoSQL AND (communication OR media OR entertainment OR healthcare OR education OR industry OR insurance OR trade OR transportation OR IoT OR agriculture OR government) AND PUBYEAR > 2016 AND (LIMIT-TO (DOCTYPE , "cp") OR LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")).

Some of the defined exclusion criteria were applied through the presented query. Here is an overview of all the identified exclusion criteria:

1. Non-English papers are excluded.

2. Papers other than "conference paper" or "journal article" should be removed.
3. It is necessary to remove papers that appear in the search results of both index databases.
4. If one author has more than one paper regarding the same approach, only one paper should be included in the review.
5. Papers that deal only with the analysis and implementation of some of the SQL databases or some auxiliary frameworks without the analysis of NoSQL databases and tools are not considered.
6. Papers that theoretically describe different NoSQL systems without analyzing system performance over specific data from any of the identified usage domains are excluded.

Identified inclusion criteria are:

1. Papers that are accessible electronically.
2. Papers that are published since 2017.
3. Papers that test the performance of various NoSQL system on data collected from a real system to propose the best NoSQL database.
4. Papers describing the implementation of a software solution from a real system, with a focus on NoSQL databases and tools.
5. Papers comparing NoSQL and SQL databases over some specific data from different domains.

For further analysis and presentation of descriptive statistics, data such as: publication year, publication type, authors, keywords, NoSQL system and usage domain mentioned in the paper, advantages and challenges that are solved by using NoSQL systems and shortcomings and challenges that NoSQL systems cannot respond to were extracted from each paper.

3.2 Conducting the literature review

The initial search of both selected index databases resulted in a total of 845 papers that meet the defined criteria. Then the papers were selected based on the title of the paper, as well as based on the abstract, and the numbers of selected papers are shown in Table 2.

Table 2. Number of selected papers by search phases

	WoS	SCOPUS
Initial search	317	528
Number of selected papers based on title	76	42
Number of selected papers based on abstract	59	29
Number of selected works based on their content	32	11
In total:	43	

The papers accepted in the previous phases were analyzed in detail and 43 papers were selected on which descriptive statistics will be based and on the basis of which the research questions will be answered.

Through all phases, the previously mentioned inclusion and exclusion criteria were applied.

3.3 Results of review

In this section descriptive statistics based on relevant information extracted from 43 selected papers is presented. For a more compact presentation of the contents of Table 3. and Table 4., ordinal numbers of papers are used, and a list of associated ordinal numbers is given in the table in the Appendix.

The origin of selected papers is shown in Table 3. It can be seen that 69.8% of papers were published at various conferences, while the rest of the papers (30.2%) were published in scientific journals.

Table 3. Source type of selected papers

Source type	Papers	Percentage
Conference proceedings	1, 2, 3, 8, 9, 10, 11, 12, 13, 15, 16, 17, 19, 20, 21, 22, 23, 26, 27, 29, 31, 32, 34, 35, 36, 37, 39, 40, 41, 42	69,8 %
Journal article	4, 5, 6, 7, 14, 18, 24, 25, 28, 30, 33, 38, 43	30,2 %

In Fig 1. is given the number of published works by year. Based on the graph shown in Fig 1, we cannot observe a trend of publishing papers on this topic in a certain period of time.

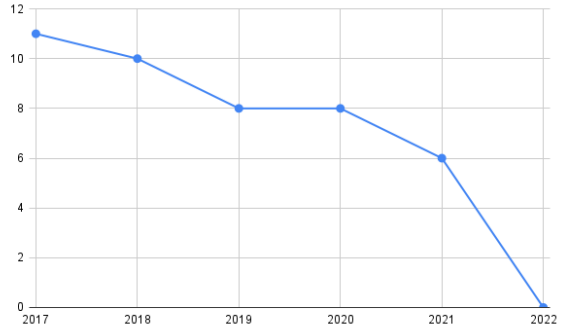


Figure 1. Number of selected studies per year

A large part of the selected papers (44.2%) refers to the comparative analysis of different NoSQL databases and testing their performance, while in a smaller number of papers (16.3%) the authors compare the performance of NoSQL and relational databases. As can be seen in Table 4, other papers (39.5%) deal with the implementation of a specific NoSQL database that they consider to be the most

suitable for the needs of the application for which the database is being created.

Table 4. Division of papers based on content

	Papers	Percentage
Comparison of SQL and NoSQL databases	5, 8, 19, 20, 24, 33, 36	16,3 %
Concrete application of a particular NoSQL database	2, 9, 13, 15, 18, 23, 27, 29, 32, 34, 37, 38, 39, 40, 41, 42, 43	39,5 %
Comparison of different NoSQL databases	1, 3, 4, 6, 7, 10, 11, 12, 14, 16, 17, 21, 22, 25, 26, 28, 30, 31, 38	44,2 %

The analysis of selected papers also recorded the key words that most often appear in papers that correspond to the chosen topic. As is shown in Fig 2., the largest share of papers is based on the key words: *NoSQL Database* and *Big Data*. In addition to them, various databases (*MongoDB*, *Cassandra* and *Couchbase*) appear in the identified keywords. When it comes to application domains, the key word "*healthcare*" stands out, which leads to the assumption that most of the papers will be based on the use of NoSQL databases within medicine. The appearance of keywords "*IoT*" is evident, which we can relate to the moment when most of the data to be stored within the database is actually generated in this way, and that the term Internet of Things can be associated with all identified application domains.

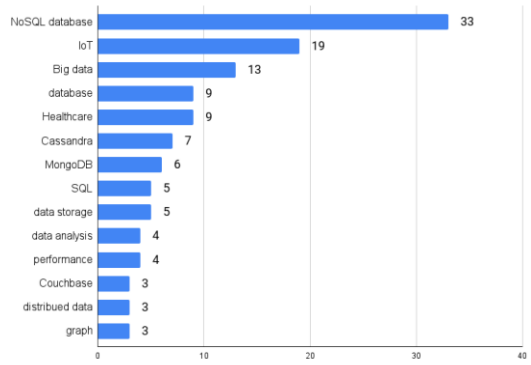


Figure 2. Keywords from selected papers

Different data structures will give different results when it comes to the performance of the database within which the data is stored. It can be concluded that for appropriate data structures some databases are suitable to a greater extent and some to a lesser extent. Assuming that relatively similar data formats are used within one domain of the application, it is justified to generalize the use of the database to the categories shown in the graph in Fig 3. As could be assumed when analyzing keywords, the number of papers based on the use of NoSQL databases for health purposes is the largest. Also, it cannot be claimed that NoSQL databases are not used or suitable for storing data from certain categories (education, insurance), but only that the search results did not find adequate work related to these application domains.

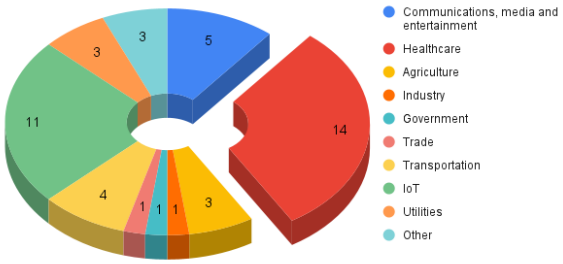


Figure 3. The number of usage domain appearances in selected papers

A detailed analysis of the papers recorded the NoSQL databases that were discussed in the largest number of papers. From the graph in Fig. 4 it can be seen that in the papers the most analyzed document databases, given that MongoDB is mentioned in 21 selected papers, which is almost half of the total number of selected papers. Cassandra, a representative of column databases, and Neo4j, which belongs to graph databases, also appear in most papers.

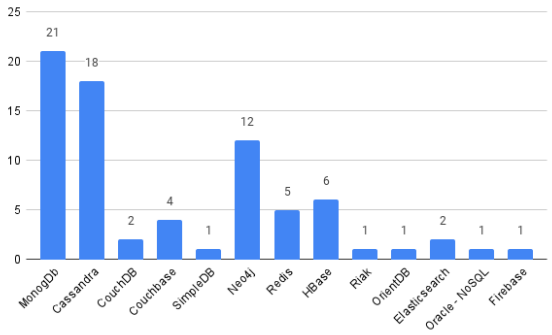


Figure 4. The number of NoSQL databases appearances in selected papers

The latest descriptive statistics shown in the graph in Fig. 5 represent in which application domain the most common type of NoSQL database is. In most cases, there is no significant difference for any of the types of NoSQL databases, which calls into question whether the analysis of types of NoSQL databases by application domains makes sense. The answer to this question will be presented through the analysis of selected papers, where the content of the papers will be described in more detail.

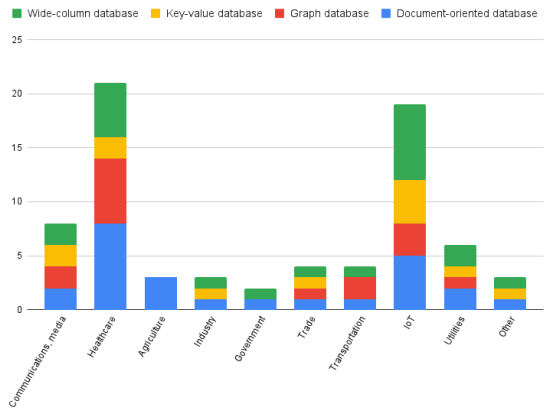


Figure 5. NoSQL types per usage domains

4 Discussion

In this chapter, previously asked research questions will be answered.

In order to give an answer to RQ1, the graphical representation from Fig. 5 is important. Given that in most cases it cannot be observed that some type of NoSQL database is predominantly used within some of the categories, it cannot be said with certainty that the domain of application is crucial when choosing a database.

The answer to RQ1.1 would be more precise, because at the level of granularity on the basis of which this question is defined, it is possible to draw conclusions – for which types of data which database is suitable. Then it could be researched which types of data are most represented in which domain of application.

In paper named "*Which NoSQL database for IoT applications?*" (Amghar et al., 2018), the authors review the literature based on which they conclude that most IoT applications use NoSQL databases for their data storage needs. They ask which of the many available NoSQL databases is the best choice and come to the conclusion that it depends on the needs of the domain. They identify 7 needs of IoT applications: data heterogeneity, semantic interoperability, scalability, real time processing, security, spatial data handling and data aggregation. The characteristics of the selected NoSQL databases are analyzed and it is concluded which database would correspond to which type of data. The results of the research mostly emphasize that in the domains of application that involve a large amount of spatial data, there is an obvious need to use NoSQL graph databases, given the importance of entity connectivity.

The domain of application described in most of the papers is health. Observing the breadth of this domain, it is expected that data formats will differ depending on the specific needs of the application for which the database is being created. It is not uncommon for different applications to need to store different data in different databases, in order to ensure the best performance of the application.

This situation has been described in (Celesti, et al., 2020b) and (Cerbulescu, C. & Cerbulescu, C. M., 2017) and (Kundu et al., 2021) describes the different needs of an application belonging to the same domain as applications on which the first two papers are based.

(Celesti, 2020b) state that in the tele-rehabilitation system in which patients send data via devices from their home, and doctors review it and return therapy, two groups of data are separated. The first group includes personal data on doctors and patients, and since this is data that is rarely changed and structured, the best option for storing this group of data is a relational database. The second group consists of patient-generated data. The volume of this data is much larger and most often this data is not structured, so in this situation the use of NoSQL database is recommended.

(Cerbulescu, C. & Cerbulescu, C. M., 2017) in their study emphasize the importance of rapid processing of critical data. Their approach is based on extracting a small amount of key data that is considered critical and this data is recorded within a relational database. This database has good performance when it comes to small amounts of data. In addition, in order for the SQL database to be as fast as possible, they set the expiration date on the data entered in it, and unnecessary data is deleted after the expiration of that period. All other data, which is not identified as critical, is stored within the NoSQL database. The previous analysis also answered research question RQ2.2 because there is a clear difference in the usage of relational and NoSQL databases. The authors cite these differences as the main reason for choosing one or another type of database.

Through the study (Kundu et al., 2021) we can confirm the statement that within one domain of application there may be needs for different databases. In order to build a system that shows "trust" to the doctor, the Neo4J database graph was used. For each new registered doctor, a new node is added that contains information about the doctor. A new node is also created for each patient who schedules an examination, and a new edge is created for each scheduled appointment with a doctor. Each scheduled appointment with a doctor increases the doctor's score (if a patient returns to the same doctor several times, the number of points added to the previous score increases). Applications based on the principle of "trust" are inherent in other domains of application, not only health.

In order to answer the research question RQ2.1, the selected papers must be analyzed in detail. The choice of database also depends on the domain in which it is applied, but mostly on the data structure that is stored there and the needs of applications for

certain performance tools. In the paper (Chen & Lee, 2019), the authors describe the method of selecting NoSQL systems depending on the purpose and needs of applications. The first case study is a shopping site that is slow because it uses a relational database to store large amounts of data. The solution to the problem is to use the NoSQL database, which will enable the merging of various relational database tables in advance and shorten the time required to execute select queries. A wide column NoSQL database was chosen because access to the database often requires searching for data by a specific feature. The choice was reduced to Apache Cassandra and Apache HBase. HBase spent much less time accessing data than Apache Cassandra so HBase was chosen as adequate for implementation. Another study chose the document store NoSQL database, MongoDB, for collecting and storing files with multimedia material for the needs of the news agency, because it is the most used in that category. In (Pramukantoro et al., 2017), the authors also select the MongoDB database for storing data such as surveillance camera images, the degree of CO gas presence, or temperature and fluidity data from different IoT sensors. The main reason is that it contains a GridFS format that allows you to store large amounts of multimedia content. The authors of the paper (Sharma & Kaur, 2019) want to determine the areas in the city where there is heavy traffic, at what time of day and how much these parts violate traffic safety. The data to be stored are geolocation points (latitude and longitude) and time data. They choose the Apache Cassandra Cluster tool whose interface is suitable for data modeling and query execution. Also, using that tool can be created a cluster of many nodes to which data can be distributed. The node is easy to add and remove because they are all independent of each other. If one of the nodes fails, the others can run smoothly, but the malfunction may increase the time required to execute the query. In each selected paper, the advantages and reasons for choosing a certain NoSQL tool are defined, but in response to the question RQ2.1. some of them are described.

The response to RQ3 is not entirely satisfactory. In order to obtain more precise information about the challenges when using NoSQL databases, it is necessary to define a different search query for index databases. From the accepted papers, it was concluded that the biggest challenge is data security. As an example, the authors of the paper (J. Aqel et al., 2019) describe weak mechanisms for data encryption. All data in MongoDB is stored as plain text and there is no encryption mechanism to encrypt data files and all passwords in Cassandra are encrypted using of MD5 hash function so that the passwords are very weak.

5 Conclusion

Based on the conducted systematic review of the literature, it can be concluded that there is no rule according to which an adequate NoSQL database would be selected based on the domain of the application for which the database is being created.

The categorization of the usage of NoSQL databases by application domains is not the most precise, but on the basis of selected papers, it is possible to obtain an image of which data features most often occur in which application domains.

In most cases, it is necessary to store data in several different types of databases for the needs of applications, because different data structures are used within one application. Although the focus of this paper is based on NoSQL databases, the usage of relational databases, which are still number one when it comes to transactions and static data structures, cannot be completely ignored.

Further research directions would include a more detailed search of the challenges that arise during the implementation of NoSQL databases, then the classification of identified challenges into those whose negative effects can be mitigated, as well as those whose negative effects cannot be avoided.

References

Ai-Sakran, A., Qattous, H., & Hijjawi, M. (2018). A proposed performance evaluation of NoSQL databases in the field of IoT. *2018 8th International Conference on Computer Science and Information Technology, CSIT 2018*, 32–37. https://doi.org/10.1109/CSIT.2018.8486199

Al Sadi, M. B., Wimmer, H., Chen, L., & Wang, K. (2017). Improving the efficiency of big forensic data analysis using NoSQL. *International Conference on Mobile Multimedia Communications (MobiMedia), 2017-July*, 240–248. https://doi.org/10.475/eai.13-7-2017.2270344

Almeida, A., Oliveira, F., Lebre, R., & Costa, C. (2020). NoSQL distributed database for DICOM objects. *Proceedings - 2020 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2020*, 1882–1885. https://doi.org/10.1109/BIBM49941.2020.9313430

Amghar, S., Cherdal, S., & Mouline, S. (2018). Which NoSQL database for IoT applications? *2018 International Conference on Selected Topics in Mobile and Wireless Networking, MoWNeT 2018*, 131–137. https://doi.org/10.1109/MoWNet.2018.8428922

Asiminidis, C., Kokkonis, G., & Kontogiannis, S. (2019). Database Systems Performance Evaluation for IoT Applications. *SSRN Electronic Journal*,

November 2019. https://doi.org/10.2139/ssrn.3360886

Braulio, L. D. C., Moreno, E. D., De MacEdo, D. D. J., Kreutz, D., & Dantas, M. A. R. (2018). Towards a Hybrid Storage Architecture for IoT. *Proceedings - IEEE Symposium on Computers and Communications*, *2018-June*(5), 470–473. https://doi.org/10.1109/ISCC.2018.8538474

Celesti, A., Buzachis, A., Galletta, A., Fazio, M., & Villari, M. (2018). A NoSQL Graph Approach to Manage IoTaaS in Cloud/Edge Environments. *Proceedings - 2018 IEEE 6th International Conference on Future Internet of Things and Cloud, FiCloud 2018*, 407–412. https://doi.org/10.1109/FiCloud.2018.00065

Celesti, A., Celesti, F., Galletta, A., Fazio, M., & Villari, M. (2020a). Improving Machine Learning Algorithm Processing Time in Tele-Rehabilization through a NoSQL Graph Database Approach: A Preliminary Study. *Proceedings - IEEE Symposium on Computers and Communications*, *2020-July*. https://doi.org/10.1109/ISCC50000.2020.9219700

Celesti, A., Lay-Ekuakille, A., Wan, J., Fazio, M., Celesti, F., Romano, A., Bramanti, P., & Villari, M. (2020b). Information management in IoT cloud-based tele-rehabilitation as a service for smart cities: Comparison of NoSQL approaches. *Measurement: Journal of the International Measurement Confederation*, *151*, 107218. https://doi.org/10.1016/j.measurement.2019.107218

Cerbulescu, C., & Cerbulescu, C. M. (2017). Proposed Architecture to manage critical states predictions in IoT applications. *ACM International Conference Proceeding Series*, *Part F1309*, 1–4. https://doi.org/10.1145/3136273.3136279

Chauhan, N., & Malik, M. (2021). Home Automation With NoSQL and Node-RED Through Message Queuing Telemetry Transport. *Advances in Intelligent Systems and Computing*, *1172*(April), 427–434. https://doi.org/10.1007/978-981-15-5566-4_37

Chen, J. K., & Lee, W. Z. (2019). An introduction of NoSQL databases based on their categories and application industries. *Algorithms*, *12*(5), 1–16. https://doi.org/10.3390/a12050106

Cordeiro, J. R., & Postolache, O. (2018). Big Data Storage for a Health Predictive System. *2018 International Symposium in Sensing and Instrumentation in IoT Era, ISSI 2018*, 1–6. https://doi.org/10.1109/ISSI.2018.8538285

Dias, L. B., Holanda, M., Huacarpuma, R. C., & De Sousa, R. T. (2018). NoSQL database performance tuning for IoT data cassandra case study. *IoTBDS 2018 - Proceedings of the 3rd International Conference on Internet of Things, Big Data and Security*, *2018-March*(IoTBDS 2018), 277–284. https://doi.org/10.5220/0006782702770284

Faridoon A., & Imran A. (2021). Big data storage tools using NoSQL databases and their applications in various domains: a systematic review. *Computing and Informatics*, *40*, 489–521. https://doi.org/10.31577/cai_2021_3_489

Ferencz, K., & Domokos, J. (2019). IoT Sensor data acquisition and storage system using raspberry Pi and apache cassandra. *CANDO-EPE 2018 - Proceedings IEEE International Conference and Workshop in Obuda on Electrical and Power Engineering*, 143–146. https://doi.org/10.1109/CANDO-EPE.2018.8601139

Gamal, A., Barakat, S., & Rezk, A. (2021). Integrated Document-based Electronic Health Records Persistence Framework. *International Journal of Advanced Computer Science and Applications*, *12*(10), 147–155. https://doi.org/10.14569/IJACSA.2021.0121017

Ghiwari, S., Sambrekar, K., & Rajpurohit, V. S. (2018). Hierarchical Storage for Agro Informatics System Using NoSQL Technology. *2017 International Conference on Computing, Communication, Control and Automation, ICCUBEA 2017*, 1–5. https://doi.org/10.1109/ICCUBEA.2017.8463693

Hak, F., Guimarães, T., Abelha, A., & Santos, M. (2020). An Exploratory Study of a NoSQL Database for a Clinical Data Repository. *Advances in Intelligent Systems and Computing*, *1161 AISC*, 476–483. https://doi.org/10.1007/978-3-030-45697-9_46

Hendawi, A., Gupta, J., Liu, J., Teredesai, A., Ramakrishnan, N., Shah, M., El-Sappagh, S., Kwak, K. S., & Ali, M. (2019). Benchmarking large-scale data management for Internet of Things. *Journal of Supercomputing*, *75*(12), 8207–8230. https://doi.org/10.1007/s11227-019-02984-6

J. Aqel, M., Al-Sakran, A., & Hunaity, M. (2019). A Comparative Study Of NoSQL Databases. *Bioscience Biotechnology Research Communications*, *12*(1), 17–26. https://doi.org/10.21786/bbrc/12.1/3

Khan, W., Ahmad, W., Luo, B., & Ahmed, E. (2019). SQL database with physical database tuning technique and NoSQL graph database comparisons. *Proceedings of 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference, ITNEC 2019*, *Itnec*, 110–116. https://doi.org/10.1109/ITNEC.2019.8729264

Khan, W., ahmed, E., & Shahzad, W. (2017).

Predictive Performance Comparison Analysis of Relational & NoSQL Graph Databases. *International Journal of Advanced Computer Science and Applications*, *8*(5). https://doi.org/10.14569/ijacsa.2017.080564

Kitchenham, B. (2004). Procedures for Performing Systematic Reviews. *Keele University Technical Report*, 33(2004), 1-26. https://www.inf.ufsc.br/~aldo.vw/kitchenham.pdf

Klein, J., Gorton, I., Ernst, N., Donohoe, P., Pham, K., & Matser, C. (2015). Performance evaluation of NoSQL databases: A case study. *PABS 2015 - Proceedings of the 1st ACM/SPEC International Workshop on Performance Analysis of Big Data Systems*, *February*, 5–10. https://doi.org/10.1145/2694730.2694731

Küçükkeçeci, C., & Yazıcı, A. (2018). Big Data Model Simulation on a Graph Database for Surveillance in Wireless Multimedia Sensor Networks. *Big Data Research*, *11*(September), 33–43. https://doi.org/10.1016/j.bdr.2017.09.003

Kundu, G., Mukherjee, N., & Mondal, S. (2021). Building a Graph Database for Storing Heterogeneous Healthcare Data. *Smart Innovation, Systems and Technologies*, *196*, 193–201. https://doi.org/10.1007/978-981-15-7062-9_19

Magdum, J., & Barhate, R. (2018). Performance Analysis of DML Operations on NoSQL Databases for Streaming Data. *Proceedings - 2018 4th International Conference on Computing, Communication Control and Automation, ICCUBEA 2018*, 1–6. https://doi.org/10.1109/ICCUBEA.2018.8697482

Mahmood, K., Orsborn, K., & Risch, T. (2019). Comparison of NoSQL datastores for large scale data stream log analytics. *Proceedings - 2019 IEEE International Conference on Smart Computing, SMARTCOMP 2019*, 478–480. https://doi.org/10.1109/SMARTCOMP.2019.00093

Naik, K., & Joshi, A. (2017). Role of Big Data in various sectors. *Proceedings of the International Conference on IoT in Social, Mobile, Analytics and Cloud, I-SMAC 2017*, 117–122. https://doi.org/10.1109/I-SMAC.2017.8058321

Nassif, E. H., Hicham, H., Yaagoubi, R., & Badir, H. (2020). Assessing NoSql Approaches for Spatial Big Data Management. *Lecture Notes in Networks and Systems*, *92*, 49–58. https://doi.org/10.1007/978-3-030-33103-0_6

Philip Chen, C. L., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, *275*, 314–347. https://doi.org/10.1016/j.ins.2014.01.015

Pramukantoro, E. S., Yahya, W., Arganata, G., Bhawiyuga, A., & Basuki, A. (2017). Topic Based IoT Data Storage Framework For Heterogeneous Sensor Data. *2017 11th International Conference on Telecommunication Systems Services and Applications (TSSA)*, 1-4. https://doi.org/ 10.1109/TSSA.2017.8272895

Pramukantoro, E. S., Primanita Kartikasari, D., & Siregar, R. A. (2019). Performance evaluation of MongoDB, cassandra, and HBase for heterogenous IoT data storage. *Proceedings of ICAITI 2019 - 2nd International Conference on Applied Information Technology and Innovation: Exploring the Future Technology of Applied Information Technology and Innovation*, 203–206. https://doi.org/10.1109/ICAITI48442.2019.8982159

Rasheed, Y., Qutqut, M. H., & Almasalha, F. (2019). Overview of the Current Status of NoSQL Database. *IJCSNS International Journal of Computer Science and Network Security*, *19*(4), 47-53. http://paper.ijcsns.org/07_book/201904/20190406.pdf

Reichardt, M., Gundall, M., & Schotten, H. D. (2021). Benchmarking the Operation Times of NoSQL and MySQL Databases for Python Clients. *IECON Proceedings (Industrial Electronics Conference)*, *2021-Octob*(October). https://doi.org/10.1109/IECON48115.2021.9589382

Sharma, M., & Kaur, G. (2019). Providing Low Cost IoT Enabled Security Solutions for High Intensity Regions. *2019 12th International Conference on Contemporary Computing, IC3 2019*, 1–7. https://doi.org/10.1109/IC3.2019.8844949

Shen, Y. (2021). Construction of a Wireless Sensing Network System for Leisure Agriculture for Cloud-Based Agricultural Internet of Things. *Journal of Sensors*, *2021*. https://doi.org/10.1155/2021/3021771

Silver, N., Ettinger, K., Kolden, J., Malgeri, M., & Weaver, E. (2016). "The suitcase": Using C4, semantics and NoSQL for managing motion picture data. *SMPTE 2016 - Annual Technical Conference and Exhibition, 2016-Janua*(June), 1–15. https://doi.org/10.5594/m001715

Stonebraker, M. (2010). SQL databases v. NoSQL databases. *Communications of the ACM*, *53*(4), 10–11. https://doi.org/10.1145/1721654.1721659

Strauch, C. (2011). NOSQL databases. In *Advanced Information and Knowledge Processing*. Lecture Notes, Stuttgart Media University, vol. 20. https://doi.org/10.1007/978-3-030-04603-3_8

Subahi, A. F. (2019). Edge-Based IoT Medical Record System: Requirements, Recommendations

and Conceptual Design. *IEEE Access*, *7*, 94150–94159. https://doi.org/10.1109/ACCESS.2019.2927958

Tokognon, A. C., Gao, B., Tian, G. Y., & Yan, Y. (2017). Structural Health Monitoring Framework Based on Internet of Things: A Survey. *IEEE Internet of Things Journal*, *4*(3), 619–635. https://doi.org/10.1109/JIOT.2017.2664072

Vanathi, R., & Shaik Abdul Khadir, A. (2017). A Robust Architectural Framework for Big Data Stream Computing in Personal Healthcare Real Time Analytics. *Proceedings - 2nd World Congress on Computing and Communication Technologies, WCCCT 2017*, 97–104. https://doi.org/10.1109/WCCCT.2016.32

Vela, B., Cavero, J. M., Caceres, P., Sierra, A., & Cuesta, C. E. (2018). Defining a NoSQL Document Database of Accessible Transport Routes. *Proceedings - 2017 IEEE International Conference on Internet of Things, IEEE Green Computing and Communications, IEEE Cyber, Physical and Social Computing, IEEE Smart Data, IThings-GreenCom-CPSCom-SmartData 2017*, *2018-Janua*, 1125–1129. https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2017.172

Venkatraman, S., Overmars, A., Fahd, K., Parvin, S., & Kaspi, S. (2020). Security Challenges for Big Data and IoT. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3378904.3378907

Wang, S., Zhu, Y., Wang, F., & Qian, Z. (2017). The selection of NoSQL database in E-government system of China meteorological administration. *2017 IEEE 2nd International Conference on Big Data Analysis, ICBDA 2017*, 105–109. https://doi.org/10.1109/ICBDA.2017.8078786

Wu, S., Bao, L., Zhu, Z., Yi, F., & Chen, W. (2018). Storage and retrieval of massive heterogeneous IoT data based on hybrid storage. *ICNC-FSKD 2017 - 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery*, 2982–2987. https://doi.org/10.1109/FSKD.2017.8393258

Appendix

No	Reference	No	Reference
1	(Amghar et al., 2018)	25	(J. Aqel et al., 2019)
2	(Ghiwari et al., 2018)	26	(Nassif et al., 2020)
3	(Wang et al., 2017)	27	(Almeida et al., 2020)
4	(Celesti, et al., 2020b)	28	(Küçükkeçeci & Yazıcı, 2018)
5	(Subahi, 2019)	29	(Vela et al., 2018)
6	(Hendawi et al., 2019)	30	(Gamal et al., 2021)
7	(Rasheed et al., 2019)	31	(Magdum & Barhate, 2018)
8	(Celesti et al., 2018)	32	(Ferencz & Domokos, 2019)
9	(Vanathi & Shaik Abdul Khadir, 2017)	33	(Asiminidis et al., 2019)
10	(Braulio et al., 2018)	34	(Dias et al., 2018)
11	(Pramukantoro et al., 2019)	35	(Venkatraman et al., 2020)
12	(Mahmood et al., 2019)	36	(Cerbulescu, C. & Cerbulescu, C. M., 2017)
13	(Celesti, et al., 2020a)	37	(Brown, 2020)
14	(Chen & Lee, 2019)	38	(Tokognon et al., 2017)
15	(Al Sadi et al., 2017)	39	(Sharma & Kaur, 2019)
16	(Ai-Sakran et al., 2018)	40	(Hak et al., 2020)
17	(Klein et al., 2015)	41	(Silver et al., 2016)
18	(Shen, 2021)	42	(Kundu et al., 2021)
19	(Reichardt et al., 2021)	43	(Chauhan & Malik, 2021)
20	(Khan et al., 2019)	44	(Faridoon & Imran, 2021)
21	(Cordeiro & Postolache, 2018)	45	(Philip Chen & Zhang, 2014)
22	(Wu et al., 2018)	46	(Naik & Joshi, 2017)
23	(Pramukantoro et al., 2017)	47	(Stonebraker, 2010)
24	(Khan et al., 2017)	48	(Kitchenham, 2004)

Global Repeat Map (GRM) algorithm as genomic technology for Higher Order Repeat identification

Algoritam globalne karte ponavljanja (Global Repeat map – GRM algorithm) kao genomska tehnologija za identifikaciju periodičnosti višeg reda

Ines Vlahović
Algebra University College
Visoko učilište Algebra
Gradišćanska 24, Zagreb
Faculty of Science, Department of Physics
Prirodoslovno-matematički fakultet, Fizički odsjek
University of Zagreb
Sveučilište u Zagrebu
Bijenička cesta 32, Zagreb
ines.vlahovic@algebra.hr

Abstract. *In recent years new technologies for DNA sequence widened research of the roles of tandem repeats in eukaryotes, especially ones that are located in centromere region in chromosomes because of their importance for chromosome segregation, microtubule attachment to kinetochores and, as recently showed, their role in some diseases such as cancers. For identification of tandem repeats and higher order repeats (HORs) we use Global Repeat Map (GRM) algorithm and AlphaSub (ASu) algorithm, developed at Department of Physics, Faculty of Science, University of Zagreb. Here we present some illustrative results obtained using this tool. Based on analysis of DNA sequences, we create a database including information on location, length and divergence for alpha satellite arrays and HORs in human, Neanderthal and chimpanzee genomes (at link http://genom.hazu.hr). We summarize analysis of alpha satellite HORs for human Y chromosome, as well as for some other HORs identified in human and other primates' chromosome 1 with novel comparison of ~1.6 kb monomer units in them. We also show preliminary results of GRM analysis of three breast cancer exomes. In this respect, we emphasize a need for employing new sequencing methods to determine complete genome assemblies for identification of more precise HOR structures.*

Keywords. genomes, higher order repeat, tandem repeat, cancer, alpha satellites, NGS, human genome, Y chromosome, NBPF genes, molecular databases

Sažetak. *Posljednjih godina nove tehnologije za sekvenciranje DNK sekvenci proširile su istraživanja uloga tandemnih repeticija u eukariotima, osobito onih koji se nalaze u području centromere u kromosomima zbog njihove važnosti za segregaciju kromosoma, spajanju mikrotubuli na kinetohore i, kao što je nedavno pokazano, u ulogama u nekim bolestima poput karcinoma. Za identifikaciju tandemnih repeticija i periodičnosti višeg reda (Higher order repeats - HORs) koristimo u našim istraživanjima algoritam globalne karte ponavljanja (Global Repeat Map (GRM) algorithm) te AlphaSub (Asu) algoritam koji su razvijeni na Fizičkom odsjeku, Prirodoslovno-matematičkog fakulteta, Sveučilišta u Zagrebu. U ovom radu predstavljamo dio ilustrativnih rezultata dobivenih pomoću ovih alate. Temeljeno na analizi DNK sekvenci, stvorili smo bazu podataka koja uključuje informacije o pozicijama, duljinama i divergencijama za nizove alfa satelita i HORova za genome čovjeka, neandertalca i čimpanzu (na http://genom.hazu.hr). U ovom radu sumiramo analizu HORova zasnovanih na alfa satelitima za ljudski kromosom Y, kao i za neke druge HORove identificiranje u kromosomu 1 čovjeka te drugih primata s novom usporedbom ~1.6 kb monomerne jedinice u njima. Također, prikazati ćemo preliminarne podatke GRM analize za egzome tri slučaja karcinoma dojke. U tom smislu, naglašavamo potrebu za primjenom novih metoda sekvencioniranja za određivanje potpunih*

1 Introduction

In recent years new sequencing methods have emerged (NGS, nanopore sequencing, single molecule real time sequencing, optical mapping) that enabled collection of huge amounts of data from sequencing projects. With the increasing amount of collected data, repositories were created for storing DNA and protein sequences as well as tools for their exploration and analysis. Recent review paper with list of existing databases, ~1600 databases categorized in 15 categories and 43 subcategories, was published in which their overview was given (Rigden & Fernández, 2022). From collection of databases, NAR online Molecular Biology Database Collection, we use data from some relevant databases for our research, like NCBI, Ensemble, ClinVar, GWAS, TCGA, NCI GDC. Although those data are mostly used in investigation of genetic content, a few are oriented toward repetitive DNA sequences (TRDB, STRBase, Plant Repeat Database, RepBase). For years, the repetitive sequences were considered as "junk" elements, but recently there are more and more evidence that they might have an important role and could be responsible for some diseases such as Huntington, schizophrenia and even cancer [Fondon & Garner, 2004, Glunčić, Vlahović, & Paar, 2019]. Repetitive sequences can emerge in DNA sequence as tandem repeats, dispersed repeats, segmental duplications, complex repeats, and higher order repeats. Some of our main results obtained with GRM algorithm and AlphaSub algorithm belong to the analysis of satellite DNA i.e., research of human alpha satellites of ~171 bp monomer unit length that form high range of HORs – like the one in chromosome 21, (34/36mer HOR) that is largest among human somatic chromosomes, and HORs in human Y chromosome. In addition, our algorithms are very effective with identification of repeats for long monomer units like NBPF monomers with length of ~1.6 kb (classified as mega satellites) as well as smaller monomers (mini and micro satellites in hornerin gene) that are used for DNA fingerprinting. Fig.1 shows a more detailed classification according to ref. (Pathak & Ali, 2012).

In human genome, repetitive elements occupy ~50% of DNA sequence (Glunčić & Paar, 2013). It is shown that tandem repeats have roles such as markers for genotyping a number of pathogens (Flèche, 2001), that copy number variation can influence molecular paths in individuals and that they could also be a reason for phenotype changes in population (Romero, Hosomichi, Nakaoka, Shibata, & Inoue, 2017). One of the most investigated tandem repeats are human centromeric alpha satellites with length of ~171 bp monomer unit that can form higher order repeats (HORs) although their exact schemes with variant copies are not given. Satellites are important for processes of chromosome segregation in cell division which might, if disrupted, cause

genomskih sekvenci za identifikaciju preciznije opisanih HOR struktura.

Ključne riječi. genomi, periodišnosti višeg reda, tandemsko ponavljanje, karcinom, alfa sateliti, NGS, ljudski genom, Y kromosom, NBPF geni, molekularne baze podataka

1 Uvod

Posljednjih godina pojavile su se nove metode sekvenciranja (NGS, sekvenciranje nanopora, sekvenciranje pojedinačne molekule u stvarnom vremenu, optičko mapiranje) koje su omogućile prikupljanje ogromnih količina podataka iz projekata sekvenciranja. S povećanjem količine prikupljenih podataka stvoreni su repozitoriji za pohranjivanje sekvenci DNK i proteina kao i alati za njihovo istraživanje i analizu. Nedavno je objavljen pregledni rad s popisom postojećih baza podataka, ~1600 baza podataka kategoriziranih u 15 kategorija i 43 potkategorija, u kojem je dan njihov pregled (Rigden & Fernández, 2022). Iz zbirke baza podataka, NAR online Molecular Biology Database Collection, koristimo neke relevantne baze podataka u našim istraživanjima, kao što su NCBI, Ensemble, ClinVar, GWAS, TCGA, NCI GDC. Iako se podaci iz tih baza uglavnom koriste u istraživanjima gena, neke od baza su orijentirane na ponavljajuće sekvence DNK (TRDB, STRBase, Plant Repeat Database, RepBase). Godinama su se ponavljajuće sekvence smatrale elementima "smeća", no nedavno je sve više dokaza da bi upravo one mogle imati važnu ulogu te da bi one mogle biti odgovorne za neke bolesti poput Huntingtonove, shizofrenije, pa čak i karcinoma [Fondon & Garner, 2004, Glunčić, Vlahović, & Paar, 2019]. Repeticije sekvenci DNK se mogu pojaviti kao tandemna ponavljanja, disperzirana ponavljanja, segmentne duplikacije, složena ponavljanja te periodičnosti višeg reda (HOR). Neki od naših glavnih rezultata dobivenih s GRM metodom i AlphaSub algoritmom pripadaju analizi satelitske DNA tj. istraživanjima ljudskih alfa satelita s ~171 bp monomernom jedinicom koja može tvoriti veliki raspon HORova – npr. u kromosomu 21 nalazi se najveći među somatskim kromosomima 34/36mer HOR, te HOR strukture u ljudskom Y kromosomu. Dodatno, naši algoritmi su vrlo učinkoviti za repeticije s duljim monomerskim jedinicama poput NBPF monomera duljine od ~1.6 kb (klasificirani kao megasateliti) ali i za manje monomerne jedinice (kao što su mikro i mini sateliti u hornerin genu) koji se koriste u metodama kao što je DNK otisak prsta. Slika 1 prikazuje detaljniju klasifikaciju prema izvoru (Pathak & Ali, 2012).

U ljudskom genomu repeticije zauzimaju ~50% sekvence DNK (Glunčić & Paar, 2013). Pokazalo se da tandemne repeticije mogu poslužiti kao markeri za genotipizaciju brojnih patogena (Flèche, 2001) te

missegregation, aneuploidy and cancer [A.K., Mourad, Kaplan, & et al., 2019, Levy-Sakin, Pastor, Mostovoy, & et al., 2019, Black & Giunta, 2018].

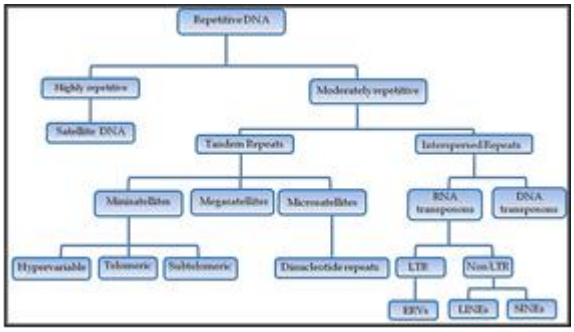
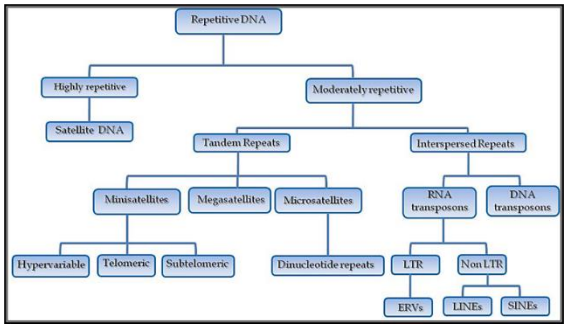


Figure 1. Repetitive elements classification from (Pathak & Ali, 2012). In this paper, we analyze satellite DNA, tandem repeats and HORs in primates for population studies (NBPF copy number in primates, Neanderthal and human genomes and example of hornerin gene in chromosome 1) and breast cancer exomes with our GRM and AlphaSub algorithm.

We organise this paper in chapters 2 Methods, 3 Results and 4 Conclusion. In the chapter Methods, we describe Global Repeat Map algorithm and AlphaSub algorithm for identification of tandem repeats and alpha satellite sequences. In 2.1. we give links to data used for our research and which databases store those datasets. In chapter 3, Results, we show analysis of human Y chromosome, chromosome 21, primates and human chromosome 1 and novel comparison of NBPF (Neuroblastoma BreakPoint Family) monomer, dimer and trimer repeat units among human, primates and Neanderthal genomes. As novel preliminary result, we show GRM diagram for three cases of whole exome sequences of breast cancer from GDC data portal from CCLE project. One possible way for using GRM analysis could be for diagnostic purposes and classification of cancer exome sequences or whole genomes from various available datasets (from projects such as TCGA, CCLE). Our chapter 4 Conclusion emphasize the need for applying novel sequencing methods in whole genome projects, especially for population studies and studies of different diseases (cancer, autism).

Related work necessary for this research paper included investigation of databases and their usage, as well as some additionally tools and algorithms implemented for DNA analysis like Blast algorithm and tools like UGENE for making consensus sequence in fasta format. The most common database for genomic sequence that we use is NCBI that host a vast number of sequenced species, including viruses, bacteria and eukaryotes that enable us to conduct population studies. Blast algorithm enabled us to gain a novel insight in sequence evolution and their emergence in primate genomes like our novel

da varijacije broja kopija u njima mogu utjecati na određene molekularne putove kod pojedinaca. Također tandemna ponavljanja bi mogla biti uzrok i u promjenama fenotipa u populaciji (Romero, Hosomichi, Nakaoka, Shibata & Inoue, 2017). Jedna grupa najistraživijih tandemskih repeticija su centromerni alfa sateliti duljine ~171 bp koje se mogu dodatno pojaviti i kao periodičnosti višeg reda (HOR) iako do sada njihove egzaktne sheme s varijantnim kopijama nisu dane. Sateliti su važni za procese segregacije kromosoma u staničnoj diobi koji bi, ako se poremete, mogli uzrokovati pogrešnu segregaciju, aneuploidiju i karcinome [A.K., Mourad, Kaplan & sur., 2019, Levy-Sakin, Pastor, Mostovoy & sur., 2019, Black & Giunta, 2018].



Slika 1. Klasifikacija repeticijskih elemenata u DNK (Pathak & Ali, 2012). U ovom radu analiziramo satelitsku DNK, tandemna ponavljanja i HOR-ove kod primata za populacijske studije (broj kopija NBPF-a kod primata, genoma neandertalaca i ljudi te primjer hornerin gena u kromosomu 1) i egzome karcinoma dojke s našim GRM i AlphaSub algoritmom.

Ovaj rad je strukturiran kroz poglavlja 2 Metode, 3 Rezultati i 4 Zaključak. U poglavlju Metode, dajemo opis GRM algoritma i algoritma AlphaSub za identifikaciju tandemnih repeticija te sekvenci alfa satelita. U 2.1. dajemo poveznice na podatke koje koristimo u našem istraživanju i podatke o bazama podataka koji te podatke pohranjuju. U poglavlju 3, Rezultati, prikazujemo analizu ljudskog Y kromosoma, kromosoma 21, kromosoma 1 za čovjeka i primate te novu usporedbu NBPF (Neuroblastoma Break Point Family) monomernih, dimerskih i trimerskih repeticijskih jedinica među genomima čovjeka, neandertalca i primata. Kao novi preliminarni rezultati, također prikazujemo GRM dijagram za tri slučaja karcinoma dojke s GDC podatkovnog portala iz CCLE projekta. Jedan mogući način korištenja GRM algoritma za analize može poslužiti u dijagnostici i klasifikaciji sekvenci egzoma karcinoma ili cijelih genoma iz različitih dostupnih podatkovnih setova (iz projekta TCGA, CCLE). U poglavlju 4 Zaključak naglašavamo potrebu za primjenom novih metoda sekvenciranja u projektima cijelih genoma, posebno u svrhu

results for NBPF gene family. Our interpretation of those novel results takes in consideration assemblies used, because of percentage of completeness of genomic sequence and sequencing techniques used. We stress out that combination of different sequencing techniques should fill sequence gaps in the future. Filling the gaps in genomic sequence will enable us to create a more complete picture of tandem repeats and HORs, especially in centromeric region that is responsible for correct chromosome segregation like for human Y, 1 and 21 chromosomes presented in this paper.

Also, the existing databases that store genome and exome data for different diseases (TCGA, CCLE, TARGET, dbGaP, ClinVar, COSMIC, MSSNG, et cetera) could help us to understand progression of diseases and their connection to structural variants and copy number variants. Examples of those kind of diseases are autism, macrocephaly, microcephaly and neuroblastoma that are mentioned earlier and they are connected to NBPF monomer indels (deletions, insertions).

Complete schemes of structural variants and polymorphic copies in HORs could be used as a diagnostic tool or for population studies. In that regard tools such as our GRM and AlphaSub algorithm can be used, but also novel algorithms based on machine learning and AI should be developed that could bring new insight in pattern recognition in studying DNA from large data sets.

2 Methods

For investigation of repeats in DNA sequences of different eukaryotes, we use Global Repeat Map algorithm developed at University of Zagreb, Faculty of Science, Physics department. This method was explained in details in ref. (Glunčić & Paar, 2013), and here we summarize the main points characterizing this tool:

1. It directly maps the DNA symbolic sequence into the frequency domain making „GLOBAL MAP" based on combination of statistical and digital processing methods (Glunčić & Paar, 2013).
2. It uses a complete k-word ensemble (global - local) for creation of global map (Fig. 2).
3. It is parameter – free.
4. It can identify repetitions of all lengths.
5. It is robust to copy deviations from the perfect sample.
6. It identifies higher-order repeats (HOR).
7. We can simply determine consensus lengths and sequences from GRM results.
8. It is „good" for use in combination with BLAST for similar sequence.

This tool is free for download at http://genom.hazu.hr/tools.html. On given link instructions for its usage are given with examples.

populacijskih istraživanja i istraživanja raznih bolesti (karcinomi, autizam).

Povezani rad neophodan za ovo istraživanje uključuje istraživanje baza podataka i njihove upotrebe, kao i neke dodatne alate i algoritme implementirane za analizu DNK kao što je Blast algoritam i alate kao što je UGENE za izradu konsenzusnog sekvence u fasta formatu. Najčešća baza podataka za genomsku sekvencu koju koristimo je NCBI koja sadrži veliki broj sekvenciranih vrsta, uključujući viruse, bakterije i eukariote koji nam omogućuju provođenje populacijskih istraživanja. BLAST algoritam nam je omogućio da steknemo novi uvid u evoluciju sekvenci i njihovu pojavu u genomima primata, poput naših novih rezultata za obitelj gena NBPF. Naše tumačenje tih novih rezultata uzima u obzir korištene sklopove genomskih sekvenci (assembly), zbog postotka potpunosti genomske sekvence i korištenih tehnika sekvenciranja. Naglašavamo da bi kombinacija različitih tehnika sekvenciranja trebala popuniti praznine u sekvencama u budućnosti. Popunjavanje praznina u genomskoj sekvenci omogućit će nam stvaranje cjelovitije slike tandemnih ponavljanja i HORova, posebno u centromernom području koje je odgovorno za ispravnu segregaciju kromosoma kao što pokazuju primjeri za ljudske Y, 1 i 21 kromosome koji su predstavljeni u ovom radu.

Također, postojeće baze podataka koje pohranjuju podatke o genomu i egzomu za različite bolesti (TCGA, CCLE, TARGET, dbGaP, ClinVar, COSMIC, MSSNG, itd.) mogle bi nam pomoći da razumijemo progresiju bolesti i njihovu povezanost sa strukturnim varijantama i varijantama broja kopija.

Primjeri takvih bolesti su autizam, makrocefalija, mikrocefalija i neuroblastom koji su ranije spomenuti i povezani su s NBPF monomernim indelima (delecije, insercije).

Potpune sheme strukturnih varijanti i polimorfnih kopija u HOR-ovima mogu se koristiti kao dijagnostički alat ili za populacijske studije. U tom smislu mogu se koristiti alati kao što su naš algoritam GRM i AlphaSub, ali također bi se trebali razviti novi algoritmi temeljeni na strojnom učenju i umjetnoj inteligenciji koji bi mogli donijeti novi uvid u prepoznavanje uzoraka u proučavanju DNK sekvenci dobivenih iz velikih skupova podataka.

2 Metode

Za istraživanje repeticija u sekvencama DNK različitih eukariota koristimo algoritam globalne karte ponavljanja - Global Repeat Map (GRM) razvijen Fizičkom odsjeku, Prirodoslovno-matematičkog fakulteta, Sveučilišta u Zagrebu. Ova metoda je detaljno objašnjena u izvoru (Glunčić & Paar, 2013), a ovdje sažimamo glavne prednosti koje karakteriziraju ovaj alat:

Example of using GRM method is shown in Fig.2 where we analysed satellites (with monomer unit lengths in range from ~331 bp to ~369 bp) from beetle T.Castaneum genome (Vlahovic, Gluncic, Rosandic, Ugarkovic, & Paar, 2017).

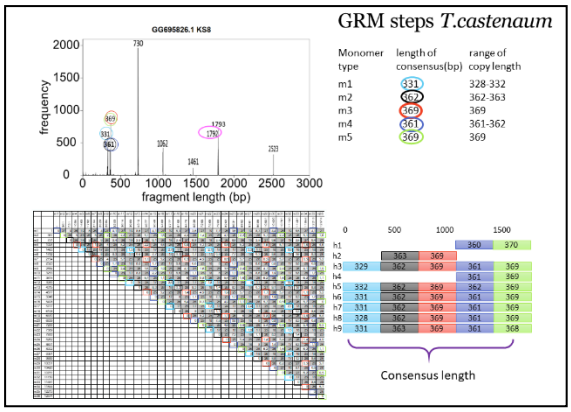


Figure 2. GRM diagram for Tcast satellites show monomer units that are organized in five monomer types as GRM diagram shows (upper left). Needleman–Wunsch algorithm is applied to monomer DNA sequence for their classification into different monomer types according to sequence divergence. Divergence between monomers in array reveals higher order structure of length ~1792 bp i.e., 5mer HOR. Monomers that belong to the same type (with low divergence) have same color in Needleman-Wunsch matrix. This result/example is published in (Vlahovic, Gluncic, Rosandic, Ugarkovic, & Paar, 2017). Currently we are developing novel algorithms for automatization of this process.

We developed and used for automatization of alpha satellites identification in chromosomes additional algorithm AlphaSub, described in ref. (Glunčić, Vlahović, & Paar, 2019). This algorithm is based on 28 bp fragment of alpha satellite consensus sequence "TGAGAAACTGCTTTGTGATGTGTGCATT" as ideal "key word" as well as its reverse complement and by using Levenshtein distance algorithm we identify position of those alpha satellites ~171 bp in chromosome DNA sequence. We also used this method for analysis of human chromosome Y as well as whole human, Neanderthal and chimpanzee genomes. Data for those genomes and their alpha satellite arrays can be found at http://genom.hazu.hr. This tool can be downloaded from the link http://genom.hazu.hr/tools.html.

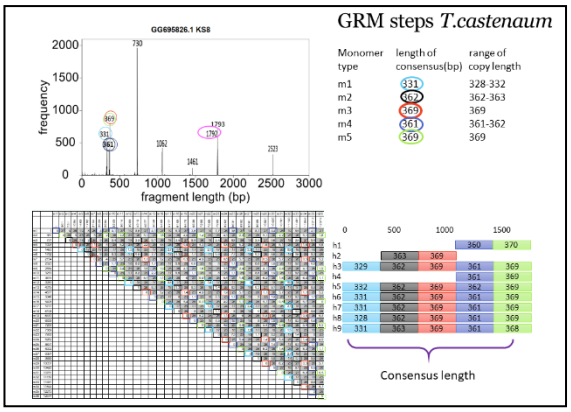
2.1 *Data*

Data used for our research are mostly downloaded from NCBI database ftp site as one of the most used databases in this field of research – NCBI (National Center for Biotechnology Information,

1. Izravno preslikava simboličku sekvencu DNK u frekvencijsku domenu čineći „GLOBALNU MAPU" na temelju kombinacije statističkih i digitalnih metoda obrade signala (Glunčić & Paar, 2013).
2. Koristi kompletan skup k-riječi (globalno - lokalno) za izradu globalne mape (Slika 2).
3. Nema ulaznih dodatnih parametara.
4. Može identificirati repeticije svih duljina.
5. Robustan je za odstupanja repeticija od savršenog uzorka.
6. Identificira periodičnosti višeg reda (HOR).
7. Jednostavno se određuju konsenzusne duljine repeticija i nizovi sekvenci iz GRM rezultata.
8. „Dobar" je za korištenje u kombinaciji s algoritmom BLAST za traženje sličnih nizova.

Ovaj je alat besplatan za preuzimanje na http://genom.hazu.hr/tools.html. Na navedenom linku dane su upute za njegovu uporabu s primjerima.

Primjer korištenja GRM metode prikazan je na Slici 2 gdje smo analizirali satelite (s duljinama monomera u rasponu od ~331 bp do ~369 bp) iz kukca *T.Castaneum* (Vlahovic, Gluncic, Rosandic, Ugarkovic, & Paar, 2017).



Slika 2. GRM dijagram za Tcast satelite koji su organizirani u pet vrsta monomera (primarnih repeticijskih jedinica) kao što prikazuje GRM dijagram (gore lijevo). Needleman-Wunsch algoritam primijenjen je na sekvencu monomerne DNA za njihovu klasifikaciju u različite tipove prema međusobnoj divergenciji monomera. Divergencija između monomera u nizu alfa satelita otkriva strukturu višeg reda duljine ~1792 bp, tj. 5mer HOR. Monomeri koji pripadaju istom tipu (s malom divergencijom) iste su boje u Needleman-Wunsch matrici. Ovaj rezultat je objavljen u radu (Vlahović, Glunčić, Rosandić, Ugarković & Paar, 2017). Trenutno razvijamo nove algoritme za automatizaciju ove procedure.

Za automatizaciju identifikacije alfa satelita u kromosomima razvili smo i koristili dodatni algoritam ALPHAsub, opisan u (Glunčić, Vlahović & Paar, 2019). Ovaj se algoritam temelji na 28 bp fragmentu konsenzusne sekvence alfa satelita

https://ftp.ncbi.nlm.nih.gov) for human and primate genomes. We also give links to Nacional Cancer Institute Genomic Data Commons Data Portal – NCI GDC portal (https://portal.gdc.cancer.gov) with cancer databases. NCI GDC portal includes most commonly used cancer databases such as TCGA (The Cancer Genome Atlas), Target (Therapeutically Applicable Research to Generate Effective Therapies) i.e., they host around 72 projects with 828104 files. It also includes Cancer Cell Line Encyclopaedia Project (CCLE) data for breast cancer sequences that we used for our novel preliminary data analysis shown in this paper.

Additional DNA sequence for GRM analysis of chromosome 1 (assembled chromosomes) in *Pan Troglodytes, Gorilla gorilla, Pongo abelii, Papio anubis, Pan paniscus, Macaca fascicularis, Macaca mulatta*, are downloaded from NCBI ftp site:

- *Pan troglodytes* - Pan_tro_3.0 (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001515.7/)
- ftp.ncbi.nih.gov/genomes/Gorilla_gorilla /Assembled_chromosomes/seq/
- ftp.ncbi.nih.gov/genomes/Pongo_abelii/ Assembled_chromosomes/seq/
- ftp.ncbi.nih.gov/genomes/Papio_anubis/ Assembled_chromosomes/seq/
- ftp.ncbi.nih.gov/genomes/Pan_paniscus/ Assembled_chromosomes/seq/
- ftp.ncbi.nih.gov/genomes/Macaca_fascicularis/ Assembled_chromosomes/seq
- ftp.ncbi.nih.gov/genomes/Macaca_mulatta/ Assembled_chromosomes/seq/
- https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.26/ (Human hg38)
- Janet Kelso group from Max Planck institute, Department of Evolutionary Genetics in Leipzig, provided DNA sequence for Neanderthal genome. Altai_Neandertal_ERP002097.
- C835.HCC1143_BL.4.bam (https://portal.gdc.cancer.gov/legacy-archive/files/7bf5df2d-0859-43bd-8d2f-b0d584d2bf47).
- C836.Hs_606.T.1.bam (https://portal.gdc.cancer.gov/legacy-archive/files/49a2f622-ee79-43fd-a593-0458a908acc7).
- C836.Hs_343.T.1.bam (https://portal.gdc.cancer.gov/legacy-archive/files/45dd789e-943f-4b25-b4b2-defbdf951ee1).

For our analysis purpose, we first download fasta sequence that we use for GRM analysis for obtaining global repeat map – GRM diagram that show us repeat unit lengths and possible HOR structures (Fig.2). For investigation of alpha satellite sequences, we additionally use AlphaSub algorithm. For breast cancer genomes, we first download raw sequence in BAM format after which we construct "TGAGAAACTGCTTTGTGATGTGTGCATT" kao idealna "ključna riječ" kao i njen obrnuti komplement te u kombinaciji s Levenshteinovim algoritmom udaljenosti identificiramo položaj alfa satelita ~171 bp u sekvenci DNK kromosoma.

Ovu smo metodu također koristili za analizu ljudskog kromosoma Y kao i cijelih genoma čovjeka, neandertalca i čimpanze. Podaci za te genome i njihove alfa satelitske nizove mogu se pronaći na http://genom.hazu.hr. Ovaj alat može se preuzeti s poveznice http://genom.hazu.hr/tools.html.

2.1 Podatci

Podaci korišteni za naše istraživanje većinom su preuzeti s ftp stranice NCBI baze podataka kao jedne od najkorištenijih baza podataka u ovom području istraživanja – NCBI (National Center for Biotechnology Information, https://ftp.ncbi.nlm.nih.gov) za genome čovjeka i primata. Također, dajemo poveznice na Nacionalni institut za karcinome Genomic Data Commons Data Portal – NCI GDC portal (https://portal.gdc.cancer.gov) s bazama podataka o karcinomima. Portal NCI GDC uključuje najčešće korištene baze podataka o karcinomima kao što su TCGA (The Cancer Genome Atlas), Target (Therapeutically Applicable Research to Generate Effective Therapies) tj. u njima se nalazi oko 72 projekta s 828104 datoteka. Također uključuje podatke Projekta Enciklopedije stanične linije karcinoma (CCLE) za sekvence karcinoma dojke koje smo koristili za našu novu preliminarnu analizu prikazanu u ovom radu.

Dodatne sekvence DNK za GRM analizu kromosoma 1 (sastavljeni kromosomi) u *Pan Troglodytes, Gorilla gorilla, Pongo abelii, Papio anubis, Pan paniscus, Macaca fascicularis, Macaca mulatta*, preuzeti su s NCBI ftp stranice:

- *Pan troglodytes* - Pan_tro_3.0 (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001515.7/)
- ftp.ncbi.nih.gov/genomes/Gorilla_gorilla /Assembled_chromosomes/seq/
- ftp.ncbi.nih.gov/genomes/Pongo_abelii/ Assembled_chromosomes/seq/
- ftp.ncbi.nih.gov/genomes/Papio_anubis/ Assembled_chromosomes/seq/
- ftp.ncbi.nih.gov/genomes/Pan_paniscus/ Assembled_chromosomes/seq/
- ftp.ncbi.nih.gov/genomes/Macaca_fascicularis/ Assembled_chromosomes/seq
- ftp.ncbi.nih.gov/genomes/Macaca_mulatta/ Assembled_chromosomes/seq/
- https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.26/ (Human hg38)
- Janet Kelso group from Max Planck institute, Department of Evolutionary Genetics in Leipzig, provided DNA sequence for Neanderthal genome. Altai_Neandertal_ERP002097.

fasta consensus sequence using tool UGENE (http://ugene.net) (Okonechnikov, Golosova, Fursov, & Ugene team, 2012) and its algorithms. We afterwards analyse obtained fasta sequences with GRM algorithm for making global repeat map and analyse it further with procedure explained in Fig. 2. For search of NBPF monomers in primate's genomes, we additionally use BLAST algorithm (https://blast.ncbi.nlm.nih.gov /Blast.cgi).

3 Results

Studies of human Y chromosome show that it is replete with different kinds of repeats. The largest alpha satellite higher order repeat unit 34/36mer (with variant units of ~5.7 kb and ~6.0 kb) was identified previously [Tyler-Smith & Brown, 1987, Jain & et al., 2018], but its polymorphic variants have not been studied [Tyler-Smith & Brown, 1987, Jain & et al., 2018, Tyler-Smith & et al., 1993, Rozen & et al., 2003, Perry, Tito, & Verrelli, 2007]. Using GRM and ALPHAsub algorithm, we analysed Y chromosome from hg38 human genome assembly in which most of the centromere gaps have been filled (Uralsky & et al., 2019). We identify 28 alpha satellite arrays ranging from three monomers to 1334 monomers in individual array. Average divergence of monomers in those arrays is ~24%. Those alpha satellite arrays in human Y chromosome occupy 0.71% (407986 bp). Of these alpha satellite arrays, the three arrays constitute HORs located in three domains of Y chromosome (Fig. 3).

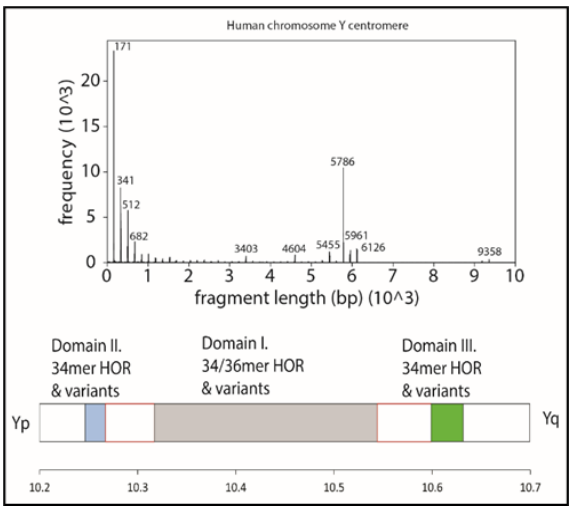


Figure 3. GRM diagram for centromere region and Y ideogram. Ideogram is modified from submitted paper (Vlahović, Glunčić, & Paar).

The novel corresponding computed GRM diagram for centromere region is presented in this paper and details of the variants are presented in (Vlahović, Glunčić, & Paar, (submitted for

- C835.HCC1143_BL.4.bam (https://portal.gdc.cancer.gov/legacy-archive/files/7bf5df2d-0859-43bd-8d2f-b0d584d2bf47).
- C836.Hs_606.T.1.bam (https://portal.gdc.cancer.gov/legacy-archive/files/49a2f622-ee79-43fd-a593-0458a908acc7).
- C836.Hs_343.T.1.bam (https://portal.gdc.cancer.gov/legacy-archive/files/45dd789e-943f-4b25-b4b2-defbdf951ee1).

U svrhu naše analize, prvo preuzimamo fasta sekvencu koju koristimo za dobivanje globalne mape ponavljanja – GRM dijagrama koji nam prikazuje duljine monomernih jedinica repeticija i moguće HOR strukture (Slika 2). Za istraživanje alfa satelitskih sekvenci dodatno koristimo AlphaSub algoritam. Za genome karcinoma dojke prvo preuzimamo "sirovu" sekvencu u BAM formatu nakon čega konstruiramo fasta konsenzusnu sekvencu pomoću alata UGENE (http://ugene.net) (Okonechnikov, Golosova, Fursov i Ugene tim, 2012) i njegovih algoritama. Nakon toga analiziramo dobivene fasta sekvence s GRM algoritmom za izradu mape globalnog ponavljanja i dalje podatke analiziramo postupkom objašnjenim na Slici 2. Za pretraživanje NBPF monomera u genomima primate dodatno koristimo BLAST algoritam (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

3 Rezultati

Studije ljudskog Y kromosoma pokazuju da je prepun različitih vrsta repeticija. Najveća ponovljena jedinica alfa satelita višeg reda 34/36mer (s varijantnim jedinicama od ~5,7 kb i ~6,0 kb) identificirana je ranije [Tyler-Smith & Brown, 1987, Jain & sur., 2018], ali njene polimorfne varijante nisu proučavane [Tyler-Smith & Brown, 1987, Jain & sur., 2018, Tyler-Smith & sur., 1993, Rozen & sur., 2003, Perry, Tito, & Verrelli, 2007]. Koristeći GRM i ALPHAsub algoritam, analizirali smo Y kromosom iz ljudskog genoma hg38 u kojem su većinom popunjene praznine centromera (Uralsky & sur., 2019). Identificirali smo 28 alfa satelitskih nizova u rasponu od 3 monomera do 1334 monomera u pojedinačnim nizovima. Prosječna divergencija monomera u tim nizovima je ~24%. Ti alfa satelitski nizovi u ljudskom Y kromosomu zauzimaju 0,71% (407986 bp). Od tih alfa satelitskih nizova, tri čine HORove smještene u tri domene na Y kromosomu (Slika 3).

publication)). These HORs are located in centromere and near centromere region (domains I., II. and III.). Domain I. constitutes the largest HOR array 34/36mer (positions 10316945-10544039 bp in the centromere). We identified polymorphic variants that have monomer duplications, deletions or insertions, or rearrangements and non-HOR insertions of up to 44-monomer length (details given in (Vlahović, Glunčić, & Paar, (submitted for publication))). These polymorphic variants are evident also from novel Needleman - Wunsch analysis of alpha satellite arrays in these three domains (dot plot in Fig 4. for domain II).

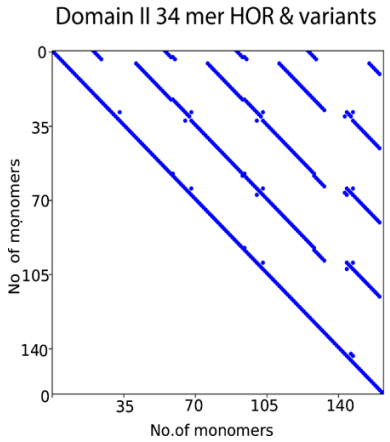
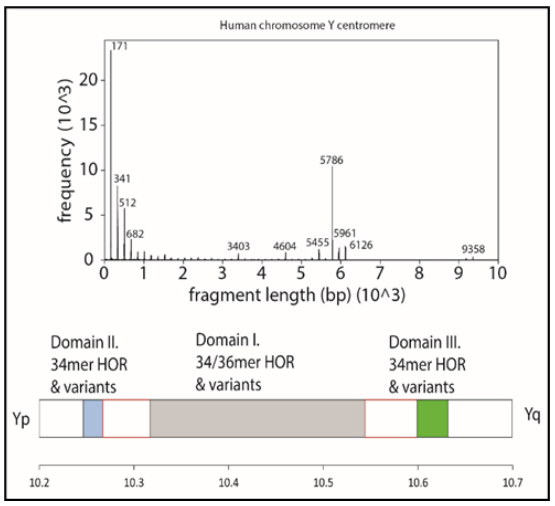


Figure 4. Needleman-Wunsch analysis of monomer divergence. Diagonals represent monomers with divergence less than 7%. Horizontal distances reveal nmer HOR copy that is in this case 34mer and its variants.

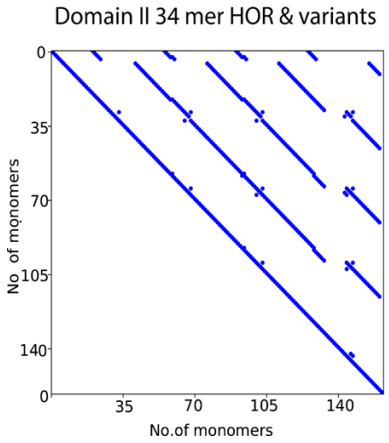
Using GRM and ALPHAsub algorithms, we analyse all alpha satellite arrays in human hg38 assembly as well as in chimpanzee and Neanderthal genomes (data stored in http://genom.hazu.hr database). These data on alpha satellite arrays and their positions within genomes provide a clue about their distribution and evolution between closely related species.

Calculation of percentage of alpha satellite content in genomes show that in chimpanzee (Pan_tro_3.0 assembly) they occupy 0.22%, in Neanderthal (Altai_Neandertal_ERP002097) 0.22% and in human genome (hg38) 2.19% of DNA sequence (Vlahović & et al., 2022). In human chromosomes 17-22 the percentage of alpha satellites is higher, ~5%. In our previous publication (Glunčić, Vlahović, & Paar, 2019) we analysed human chromosome 21 and we discovered the largest alpha satellite HOR unit among all somatic chromosomes - 33mer HOR, and that this chromosome is also replete with numerous other HOR arrays with monomer units composed of eight and more monomer types (Fig. 5). These HORs, based on alpha satellite arrays, could be used as unique chromosome 21 probe in molecular



Slika 3. GRM dijagram za područje centromere i Y ideogram. Ideogram je modificiran prema (Vlahović, Glunčić, & Paar, 2011, predano za objavu).

Novi odgovarajući izračunati GRM dijagram za područje centromera predstavljen je u ovom radu, a detalji varijanti prikazani su u (Vlahović, Glunčić, & Paar, (predano za objavu)). Ovi HOR-ovi se nalaze u području centromere i blizu centromere (domene I., II. i III. Domena I. čini najveći HOR niz 34/36mer (pozicije 10316945-10544039 bp u centromeri). Identificirali smo polimorfne varijante koje imaju duplikacije monomera, delecije ili insercije, ili preraspodjele te insercije koje ne pripadaju repeticijskim elementima HOR-a do duljine od 44 monomera (pojedinosti su dane u (Vlahović, Glunčić, & Paar, (predano za objavu))). Ove polimorfne varijacije također su vidljive iz Needleman-Wunsch analize alfa satelitskih nizova u ove tri domene (dot plot dijagram na slici 4. za domenu II).



Slika 4. Needleman-Wunsch analiza divergencije monomera. Dijagonale predstavljaju monomere s divergencijom manjom od 7%. Horizontalne udaljenosti otkrivanju nmer HOR kopije koje su u ovom slučaju 34 mer HOR te njegove varijante.

cytogenetics for diagnostics purposes (Glunčić, Vlahović, & Paar, 2019).

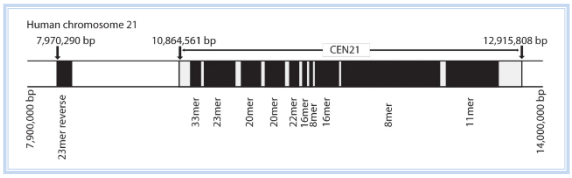


Figure 5. Positions of alpha satellites and nmer HORs in centromere of human chromosome 21 obtained with ALPHAsub algorithm (Glunčić, Vlahović, & Paar, 2019). In centromere region, there are arrays of 33mer, 23mer, 20mer, 20mer, 22mer, 16mer, 8mer, 16mer, 8mer, 11mer HORs. Outside of centromere region in chromosome 21 there is one array of 23mer HOR in its reverse form. HOR arrays of alpha satellites outside of centromere region in human genome are not common.

Global Repeat Map algorithm is well suited for the study of other tandem repeats and HOR identification in genomes. One of our results is identification of novel HORs in human and chimpanzee and other primate's chromosome 1. Of particular interest is our discovery of HORs, which are embedded in Neuroblastoma Breakpoint Family Genes and in hornerin gene (Paar, Glunčić, Rosandić, Basar, & Vlahović, 2011). Recently Romero et al. (Romero & al, 2018) confirmed our previous discovery (Paar, Glunčić, Rosandić, Basar, & Vlahović, 2011) of 1410-bp quartic HOR repeat in human hornerin gene with phylogenetic analysis based on monomer unit of ~39 bp (Fig. 6.).

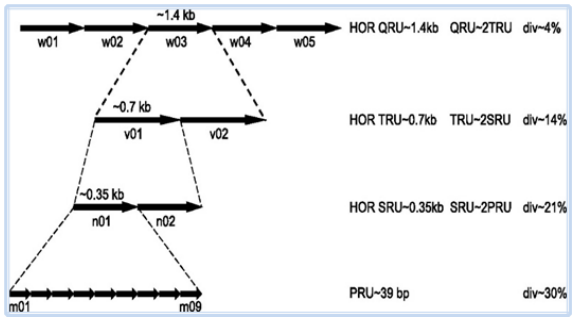
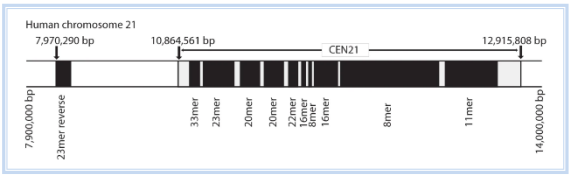


Figure 6. Scheme of 1410 bp HOR in human chromosome 1. Figure is taken from (Paar, Glunčić, Rosandić, Basar, & Vlahović, 2011).

Hierarchical structure of 1410-bp quartic HOR is composed of nine 39-bp primary repeat units (divergence ~30%) that form a ~0.35-kb secondary repeat unit. The two ~0.35-kb secondary repeat units (divergence ~21%) form a ~0.7-kb tertiary repeat unit, two ~0.7-kb tertiary repeat units (divergence ~14%) form ~1.4-kb quartic HOR repeat unit. HOR copies average divergence is ~4% (Paar, Glunčić,

Koristeći algoritme GRM i ALPHAsub, analizirali smo sve alfa satelitske nizove u ljudskom hg38, kao i u genomima čimpanze i neandertalca (podaci pohranjeni u bazi podataka http://genom.hazu.hr). Ovi podaci o alfa satelitskim nizovima i njihovim položajima unutar genoma daju trag o njihovoj distribuciji i evoluciji između blisko povezanih vrsta.

Izračun postotka sadržaja alfa satelita u genomima pokazuje da kod čimpanze (Pan_tro_3.0) oni zauzimaju 0,22%, kod neandertalca (Altai_Neandertal_ERP002097) 0,22%, a kod genoma čovjeka (hg38) 2,19% sekvence DNK (Vlahović & sur., 2022). U ljudskim kromosomima 17-22 postotak alfa satelita je veći, ~5%. U našoj prethodnoj publikaciji (Glunčić, Vlahović & Paar, 2019) analizirali smo ljudski kromosom 21 i otkrili u njemu najveću alfa satelitsku HOR jedinicu među svim somatskim kromosomima - 33mer HOR, te da je taj kromosom također prepun brojnih HOR jedinica sastavljenih od osam i više monomerskih tipova (Slika 5). Ovi HOR-ovi temeljeni na alfa satelitskim nizovima mogu se koristiti kao jedinstveni identifikator kromosoma 21 u molekularnoj citogenetici za potrebe dijagnostike (Glunčić, Vlahović & Paar, 2019).



Slika 5. Položaji alfa satelita u ljudskom kromosomu 21 dobiveni su algoritmom ALPHAsub (Glunčić, Vlahović & Paar, 2019). U području centromera postoje nizovi 33mer, 23mer, 20mer, 20mer, 22mer, 16mer, 8mer, 16mer, 8mer, 11mer HOR-ova. Izvan područja centromere u kromosomu 21 postoji jedan niz 23mer HOR u svom obrnutom obliku. HOR nizovi alfa satelita izvan područja centromera u ljudskom genomu nisu uobičajeni.

Algoritam Global Repeat Map prikladan je za proučavanje drugih tandemskih ponavljanja (osim alfa satelita) i identifikaciju HORova u genomima. Jedan od naših rezultata je identifikacija novih HORova u kromosomu 1 čovjeka i čimpanze te ostalih primata. Od posebnog je interesa naše otkriće HOR-ova koji su ugrađeni u gene obitelji neuroblastoma (Neuroblastom Breakpoint Family) i u genu hornerin (Paar, Glunčić, Rosandić, Basar & Vlahović, 2011). Nedavno su Romero i sur. (Romero & sur., 2018) potvrdili naše prethodno otkriće (Paar, Glunčić, Rosandić, Basar & Vlahović, 2011) kvartičnog HOR ponavljanja od 1410 bp u ljudskom hornerin genu filogenetskom analizom temeljenom na monomernoj jedinici od ~39 bp (Slika 6.)

Rosandić, Basar, & Vlahović, 2011). Our explanation of the hornerin gene HOR structure was accompanied by our initial discovery of the NBPF HORs (Paar, Glunčić, Rosandić, Basar, & Vlahović, 2011). We hypothesized (Paar, Glunčić, Rosandić, Basar, & Vlahović, 2011) that the NBPF HORs have important roles in human brain development and cognitive abilities. Later research by other scientists has shown the role of NBPF monomers (also called DUF1220/ Olduvai domain) are related to neurodegenerative diseases schizophrenia, autism, microcephaly, macrocephaly (O'Bleness & et al., 2014) and neuroblastoma (Andries & et al., 2015).

Our novel results for analysed sequences of several primates (*Pan Troglodytes, Gorilla gorilla, Pongo abelii, Papio anubis, Pan paniscus, Macaca fascicularis, Macaca mulatta*), as well as for the human and Neanderthal chromosome 1 show peaks for the ~1530 bp NBPF monomer units, ~3180 bp dimer units and for the ~4770 bp 3mer HOR NBPF units in our GRM diagrams. Fig.7 shows the most significant peaks are for human and Neanderthal chromosome 1 and that the corresponding weak signal is observed for Gorilla chromosome 1.

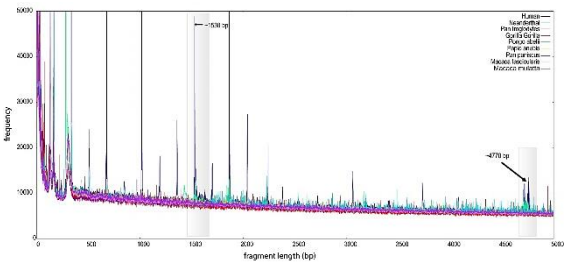
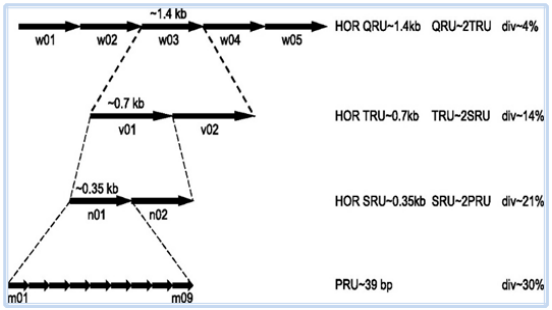


Figure 7. GRM diagrams of chromosome 1 from human, Neanderthal, Pan troglodytes, Gorilla gorilla, Pongo abelii, Papio anubis, Pan Paniscus, Macaca fascicularis and Macaca mulatta with pronounced peaks at ~1530 bp (monomer unit) and ~4770 bp (NBPF HOR).

Additionally, using BLAST algorithm we investigated the occurrence of human and Neanderthal monomer units in three GRM length ranges, <2500 bp, 2500 - 4000 bp and >4000 bp, to identify NBPF monomers, dimers and trimers that we identify by GRM analysis in different primates (Table 1). This novel results in Table 1 shows that most copies of those triplets are found in Gorilla genome that represent almost full triplet HOR copies. From Table 1 we can see that for genome assemblies used for this analysis, Gorilla gorilla genome have seven monomer copies with identity of ~82% when compared to human and Neanderthal NBPF consensus monomer sequences. There are 5 dimers when compared to human NBPF dimers but only 2 when compared to Neanderthal NBPF dimer sequence with identity of ~81% for human NBPF dimer and 80% for Neanderthal dimer. Interestingly



Slika 6. Shema 1410 bp HOR u ljudskom kromosomu 1. Slika je preuzeta iz (Paar, Glunčić, Rosandić, Basar & Vlahović, 2011).

Hijerarhijska struktura kvartičnog HOR-a od 1410 bp sastoji se od devet primarnih repeticijskih jedinica od 39 bp (divergencija ~30%) koje tvore sekundarnu repeticijsku jedinicu od ~0,35 kb. Dvije sekundarne repeticijske jedinice od ~0,35 kb (divergencija ~21%) tvore tercijarnu repeticijsku jedinicu od ~0,7 kb, dvije tercijarne repeticije od ~0,7 kb (divergencija ~14%) tvore kvartičnu HOR jedinicu od ~1,4 kb. Prosječna divergencija HOR kopija je ~4% (Paar, Glunčić, Rosandić, Basar & Vlahović, 2011). Naše objašnjenje strukture HOR gena hornerina popraćeno je i našim prvim otkrićem NBPF HORova (Paar, Glunčić, Rosandić, Basar & Vlahović, 2011). Postavili smo hipotezu (Paar, Glunčić, Rosandić, Basar & Vlahović, 2011) da NBPF HORovi imaju važnu ulogu u razvoju ljudskog mozga i kognitivnih sposobnosti. Kasnija istraživanja drugih znanstvenika pokazala su da uloga monomera NBPF gena (također nazvanih DUF1220/ Olduvai domene) su povezana s neurodegenerativnim bolestima kao što su shizofrenija, autizam, mikrocefalija, makrocefalija (O'Bleness, & sur., 2014) te neuroblastomom (Andries & sur., 2015).

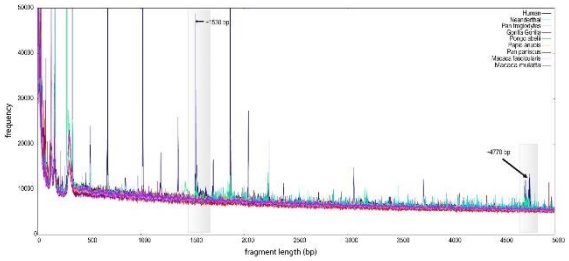
Naši novi rezultati za analizirane sekvence nekoliko primata *(Pan Troglodytes, Gorilla gorilla, Pongo abelii, Papio anubis, Pan paniscus, Macaca fascicularis, Macaca mulatta*), kao i za ljudski i neandertalski kromosome 1 pokazuju pikove za ~1530 bp NBPF monomernu jedinicu, ~3180 bp dimer jedinicu i za ~ 4770 bp 3mer HOR NBPF jedinicu na GRM dijagramu. Slika 7 prikazuje najznačajnije vrijednosti pikova za kromosom 1 čovjeka i neandertalaca, te odgovarajući slabi signal u kromosomu 1 gorile.

we identify nine NBPF trimers with ~81% identity for human NBPF 3mer HOR copy and nine trimers for Neanderthal NBPF 3mer HOR copy in *Gorilla gorilla* chromosome 1. Interestingly *Pan Troglodytes* (chimpanzee) have much less identified monomer, dimer and trimer units when compared to human and Neanderthal consensus monomers from NBPF 3mer HORs. In addition, we can see from Table 1 that some primates have greater number of monomers, dimers and trimers when compared to Neanderthal NBPF consensus sequences than to human NBPF monomer, dimer and trimer units. Examples are *Pongo abelii* with six trimer units when compared to Neanderthal vs five trimers compared to human, *Pan Troglodytes* have five dimers compared to Neanderthal vs 4 dimers when compared to human NBPF consensus sequences. *Macaca Fascularis* have two monomers when compared to Neanderthal vs one monomer when compared to human NBPF monomer consensus sequence. We have to emphasize that this big difference can be due to the percentage of completeness of primate genomes assemblies so this result is not finite – new assemblies with more complete genomes will give more detailed picture of NBPF monomer like copy numbers and HOR structures in primates.

Table 1. BLAST hit results for comparison of human and Neanderthal HOR consensuses with Pan Troglodytes, Gorilla Gorilla, Pongo abelii, Papio Anubis, Macaca fascicularis, Macaca mulatta and Pan Paniscus.

Species	<2500 bp			
	Human		Neanderthal	
	No.of copies	% identity	No.of copies	%identity
Pan troglodytes	3	81.57	2	81.26
Pan paniscus	1	81.56	1	79.82
Gorilla gorilla	7	82.29	7	81.14
Pongo abelii	4	81.40	3	80.80
Papio Anubis	2	79.29	2	78.03
Macaca Fascularis	1	79.56	2	79.56
Macaca mulatta	1	79.58	1	78.76
Species	**2500 - 4000 bp**			
	Human		Neanderthal	
	No.of copies	% identity	No.of copies	%identity
Pan troglodytes	4	80.44	5	78.37
Pan paniscus	1	78.93	1	78.60
Gorilla gorilla	5	81.81	2	80.10
Pongo abelii	3	79.84	1	78.92
Papio Anubis	3	80.30	2	79.23
Macaca Fascularis	1	80.85	1	80.79
Macaca mulatta	1	80.51	1	78.60
Species	**>4000 bp**			
	Human		Neanderthal	
	No.of copies	% identity	No.of copies	%identity
Pan troglodytes	3	80.29	3	81.77
Pan paniscus	2	80.25	2	81.22
Gorilla gorilla	9	81.24	9	81.76
Pongo abelii	5	81.54	6	82.07
Papio Anubis	1	77.44	1	77.81
Macaca Fascularis	2	79.21	2	79.59
Macaca mulatta	2	79.19	2	80.40

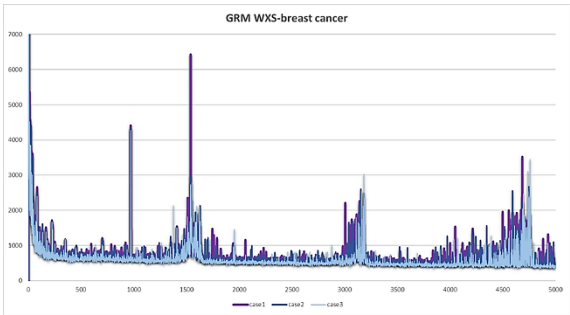
Another direction of our research is identification of structural variants in cancer genomes related with tandem repeats. Our preliminary novel study of the



Slika 7. GRM dijagrama za kromosom 1 čovjeka, neandertalca, Pan troglodytes, Gorilla gorilla, Pongo abelii, Papio anubis, Pan Paniscus, Macaca fascicularis i Macaca mulatta s naglašenim pikovima na ~1530 bp (monomerna jedinica) te ~4770 bp (NBPF HOR jedinica).

Dodatno, pomoću BLAST algoritma istražili smo pojavu ljudskih i neandertalskih monomernih jedinica za ove vrste u tri GRM raspona duljina, <2500 bp, 2500 - 4000 bp i >4000 bp, kako bismo identificirali NBPF monomere, dimere i trimere koje smo pronašli s GRM metodom za različite primate (Tablica 1). Ova tablica pokazuje naše nove rezultate da se većina kopija tih tripleta nalaze u genomu gorile koji predstavljaju gotovo pune triplet NBPF HOR kopije. Iz Tablice 1 možemo vidjeti da za sklopove genoma sekvenci korištene za ovu analizu, genom Gorilla gorilla ima sedam monomernih kopija s identičnošću od ~82% u usporedbi s ljudskim i neandertalskim NBPF konsenzusnim sekvencama monomera. U njemu postoji 5 dimera u usporedbi s ljudskim NBPF dimerima, ali samo 2 u usporedbi s neandertalskim NBPF dimerskom sekvencom s identičnošću od ~81% za ljudski NBPF dimer i 80% za neandertalski dimer. Zanimljivo je da smo identificirali devet NBPF trimera s ~81% identičnosti za ljudsku NBPF 3mer HOR kopiju i devet trimera za neandertalsku NBPF 3mer HOR kopiju u *Gorila gorila* kromosomu 1. Zanimljivo je da *Pan Troglodytes* (čimpanze) imaju puno manje identificiranih monomernih, dimernih i trimernih jedinica u usporedbi s ljudskim i neandertalskim konsenzusnim monomerima iz NBPF 3mer HORova. Osim toga, iz Tablice 1 možemo vidjeti da neki primati imaju veći broj monomera, dimera i trimera u usporedbi s neandertalskim NBPF konsenzusnim sekvencama u odnosu na ljudske NBPF monomerne, dimerne i trimerne jedinice. Primjeri su *Pongo abelii* sa šest trimernih jedinica u usporedbi s neandertalcem nasuprot pet trimera u usporedbi s čovjekom, *Pan Troglodytes* imaju pet dimera u usporedbi s neandertalcem nasuprot 4 dimera u usporedbi s ljudskim NBPF konsenzusnim sekvencama. *Macaca Fascularis* ima dva monomera u usporedbi s neandertalcem naspram jednog monomera u usporedbi s konsenzusnom sekvencom humanog NBPF monomera. Moramo naglasiti da ova velika razlika može biti posljedica postotka potpunosti sklopova genomske sekvence primata tako da ovaj rezultat nije konačan – novi sklopovi

whole exome sequence for three cases of breast cancer from CCLE project indicate that we could, on more extensive data sets, apply those methods for classification of DNA sequences in patients. This usage could be of interest in diagnostic and drug appliances (Fig.8). Fig. 8 shows three breast cancer cases of chromosome 1 and theirs associated GRM diagrams obtained for exome sequences. Novel results show that for even those three cases existence of structural variants i.e., copy number variants (CNVs) in coding part of chromosome 1 is identified. This conclusion comes from peak differences in frequency of individual fragment lengths (that represent monomer, dimer and trimer units) for those cancer cases. Examples of differences are seen, interestingly, in the area of fragment lengths of ~1530 bp, ~3180 bp and ~4770 bp that correspond to NBPF monomer, dimer and trimer units like for NBPF 3mer HOR GRM signals in Fig. 7 (although with lower frequency). The difference between GRM diagrams for human shown on Fig. 7 and GRM diagram for three exome cancer cases (Fig.8) is wider scattering around those fragment lengths that characterize NBPF peaks in cancer cases compared to hg38 human assembly (Fig.7). Precisely that scattering is an indicator of structural variants, partial deletions/insertions or copy number variants, so detailed analysis of that part of the sequence and construction of HOR variant schemes (like on Fig.2. for HOR copy construction) is needed as well as analysis of greater number of cases for obtaining exact structural variants and CNVs for classification purposes. It is not coincidence that NBPF signals for monomer, dimer and trimer units are most pronounced signals on GRM diagram. If we take in consideration that length of NBPF genes (in form of tandem repeats that we identify – NBPF20, NBPF10, NBPF12, NBPF14 and NBPF19 of 351410 bp) occupy 0.14% of chromosome 1 overall length (248956422 bp) while in coding part of chromosome 1 this percentage is 1.36% of overall 10.40% of coding part in whole chromosome 1. For other peaks to distinguish between those cancer cases we will have to use GRM algorithm ability of magnifying glass (zoom in) to smaller ranges of fragment lengths and frequency on diagram (Glunčić & Paar, 2013).



genomskih sekvenci s potpunijim genomima dat će detaljniju sliku NBPF monomera poput broja kopija i HOR struktura u primata.

Tablica 1. Rezultati BLAST pogodaka za usporedbu ljudskih i neandertalskih koncenzusnih HOR sekvenci s primatima Pan Troglodytes, Gorilla Gorilla, Pongo abelii, Papio Anubis, Macaca fascicularis, Macaca mulatta and Pan Paniscus.

Species	<2500 bp			
	Human		Neanderthal	
	No.of copies	% identity	No.of copies	%identity
Pan troglodytes	3	81.57	2	81.26
Pan paniscus	1	81.56	1	79.82
Gorilla gorilla	7	82.29	7	81.14
Pongo abelii	4	81.40	3	80.80
Papio Anubis	2	79.29	2	78.03
Macaca Fascularis	1	79.56	2	79.56
Macaca mulatta	1	79.58	1	78.76
Species	**2500 - 4000 bp**			
	Human		**Neanderthal**	
	No.of copies	% identity	No.of copies	%identity
Pan troglodytes	4	80.44	5	78.37
Pan paniscus	1	78.93	1	78.60
Gorilla gorilla	5	81.81	2	80.10
Pongo abelii	3	79.84	1	78.92
Papio Anubis	3	80.30	2	79.23
Macaca Fascularis	1	80.85	1	80.79
Macaca mulatta	1	80.51	1	78.60
Species	**>4000 bp**			
	Human		**Neanderthal**	
	No.of copies	% identity	No.of copies	%identity
Pan troglodytes	3	80.29	3	81.77
Pan paniscus	2	80.25	2	81.22
Gorilla gorilla	9	81.24	9	81.76
Pongo abelii	5	81.54	6	82.07
Papio Anubis	1	77.44	1	77.81
Macaca Fascularis	2	79.21	2	79.59
Macaca mulatta	2	79.19	2	80.40

Drugi smjer našeg istraživanja je identifikacija strukturnih varijanti u genomima raka povezanih s tandemskim ponavljanjima. Naša preliminarna nova istraživanja cijele sekvence egzoma za tri slučaja karcinoma dojke iz projekta CCLE ukazuje da bismo mogli, na opsežnijim skupovima podataka, primijeniti te metode za klasifikaciju sekvenci DNK kod pacijenata. Ova bi uporaba mogla biti od interesa za dijagnostiku i pametne lijekove (Slika 8). Slika 8 prikazuje tri slučaja raka dojke kromosoma 1 i njihove povezane GRM dijagrame dobivene za sekvence egzoma. Novi rezultati pokazuju da čak i za ta tri slučaja postoji indikacija za strukturne varijante, tj. varijante broja kopija (CNV) u kodirajućem dijelu kromosoma 1. Ovaj zaključak proizlazi iz razlika pikova u frekvenciji individualnih fragmentnih duljina (koje predstavljaju monomerne, dimerne i trimerne jedinice) za ove slučajeve karcinoma. Primjeri ovih razlika se vide, zanimljivo, u području fragmentnih duljina od ~1530 bp, ~3180 bp i ~4770 bp koje odgovaraju signalima NBPF monomernih, dimernih i trimernih jedinica sa Slike 7 (iako nižih frekvencija pošto se radi o kodirajućem dijelu kromosoma 1). Razlike između dijagrama prikazanog za čovjekov kromosom 1 i GRM dijagrama za 3 slučaja raka dojke i njihovog exoma (Slika 8) je šire raspršenje oko fragmentnih duljina

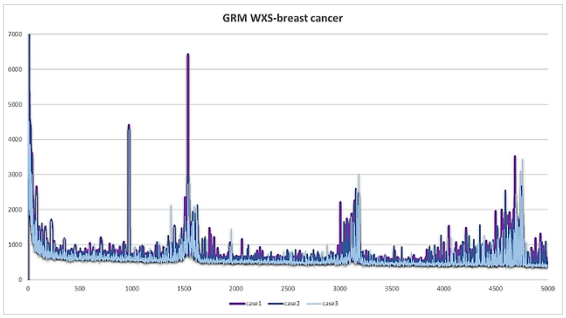
Figure 8. GRM diagram for three cases of cancer exome data for chromosome 1 downloaded from NCI GDC database, CCLE project. From these diagrams, we can conclude that structural variants, CNVs and partial HOR copies exists based on scattering around peaks in GRM diagram for NBPF 3mer HOR, corresponding to peak size in frequency (y-axis) at specific fragment lengths (x-axis).

4 Conclusion

A combination of novel methods for more precise genome assemblies involving combination of short-read, long-read and optical mapping (Weissensteiner, Pang, AWC, Bunikis, & et al., 2017) can cover large-scale tandem repeat arrays to avoid assembly errors [Tørresen, et al., 2019, Levy-Sakin & et al., 2019] especially in complex chromosome regions such as centromere. That could help in diagnostics as well as in search of new drugs and therapies, in combination with growing field of AI and machine learning methods for pattern recognition in bioinformatics.

Our GRM tool gives an advantage and opportunity to identify HORs in genomes of various species. With new sequencing technologies, our bioinformatics algorithms can provide even greater insight into tandem and HOR structure variants and CNVs in more precisely assembled genomes. Such a case is seen in hornerin gene HOR in human chromosome 1 for which we proposed a structure in ref. (Paar, Glunčić, Rosandić, Basar, & Vlahović, 2011), and later Romero found similar structure in other primates, using novel genome assembly (Romero & et al., 2018). Analysing human hg37 and hg38 genomes and chimpanzee genomes, we obtained novel results for HOR variants in human Y chromosome (Vlahović, Glunčić, & Paar, (submitted for publication)) with respect to earlier results for previous incomplete genomic assemblies (Paar & et al., 2011). Of particular interest for evolution of cognitive abilities and of neurodegenerative diseases are our discoveries of NBPF 3mer HORs in human and Neanderthal genomes and their absence/low occurrence in other primates [Paar, Glunčić, Rosandić, Basar, & Vlahović, 2011, Vlahović, Glunčić, Rosandić, & Paar, submitted for publication] probably due to incomplete primate genomes used for research. Example of assembly difference is for human hg37 that have in its chromosome 1 assembly 165 monomers of NBPF monomer unit length, organized in 57 HOR copies, while in hg38 there are 177 monomer units organized in 65 HOR copies. In hg38 chromosome Y sequence there are 51 HOR copies based on alpha satellite monomer unit with 1616 monomers, while in hg37 there are 11 HOR copies with 329 monomers, but bigger difference with newer assemblies of sequenced genome is changed HOR scheme for this chromosome. In hg37 it was 45mer HOR form of the koje karakteriziraju NBPF pikove kod slučaja karcinoma u usporedbi s hg38 ljudskim sklopom DNK sekvence kromosoma 1 (Slika 7).

Upravo je to raspršenje indikator za postojanje strukturalnih varijanti, djelomičnih delecija/insercija ili varijanti broja kopija, stoga je detaljna analiza tog dijela sekvence potreba i konstrukcija shema HOR varijanti (kao što je prikazano na Slici 2 za konstrukciju HOR kopija) te je potrebna analiza većeg broja slučajeva za dobivanje egzaktnih strukturalnih varijanti i CNV-ova za potrebe klasifikacije. Nije slučajno da su NBPF signali za monomerne, dimerne i trimerne jedinice najizraženiji signali na GRM dijagramu. Ako uzmemo u obzir da duljina NBPF gena (u obliku tandemskih ponavljanja koje smo identificirali – NBPF20, NBPF10, NBPF12, NBPF14 i NBPF19 od 351410 bp) zauzima 0.14% ukupne duljine kromosoma 1 (248956422 bp) te u kodirajućem dijelu kromosoma 1 ovaj postotak iznosi 1.36% od ukupnih 10.40% kodirajućeg dijela cijelog kromosoma 1. Za ostale pikove na GRM dijagramu, da bismo ih razlikovali u ovim slučajevima karcinoma moramo upotrijebiti sposobnost GRM algoritma - povećalo (zoom in) za uvećanje dijagrama na manje raspone fragmentnih duljina i frekvencije na dijagramu (Glunčić & Paar, 2013).



Slika 8. GRM dijagram za tri slučaja egzoma karcinoma za kromosom 1 preuzet iz NCI GDC baze podataka, s CCLE projekta. Iz ovih dijagrama možemo zaključiti da strukturne varijante, CNV i djelomične HOR kopije postoje na temelju raspršenja oko pikova za NBPF 3mer HOR, što odgovara veličini pika frekvencije (y-os) na specifičnim fragmentnim duljinama (x-os).

4 Zaključak

Kombinacija novih metoda za preciznije slaganje genoma koje uključuju kombinacije kratkog čitanja, dugog čitanja te optičkog mapiranja (Weissensteiner, Pang, AWC, Bunikis & sur., 2017) mogu pokriti nizove tandemnih repeticija velikih razmjera kako bi se izbjegle pogreške u slaganju genoma [Tørresen, & sur., 2019, Levy-Sakin, & sur., 2019], posebno u područjima kromosoma s velikim nizovima repeticija kao što je centromera. To bi moglo pomoći

scheme for Y chromosome, while within hg38 the HOR scheme is in 34/36mer HOR form.

Our future work in this field will be oriented to analysis of complex tandem repeats and HORs in eukaryotes, both plants and mammals, as well as in study of human genomes related to cancer (breast cancer, colorectal cancer) and neurodegenerative diseases (autism spectrum disorder) for which genomic data are available. In addition, we will use our tools to investigate population genomics/genetics of closely relates species such as humans and primates and theirs corresponding structural variants, CNVs in HOR structures. For those studies, we will make exact variant schemes in HOR structures that could be used for comparison and classification of available large data sets collected through different sequencing projects with machine learning and AI algorithms.

Acknowledgments

This work was supported by IP-2019-04-2757 from Croatian Science Foundation.

References

A.K., S., Mourad, M., Kaplan, M., & al., e. (2019). The Genomic Landscape of Centromeres in Cancers. *9*, str. 11259. doi: https://doi.org/10.1038/s41598-019-47757-6

Andries, V., Vandepoele, K., Staes, K., Berx, G., Bogaert, P., Isterdael, G. V., . . . Roy, F. v. (2015). NBPF1, a tumor suppressor candidate in neuroblastoma, exerts growth inhibitory effects by inducing a G1 cell cycle arrest. *BMC Cancer, 10*. doi:doi: 10.1186/s12885-015-1408-5

Black, E., & Giunta, S. (2018). Repetitive Fragile Sites: Centromere Satellite DNA As a Source of Genome Instability in Human Diseases. *Genes, 9*, str. 615. doi:doi:10.3390/genes9120615

Flèche, P. (2001). tandem repeats database for bacterial genomes: application to the genotyping of Yersinia pestis and Bacillus anthracis. *MC Microbiol, 1*. doi:doi:10.1186/1471-2180-1-2.

Fondon, J., & Garner, H. (2004). Molecular origins of rapid and continuous morphological evolution. *PNAS, 101*, str. 18058–63.

Glunčić, M., & Paar, V. (2013). Direct mapping of symbolic DNA sequence into frequency domain in global repeat map algorithm. *Nucleic Acid Research, 41*, str. e17. doi:doi:10.1093/nar/gks721

Glunčić, M., Vlahović, I., & Paar, V. (2019). Discovery of 33mer in chromosome 21 – the largest alpha satellite higher order repeat unit

u dijagnostici, kao i u potrazi za novim lijekovima i terapijama, u kombinaciji s rastućim područjem umjetne inteligencije i metoda strojnog učenja za prepoznavanje uzoraka u području bioinformatike.

Naš GRM algoritam daje prednost i priliku za identifikaciju HORova u genomima različitih vrsta. S novim tehnologijama sekvenciranja, naši bioinformatički algoritmi mogu pružiti još bolji uvid u tandemne varijacije i HOR strukture u preciznije sastavljenim genomima. Takav se slučaj vidi u HORu hornerin gena u ljudskom kromosomu 1 za koji smo predložili strukturu u ref. (Paar, Glunčić, Rosandić, Basar & Vlahović, 2011), a kasnije su i Romero & sur. pronašli sličnu strukturu u drugih primata, koristeći nove sekvence genoma (Romero et al, 2018). Analizirajući ljudske hg37 i hg38 sklopove genomskih sekvenci i genom čimpanze, dobili smo nove rezultate za HOR varijante u ljudskom Y kromosomu (Vlahović, Glunčić i Paar, predano za objavu) u odnosu na ranije rezultate za prethodne nepotpune sklopove genomskih sekvenci (Paar & sur., 2011). Od posebnog interesa za evoluciju kognitivnih sposobnosti i neurodegenerativnih bolesti su naša otkrića NBPF 3mer HORa u genomima čovjeka i neandertalca te njihova odsutnost/mala pojavnost kod drugih primata [Paar, Glunčić, Rosandić, Basar & Vlahović, 2011, Vlahović, Glunčić, Rosandić & Paar, predano na objavu] vjerojatno zbog nepotpunih genoma primata korištenih za istraživanje. Primjer razlike sklopa genomske sekvence je za ljudski hg37 koji u svom sklopu DNK sekvence kromosoma 1 ima 165 monomera duljine NBPF monomerne jedinice, organiziranih u 57 HOR kopija, dok u hg38 postoji 177 monomernih jedinica organiziranih u 65 HOR kopija. U sekvenci kromosoma Y iz hg38 postoji 51 HOR kopija bazirana na alfa satelitskoj monomernoj jedinici sa 1616 monomera, dok u hg37 postoji 11 HOR kopija sa 329 monomera, ali je veća razlika u odnosu na novije sklopove sekvenciranog genoma promijenjena HOR shema za ovaj kromosom. U hg37 to je bio 45mer-ni HOR oblik sheme za Y kromosom, dok je unutar hg38 HOR shema u 34/36mer HOR obliku.

Naš budući rad u ovom području bit će usmjeren na analizu složenih tandemskih ponavljanja i HORova u eukariota, kako biljaka tako i sisavaca, kao i na proučavanje ljudskih genoma povezanih s karcinomima (karcinom dojke, kolorektalni karcinom) i neurodegenerativnim bolestima (poremećaj iz autističnog spektra) za koje su dostupni genomski podaci. Osim toga, koristit ćemo naše alate za istraživanje populacijske genomike/genetike blisko srodnih vrsta kao što su ljudi i primati i njihovih odgovarajućih strukturnih varijanti, CNVova u HOR strukturama. Za te ćemo studije izraditi točne sheme varijanti HOR struktura koje bi se mogle koristiti za usporedbu i klasifikaciju dostupnih velikih skupova podataka prikupljenih

among all human somatic chromosomes. *Sci Rep*. doi:doi:10.1038/s41598-019-49022-2

Jain, M., Olsen, H., Turner, D., Stoddart, D., Bulazel, K., Paten, B., & et al. (2018). Linear assembly of a human centromere on the Y chromosome. *Nat Biotechnol., 4*, str. 321-3. doi:doi: 10.1038/nbt.4109

Levy-Sakin, M., & al., e. (2019). Genome maps across 26 human populations reveal population-specific patterns of structural variation. *Nature Communications, 10*, str. 1025. doi:https://doi.org/10.1038/s41467-019-08992-7

Levy-Sakin, M., Pastor, S., Mostovoy, Y., & al, e. (2019). Genome maps across 26 human populations reveal population-specific patterns of structural variation. *Nat Commun, 10*, str. 1025. doi:https://doi.org/10.1038/s41467-019-08992-7

Okonechnikov K, Golosova O, Fursov M & the GENE team. Unipro UGENE: a unified bioinformatics toolkit . Bioinformatics 2012 28: 1166-1167. doi:10.1093/bioinformatics/bts091

O'Bleness, M., Searles, V. B., Dickens, C. M., Astling, D., Albracht, D., Mak, A. C., . . . Sikela, J. M. (2014). Finished sequence and assembly of the DUF1220-rich 1q21 region using a haploid human genome. *BMC Genomics, 387*. doi:https://doi.org/10.1186/1471-2164-15-387

Paar, V., Glunčić, M., Rosandić, M., Basar, I., & Vlahović, I. (2011). Intragene Higher Order Repeats in Neuroblastoma BreakPoint Family Genes Distinguish Humans from Chimpanzees. *Molecular Biology and Evolution, 28*, str. 1877–1892. doi:https://doi.org/10.1093/molbev/msr009

Paar, V., x Glunčić, V., Basar, I., Rosandic, M., Paar, P., & Cvitković, M. (2011). Large Tandem, Higher Order Repeats and Regularly Dispersed Repeat Units Contribute Substantially to Divergence Between Human and Chimpanzee Y Chromosomes. *Journal of Molecular Evolution, 1*, str. 34-55.

Pathak, D., & Ali, S. (2012). Repetitive DNA: A Tool to Explore Animal Genomes/Transcriptomes, Functional Genomics. (D. G. Meroni, Ur.) *InTech*. doi:DOI: 10.5772/48259

Perry, G., Tito, R., & Verrelli, B. (2007). The evolutionary history of human and chimpanzee Y-chromosome gene loss. *Mol Biol Evol., 3*, str. 853-9. doi:doi: 10.1093/molbev/msm002

Rigden, D. J., & Fernández, X. M. (2022). The 2022 Nucleic Acids Research database issue and the online molecular biology database collection. Nucleic acids research, 50(D1), D1–D10. https://doi.org/10.1093/nar/gkab1195.

Romero, V., & al, e. (2018). High Order Formation and Evolution of Hornerin in Primates. *Genome Biology and Evolution, 10*, str. 3167–3175.

kroz različite projekte sekvenciranja s algoritmima strojnog učenja i umjetne inteligencije.

Zahvale

Ovaj rad je izrađen pod projektom IP-2019-04-2757 od Hrvatske zaklade za znanost.

Reference

A.K., S., Mourad, M., Kaplan, M., & al., e. (2019). The Genomic Landscape of Centromeres in Cancers. *9*, str. 11259. doi: https://doi.org/10.1038/s41598-019-47757-6

Andries, V., Vandepoele, K., Staes, K., Berx, G., Bogaert, P., Isterdael, G. V., . . . Roy, F. v. (2015). NBPF1, a tumor suppressor candidate in neuroblastoma, exerts growth inhibitory effects by inducing a G1 cell cycle arrest. *BMC Cancer, 10*. doi:doi: 10.1186/s12885-015-1408-5

Black, E., & Giunta, S. (2018). Repetitive Fragile Sites: Centromere Satellite DNA As a Source of Genome Instability in Human Diseases. *Genes, 9*, str. 615. doi:doi:10.3390/genes9120615

Flèche, P. (2001). tandem repeats database for bacterial genomes: application to the genotyping of Yersinia pestis and Bacillus anthracis. *MC Microbiol, 1*. doi:doi:10.1186/1471-2180-1-2.

Fondon, J., & Garner, H. (2004). Molecular origins of rapid and continuous morphological evolution. *PNAS, 101*, str. 18058–63.

Glunčić, M., & Paar, V. (2013). Direct mapping of symbolic DNA sequence into frequency domain in global repeat map algorithm. *Nucleic Acid Research, 41*, str. e17. doi:doi:10.1093/nar/gks721

Glunčić, M., Vlahović, I., & Paar, V. (2019). Discovery of 33mer in chromosome 21 – the largest alpha satellite higher order repeat unit among all human somatic chromosomes. *Sci Rep*. doi:doi:10.1038/s41598-019-49022-2

Jain, M., Olsen, H., Turner, D., Stoddart, D., Bulazel, K., Paten, B., & et al. (2018). Linear assembly of a human centromere on the Y chromosome. *Nat Biotechnol., 4*, str. 321-3. doi:doi: 10.1038/nbt.4109

Levy-Sakin, M., & al., e. (2019). Genome maps across 26 human populations reveal population-specific patterns of structural variation. *Nature Communications, 10*, str. 1025. doi:https://doi.org/10.1038/s41467-019-08992-7

Levy-Sakin, M., Pastor, S., Mostovoy, Y., & al, e. (2019). Genome maps across 26 human

Romero, V., Hosomichi, K., Nakaoka, H., Shibata, H., & Inoue, I. (2017). Structure and evolution of the filaggrin gene repeated region in primates. *BMC Evol Biol, 17*. doi:doi:10.1186/s12862-016-0851-5

Rozen, S., Skaletsky, H., Marszalek, J., Minx, P., Cordum, H., Waterston, R., & al, e. (2003). Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature, 6942*, str. 873-6. doi:doi: 10.1038/nature01723

Tørresen, O. K., Star, B., Mier, P., Andrade-Navarro, M. A., Bateman, A., Jarnot, P., . . . Linke, D. (2019). Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. *Nucleic Acids Research, 47*(21), str. 10994–11006 . doi:doi: 10.1093/nar/gkz841.

Tyler-Smith, C., & Brown, W. (1987). Structure of the major block of alphoid satellite DNA on the human Y chromosome. *J Mol Biol, 3*, str. 457-70.

Tyler-Smith, C., Oakey, R., Larin, Z., Fisher, R., Crocker, M., Affara, N., & al, e. (1993). Localization of DNA sequences required for human centromere function through an analysis of rearranged Y chromosomes. *Nat Genet., 4*, str. 368-75. doi:doi: 10.1038/ng1293-368

Uralsky, L., Shepelev, V., Alexandrov, A., Yurov, Y., Rogaev, E., & Alexandrov, I. (2019). Classification and monomer-by-monomer annotation dataset of suprachromosomal family 1 alpha satellite higher-order repeats in hg38 human genome assembly. *Data Brief., 24*, str. 103708. doi:doi: 10.1016/j.dib.2019.103708

Vlahovic, I., Gluncic, M., Rosandic, M., Ugarkovic, Đ., & Paar, V. (2017). Regular Higher Order Repeat Structures in Beetle Tribolium castaneum Genome. *Genome Biol Evol, 9*, str. 2668–2680. doi:doi:10.1093/gbe/evw174

Vlahović, I., Glunčić, M., & Paar, V. (n.d.). Rich polymorphic variants of alpha satellite 34mer higher order repeats in hg38 assembly of human chromosome Y (submitted for publication).

Vlahović, I., Glunčić, M., Dekanić, K., Mršić, L., Jerković, H., Martinjak, I., & Paar, V. (2022). Global repeat map algorithm (GRM) reveals differences in alpha satellite number of tandem and higher order repeats (HORs) in human, Neanderthal and chimpanzee genomes – novel tandem repeat database. *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, str. 237-242. doi:doi: 10.23919/MIPRO48935.2020.9245278.

Vlahović, I., Glunčić, M., Rosandić, M., & Paar, V. (n.d.). Higher Order Repeats (HORs) in Neuroblastoma BreakPoint Family (NBPF) genes distinguish Neanderthal and Human populations reveal population-specific patterns of structural variation. *Nat Commun, 10*, str. 1025. doi:https://doi.org/10.1038/s41467-019-08992-7

Okonechnikov K, Golosova O, Fursov M & the GENE team. Unipro UGENE: a unified bioinformatics toolkit . Bioinformatics 2012 28: 1166-1167. doi:10.1093/bioinformatics/bts091

O'Bleness, M., Searles, V. B., Dickens, C. M., Astling, D., Albracht, D., Mak, A. C., . . . Sikela, J. M. (2014). Finished sequence and assembly of the DUF1220-rich 1q21 region using a haploid human genome. *BMC Genomics, 387*. doi:https://doi.org/10.1186/1471-2164-15-387

Paar, V., Glunčić, M., Rosandić, M., Basar, I., & Vlahović, I. (2011). Intragene Higher Order Repeats in Neuroblastoma BreakPoint Family Genes Distinguish Humans from Chimpanzees. *Molecular Biology and Evolution, 28*, str. 1877–1892. doi:https://doi.org/10.1093/molbev/msr009

Paar, V., x Glunčić, V., Basar, I., Rosandic, M., Paar, P., & Cvitković, M. (2011). Large Tandem, Higher Order Repeats and Regularly Dispersed Repeat Units Contribute Substantially to Divergence Between Human and Chimpanzee Y Chromosomes. *Journal of Molecular Evolution, 1*, str. 34-55.

Pathak, D., & Ali, S. (2012). Repetitive DNA: A Tool to Explore Animal Genomes/Transcriptomes, Functional Genomics. (D. G. Meroni, Ur.) *InTech*. doi:DOI: 10.5772/48259

Perry, G., Tito, R., & Verrelli, B. (2007). The evolutionary history of human and chimpanzee Y-chromosome gene loss. *Mol Biol Evol., 3*, str. 853-9. doi:doi: 10.1093/molbev/msm002

Rigden, D. J., & Fernández, X. M. (2022). The 2022 Nucleic Acids Research database issue and the online molecular biology database collection. Nucleic acids research, 50(D1), D1–D10. https://doi.org/10.1093/nar/gkab1195.

Romero, V., & al, e. (2018). High Order Formation and Evolution of Hornerin in Primates. *Genome Biology and Evolution, 10*, str. 3167–3175.

Romero, V., Hosomichi, K., Nakaoka, H., Shibata, H., & Inoue, I. (2017). Structure and evolution of the filaggrin gene repeated region in primates. *BMC Evol Biol, 17*. doi:doi:10.1186/s12862-016-0851-5

Rozen, S., Skaletsky, H., Marszalek, J., Minx, P., Cordum, H., Waterston, R., & al, e. (2003). Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. *Nature, 6942*, str. 873-6. doi:doi: 10.1038/nature01723

Tørresen, O. K., Star, B., Mier, P., Andrade-Navarro, M. A., Bateman, A., Jarnot, P., . . . Linke, D. (2019). Tandem repeats lead to

lineage with respect to cognitive capabilities (submitted for publication).

Weissensteiner, M., Pang, AWC, Bunikis, I., & al, e. (2017). Combination of short-read, long-read, and optical mapping assemblies reveals large-scale tandem repeat arrays with population genetic implications. *Genome Res., 5*, str. 697–708. doi:doi:10.1101/gr.215095.116.

sequence assembly errors and impose multi-level challenges for genome and protein databases. *Nucleic Acids Research, 47*(21), str. 10994–11006 . doi:doi: 10.1093/nar/gkz841.

Tyler-Smith, C., & Brown, W. (1987). Structure of the major block of alphoid satellite DNA on the human Y chromosome. *J Mol Biol, 3*, str. 457-70.

Tyler-Smith, C., Oakey, R., Larin, Z., Fisher, R., Crocker, M., Affara, N., & al, e. (1993). Localization of DNA sequences required for human centromere function through an analysis of rearranged Y chromosomes. *Nat Genet., 4*, str. 368-75. doi:doi: 10.1038/ng1293-368

Uralsky, L., Shepelev, V., Alexandrov, A., Yurov, Y., Rogaev, E., & Alexandrov, I. (2019). Classification and monomer-by-monomer annotation dataset of suprachromosomal family 1 alpha satellite higher-order repeats in hg38 human genome assembly. *Data Brief., 24*, str. 103708. doi:doi: 10.1016/j.dib.2019.103708

Vlahovic, I., Gluncic, M., Rosandic, M., Ugarkovic, Đ., & Paar, V. (2017). Regular Higher Order Repeat Structures in Beetle Tribolium castaneum Genome. *Genome Biol Evol, 9*, str. 2668–2680. doi:doi:10.1093/gbe/evw174

Vlahović, I., Glunčić, M., & Paar, M. (n.d.). Rich polymorphic variants of alpha satellite 34mer higher order repeats in hg38 assembly of human chromosome Y (poslano za objavu).

Vlahović, I., Glunčić, M., Dekanić, K., Mršić, L., Jerković, H., Martinjak, I., & Paar, V. (2022). Global repeat map algorithm (GRM) reveals differences in alpha satellite number of tandem and higher order repeats (HORs) in human, Neanderthal and chimpanzee genomes – novel tandem repeat database. *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, str. 237-242. doi:doi: 10.23919/MIPRO48935.2020.9245278.

Vlahović, I., Glunčić, M., Rosandić, M., & Paar, V. (n.d.). Higher Order Repeats (HORs) in Neuroblastoma BreakPoint Family (NBPF) genes distinguish Neanderthal and Human lineage with respect to cognitive capabilities (poslano za objavu).

Weissensteiner, M., Pang, AWC, Bunikis, I., & al, e. (2017). Combination of short-read, long-read, and optical mapping assemblies reveals large-scale tandem repeat arrays with population genetic implications. *Genome Res., 5*, str. 697–708. doi:doi:10.1101/gr.215095.116.

Digital Response to Pandemic

Vjeran Bušelić
Teaching Information Literacy and Critical Thinking Skills in Coronavirus Pandemic Time

Katarina Tomičić-Pupek, Ana Kutnjak, Vjeran Strahonja
Insights into understanding the perspective of digital platform providers

Nikolina Žajdela Hrustek, Diana Šimić, Žarko Čižmar
An instrument for measuring needs of vulnerable groups in terms of digital inclusion

Azra Staščik, Amanda Glavaš, Filip Glavaš
Regulation of the legal protection of individual in digital environment within the Croatian educational system

Ema Novak, Ana Kutnjak, Larisa Hrustek
Opportunities for a Digital Transformation of the Global Supply Chains: Case Study of the Suez Canal Blockade

Goran Bubaš, Antonela Čižmešija
Possible Changes in Academic Teaching as a Consequence of the COVID-19 Pandemic: International Findings and a Case Study from Croatia

Teaching Information Literacy and Critical Thinking Skills in Coronavirus Pandemic Time

Vjeran Bušelić
Zagreb University of Applied Sciences
Vrbik 8, 10000 Zagreb
vbuselic@tvz.hr

Abstract. *The aim of the research is to demonstrate the author's intention of designing a successful high school freshmen course that can improve students' generic skills and competences needed to empower them on a journey to become valuable contributors of modern society. Even in global COVID-19 pandemic circumstances, when all education went on-line. Whole six-year educational journey is documented and explained through two type of students' surveys, one official, institutional and the authors one, proving some good initial and pandemic Course design decisions. In discussion part some authors recommendations and further challenges are opened.*

Keywords. information literacy, critical thinking, teaching, Coronavirus pandemic, freshman course, generic skills

1 Introduction

Development of students' generic skills and competencies for better employability and personal prosperity along with primary (ICT) vocation is author's preoccupation since 2012, upon finishing business and starting an academic career.

The idea is permanent empowerment of students to develop or enhance 21st century (generic) skills and related competencies. Skills needed for their University learning environment, and consequently, their (ICT) future in the upcoming VUCA[1] industry 4.0 / digital economy world (WEF, 2015), (WEF, 2016), (Fadel et al., 2015).

With a general lack of conceptual clarity, questions about the classification of generic competencies as domain-specific or domain-general, and whether their development should be within the curriculum or co-curricular and extra-curricular activities. As there are so many of them, very few of them, the ones with great importance for specific vocation fields (like Communication or Creativity) have their own, full course within the academic syllabus.

2 Teaching Information Literacy and Critical Thinking Skills Class

In the academic year 2016/17 the author offered the full freshmen course "Information Literacy and Critical Thinking" originally developed for the students of Undergraduate study of Information Technology at Zagreb University of Applied Sciences[2].

It is important to underline that in most developed countries Information Literacy and Critical Thinking Skills learning outcomes are often incorporated into the learning outcomes of a whole study, not just a single class. For example, on a list of The University of Kansas first Core Goal – "Build core skills of critical thinking and quantitative literacy", there are as many as 204 approved courses[3], which would (partially) develop desired generic IL&CT skills, and not one single class with its name. Further on, looking for single classes experience, they are often elective, student orientation courses like the ones in San Diego State University (Harley, 2001), or as a block of two "thought and expression" courses "Advancing critical thinking and information literacy skills" for the first-year college students at Gonzaga University, Spokane (Alfino et al. 2008).

Zagreb University of Applied Sciences decision to dedicate the whole class to the development of students' generic IL&CT skills through teaching dedicated Class, turned into measurable success for at least the first two-generation (Bušelić, 2019).

2.1 Coronavirus pandemic online shift

The Class of 2019/20 started on March 2020, at the very beginning of the coronavirus pandemic online class delivery regime in Croatia. The author did not

[1] VUCA is a frequently used managerial acronym explaining attributes of today's world. It stands for Volatile, Uncertain, Complex and Ambiguous.

[2] All design principles of course syllabus, together with first two generations satisfaction and skills improvement surveys are fully elaborated in the 2019 MIPRO conference paper (Bušelić, 2019).

[3] Fulfilling the KU Core, The University of Kansas, https://kucore.ku.edu/ku-core-goals, accessed 7/22/22.

have any online experience teaching so far, so there were many open questions on institutionally allowed course redesign, in line with new circumstances. There was not an option to change learning outcomes, nor general teaching method – project-based learning, but there were a lot of unknowns on appropriate usage of online learning environment.

The basic literature study was from the field of business and HE distance education field (Grandzol & Grandzol, 2006), (Bernard et al., 2004); materials concerning online learning best practices (Keengwe & Kidd, 2010), (Palvia et al., 2018), (Wilson, 2018) and re-reading of some new CT&IL courses practice to better figure out what applies to online environment (Choy, & Cheah, 2009), Weiner, 2011), (Rezaee et al., 2012).

The main difference between graduate students of business education and undergraduate students lies primarily in **motivation** (they are aware of their education cost), and much more already developed skills, like time management, teamwork, ability to work/learn under the pressure, most skills undergraduate students need to develop yet.

On the Lecturer side two most important Lecturer challenges, successfully mastered in a classroom setting, are **the importance of interactivity** and **the ability to critically think himself**, thus able to influence the students.

In a very informative study, with the plethora of cited research on the role of Instructor (teacher) in online and face-to-face Critical Thinking classes "… results indicate that the mode of instructional delivery (face-to-face or online) is not as influential as the **instructor's level of interactivity** in promoting active engagement with course material" (Mandernach et al, 2009).

In an even older, the more cited Case study on teaching critical thinking skills using course content material, not through dedicated class, Thomas Lauer in introductory biology freshmen course "… taught the concept by not specifically identified or labeled critical-thinking terminology in a class. It was not the intent of my study to have students recognize or list critical-thinking components, but rather to improve cognitive abilities from a conceptual viewpoint" (Lauer, 2005). He specifically pinpointed that "… fortunately, we can use critical-thinking skills to teach critical thinking", claiming that if teachers have little knowledge about critical thinking skills; therefore, it is difficult for them to engage students in different activities to develop their high-order thinking skills.

The author of this Course, Lecturer himself, MSc. in Mathematical logic, had no problem with the usage of critical thinking skills to influence the students. They indirectly confirmed in a survey with comments like "*I really like the approach of the professor (through jokes, laughter, interaction …), he explains the topics very well, and along the way, a lot of people* ***wonder about some things****, I have no complaints!".*

Basic Lauer principle, **not to teach students about topic, but simply guide them to learn by themselves** was fully applied, and proved very successful. The author intentionally decided to use "nowdays very popular buzzwords" as Information Literacy and Critical Thinking, making a full course about them, but never explained, mentioned, or "preached/taught" **any other skill** they mastered. This research demonstrated they learned a lot of valuable (generic) skills as a "Class byproduct".

2.2 Online environment changes

From an institutional point, each lecturer was able to choose their own online delivery platform and slightly modify assessments. Standard Croatian tools for High Education online education are LMS Moodle and Microsoft Teams collaboration suite. The author immediately picked **MS Teams** with no LMS usage at all. As already pointed out in the literature review recommendations, the primary reason was **collaboration** and **interactivity**.

Another important design factor was **not to record any lessons**, for the same reason. This Course is about **active participation** and **teamwork**, so anyone not able to participate should find their way to update himself through teammates, not video lectures.

PowerPoint presentations, as **class notes, were available** immediately after the sessions, with all other working materials, and all the students were from the beginning divided into (random) teams working on individual project themes. During the Class **MS Teams Team Channels were used**, together with any additional collaboration communication tools students were used to (Slack, WhatsApp, …).

To maintain almost the same learning outcomes, a simple **knowledge test of understanding the theoretical background** of Information Literacy and Critical Thinking **was omitted**, so all the focus could be transferred to **skill practice**. By doing that course design changes, lessons become even more interactive, with **fever lectures** and much **more workshop time**.

To keep **relevance to the present situation,** most the examples were transformed into pandemic context. For instance, the problem of fake news and difficulties to figure out the truth was replaced with a problem to understand the constant **overflooding of COVID-19 related information** and recommendations on what to do, or not to do to stay well and healthy.

Online teamwork was maybe even more suitable than classroom practice because the class could switch in a second to teams' private channel mode, so they can immediately continue working on an **individual task inside each group**, and report back to Classroom mode with no interruption of others.

Keeping interaction and **motivation of a large group** of 50 – 60, up to almost 100 students, online was much more difficult than looking at them directly, instantly responding to any sign of nonunderstanding

or boredom, but manageable, requiring **much more side chatting, commenting, questioning, joking**, **…**

The last design delivery difference was 25-minute uninterrupted sessions with mandatory five-minute breaks, as the author explained and introduced the **Pomodoro learning technique**, teaching them as well some new methods of learning.

According to mandatory student class feedback, which is conducted at the beginning and after the end of every class, the author was satisfied with overall Class feedback and continued to deliver it for two more generations with really **minor changes**.

3 Research

This research aims to confirm the author's intention of designing a successful high school freshmen course that can improve students' generic skills and competences needed to empower them on a journey to become valuable contributors to modern society. Despite severe pandemic circumstances which affected the society, high education, students and teachers themselves[4].

3.1 Methodology

To understand student's attitudes towards the Class **two surveys were conducted**. The first one is an institutional, official survey, each student has to fill at the beginning and the end of each semester, and the other is created by the author to get additional, more insightful information about the Class and acquired and/or improved generic skills.

The second one, the **author's unannounced survey**, was performed in the period from 2nd to 7th of May 2022, for all three "pandemic" generations. As the Class finished more than half a year, or even two and a half years ago, and as the author did not have any contact nor he will conduct any class till the end of their study, students were not biased at all. They just recollected their thoughts about the things they have learned quite some time ago.

The survey was the same as described in (Bušelić, 2019), with additional four questions concerning pandemic situation. It reached 201 students out of 205 who completed the Course (four of them left the University), and got an exceptional return rate of 48,26 % answers (97 out of 201). Ten of them were invalid, and nine more were uncompleted, so there is a 43,28 % return rate (87 out of 201).

This sample size gives a confidence interval of 0,08, with std. error of 0,04, within a confidence level of 95%. Even year by year samples of 31,36 and 20 answers (37,80% for 2019/20, 55,38% for 2020/21 and 37,04% for 2021/22) within same confidence level of 95% are acceptable. Confidence interval for each year is 0,14; 011 and 0,18 with std. error of 0,07, 0,06 and 0,09. That validates conclusions performed for insight made for each pandemic generation as well.

Survey repeated the exactly the same set of questions as already published analysis of Classes of 2016/17 and 17/18, with methodology and results publicly discussed on 42 International Conference MIPRO and published in IEEE Proceedings, ensuring from the field experts' side that survey does not contain a common type of errors like leading, confusing or double-barreled questions. In its afterlife, according to the Google Scholar, the article is cited 7 times, and according to the authors ResearchGate page was downloaded and acknowledged more than 300 times, with no expert objections or comments on methodology.

The high level of reliability and internal consistency of the main 20 skills questionnaire presented in Table 1. (Improvement of skills and competences after the Class) is confirmed by Cronbach's Alpha coefficient of 0,962.

The other survey that is used in this research is an **official, the institutional survey** that is performed at the beginning and the end of each class, so all institutional classes and lecturers can be compared. It has two groups of questions: about topics of a course and lecturer performance. At the beginning of each class, 14 questions are asked - 10 about the class and four about the lecturer. At the end of each class, 27 questions are asked: 18 about a Class and nine about lecturer. Scale is classical school grading (1 – 5). For this research 13 selected questions from the end of class, survey are used to analyze students' satisfaction and comparison with other classes during their study – eight about the class and five about the lecturer[5].

Although both surveys are mandatory, students can fill in the blanks, so there is an usual low return rate. The return rate for class 2019-20 was 54.65 % (47 out of 86 students), for 2020-21 was 61.54 % (40 out of 65) and for 2019-20 only 31.48 % (17 out of 54 students).

4 Results

In this Chapter findings of the institutional "End of Class survey" and author's "Student's improvements survey" are given together with a comparation of "pre-pandemic" and "pandemic" generations results.

4.1 End of Class Institutional survey

The comparison of students' satisfaction with a Class vs. average satisfaction with a whole Study, prior to and during the pandemic, is given in Fig 1.

[4] This research should be interpreted as an extension to (Bušelić, 2019) research, showing that even in pandemic circumstances Information Literacy and Critical Thinking online class can bring substantial value to students.

[5] Questions selected for this analysis are processed in Fig. 2.

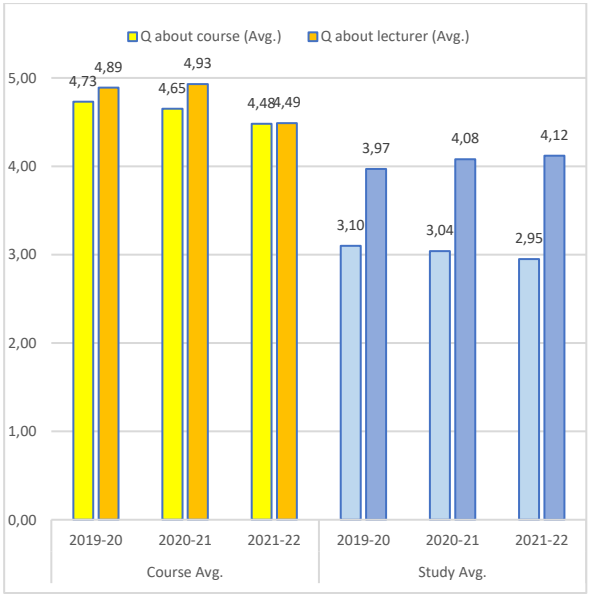


Figure 1. Official student's Avg. Course performance satisfaction survey with Study Avg comparison.

It is clear that the overall Course average is much higher than the study average (4,62 vs. 3,03), and Lecturer satisfaction is also significantly higher than the satisfaction of all lecturers (4,77 vs. 4,06), although lecturers are rated much better than study (4,06 vs. 3,03).

The average score for this analysis is calculated out of 13 selected questions from the institutional, mandatory survey. Eight are about topics and five are about a lecturer. When looking at Fig 2. - individual question Class averages, it is visible that all grades are in an upper range, distributed from 4,2 to 4,95 (4 questions in 2020-21), with the obvious drop in the third year (2021-22) - minimum scores for each question.

This visible drop in students' Class satisfaction in 2021-22, is a very good indicator of their exhaustion from the pandemic situation.

4.2 Student improvements survey

As official students' satisfaction study intention is not to go into the actual topics of each class, to understand their attitude towards this Class, especially during the pandemic time, the author conducted the same survey as in 2019, with more details, looking for their personal improvement of most important skills and competences acquired during the Course.

As in the class, the majority of students were male (81,52 %), with little more (50:42) students with general (gymnasium) than the vocational backgrounds. Majority of them (69,57 %) are 20 – 25-year-old. Majority (59,79) are very good graded students, 26,09% good and 14,13% excellent. Most of them (76.09 %) are regular students, but only 26,09% have no working experience. The rest of them are either temporary working or full-time employed.

The survey structure is simple: after asking them one simple open question **"Have you learned something useful during the course?"** with one-third (33,33%) answered No, the rest of them **elaborate on what it was**. There were 44 individual answers (78,57% response rate) and most of the answers will be presented during the discussion part, right after the structured part of the survey.

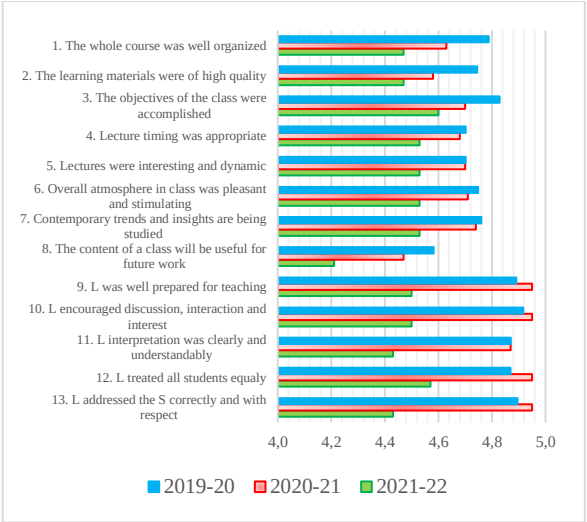


Figure 2 Selected 13 questions, comparison for a three-year period

4.2.1 Pandemic influence

The only difference between the 2019 survey was **four questions concerning pandemic influence.** The three questions were Likert 7-point agreement scale about their experience with their individual external and internal circumstances plus their relation/treatment of this Class:

ExtQ: *My external environment (everything that surrounded me during the pandemic – the general public, work, family, friends, working conditions, college, ...).*
The overall external corona pandemic situation (objectively or subjectively) I found myself in was:

IntQ: *My inner state (everything that happened to me, how I felt, behaved, acted, worked, ...).*
My state (how I personally coped) during the crown of the pandemic was:

The scale went from (1) Extremely negative (stressful), (2) Very negative (very disturbing), (3) Quite negative (often in a state of anxiety), (3) Neutral (as usual), (5) Extra motivating (sometimes), (6) Increased motivation (quite common), to (7) Extremely motivating (almost constantly, kind of positive stress).

ClassQ: *Situation related to this particular course*.
If you would compare your attitude/commitment/work in classical, live teaching and the situation we worked on (completely independent of the situation, looking back objectively), please evaluate YOUR personal relationship/ contribution. It was:

The scale for the Class experience was a little bit different, but also spread from a negative to a positive range: (1) Minor (I didn't really care), (2) Ignoring (to be done, if have spare time), (3) Reduced (fulfill obligation, get a grade), (4) Standard (attended, listened, worked as usual), (5) Increased (spend more time, read/worked more), (6) Excited (looking forward, engaged), and (7) Delighted (happily waited for, extra work/read).

At the end of this block, there was an **open comment section about education in a pandemic situation** or any personal remark, so they can elaborate if wanted. There were 41 (47,13%) very valuable observations, the author will take into consideration for Class improvement.

Most answers stand out online benefits like no travel time (students living out of Zagreb) and/or no problem with working schedule (3 out 4 are working, fully or temporarily), and on the negative side No.1 was lack of interaction with colleagues and professors, followed by lower quality of practical work during mandatory exercises. Most of them pointed out that according to the circumstances the study was well organized.

Upon looking at student's relationship with the Class in time, two things can be seen from the distribution of answers to ClassQ:

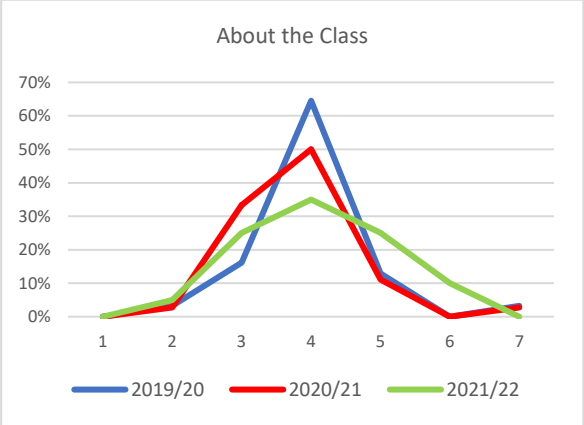


Figure 3. Class question, distribution by year

Distribution is quite normal, flattening the top as time goes by, with a slightly lower attitude in the second year and slightly better in the last one. As if they had adapted to the situation.

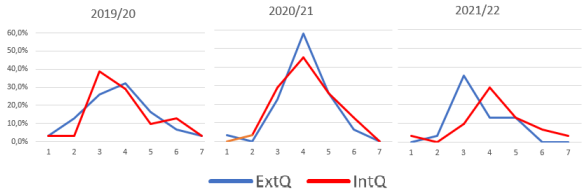


Figure 4. Yearly distribution of student's external and internal experience

Their relationship toward pandemia is even better visible though comparison between external (difficulties like pandemia, earthquake in Zagreb at the beginning of 2019/20 class) and internal state. Average of classes for internal state went positive from 3,90 to 4,14 and 4,25, but for external circumstances after 2020/21 consolidation went down again – from 3,77, 4,06 sharp drop to 3,55, showing that most of them considered the third year of pandemia even worse than initial pandemia plus earthquake.

4.2.2 Skills improvement and importance

Survey continues as previous one with simple open question "**Have you learned something useful during the course?**". And if any, **what it is**? This time exactly one-third of them (surprisingly) answered "No"[6], and out of 56 "Yes", 40 of them (71,42%) elaborate what it was in an open form, with their own words. Most of "Yes" elaboration will be commented on during the discussion part, right after the structured part of the survey.

After this free-form comments of what they have learned, they were just asked a question to **choose the area of their life they improved the most**: faculty, family, friends, work, personal development, all of the above or none. Most of them (36,90%) attributed it to the area of personal development.

In the structured part of the survey, they were offered a list of 20 skills and competences to grade them according to criteria "**Evaluate how much you have developed and improved your skills and competencies during the course:**"

Criteria was Likert 7-point agreement scale coded from (I) not at all, (II) very little, (III) little, (IV) average, (V) above average, (VI) very much, to (VII) extremely.

Table 1. Improvement of skills and competences after the Class

Skills and competences	I	II	III	IV	V	VI	VII
Taking the stance/attitude	3	3	8	14	23	18	9
Critical evaluation of stance (own incl.)	3	1	12	19	17	18	8
Responsibility	2	4	7	23	20	14	8
Result orientation	4	2	10	21	14	21	6
Adaptability	3	5	6	23	23	12	6
Argumentation	3	3	12	14	29	14	3
Information presentation (ppt & seminar)	2	5	14	18	19	13	7
Communication	2	8	6	25	16	15	6
Awareness of information need	2	2	10	29	19	12	4
Openness	4	4	11	19	21	13	6
Self-criticism	2	7	9	26	15	10	9
Teamwork	2	9	8	25	13	14	7
The influence of emotion on attitude	5	4	10	21	24	7	7
Logical component of argumentation	5	1	11	23	26	9	3
Idea and formulation	2	6	12	22	24	10	2
Persistence	5	3	14	24	17	12	3

[6] Author could not identify any correlation between answer "No" and their gender, age, education or working relation.

Learning by participation	4	9	11	25	15	8	6
Tools (Mind mapping)	6	8	11	22	20	7	4
Information evaluation (CRAPP method)	7	7	14	19	16	13	2
ERR learning method	4	13	10	27	11	8	5

The results are sorted out according to most developed or improved skills and showed progress in all 20 skills covered in this questionnaire. Even last three have shown average development / improvement (4,01; 3,99 and 3,92). Upon looking at overall extremes, it is exactly 2,06 times more (VI) very much and (VII) extremely answers than (I) not at all and (II) very little.

After this, a simple check box criterion is offered to them: "**Out of all the above-mentioned features, select a few (5-10) of the ones that you personally consider being the most valuable and which you believe are important for the next generation of students**". Before the list, a comment has to be made, that in average they pointed out around eight skills each (8,27), proving a point that they learned a lot of very valuable skills. A list of top 10 acquired skills, all mentioned from more than a half of students, is given in Table 2. The percentage is calculated as percentage of students nominated particular skill in an arbitrary group they believe are most valuable.

Table 2. Most valuable acquired skills according to student's recommendation

#	Top Skills	%
1	Responsibility	67,95%
2	Taking the stance/attitude	62,82%
3	Critical evaluation of stance (own incl.)	57,69%
4	Argumentation	57,69%
5	Teamwork	57,69%
6	Awareness of information need	51,28%
7	Communication	48,72%
8	Self-criticism	48,72%
9	The influence of emotion on attitude	39,74%
10	Persistence	39,74%

The number one skill on the list is Responsibility. This means that it was selected/recommended by 53 out of 78 students. So, most of them think it is important, and as it was 3rd on the previous list (of skills they learned or improved) it is pretty much clear they consider it important, but more important is a learning process per se. During a Class, there was no single lecture on responsibility, so obviously students themselves learned its importance by fulfilling set obligations. This is the same with the first three skills on this list, but the other two were much more mentioned during Class.

Comparing these two tables is interesting in the position of "Result orientation", skill No.4. of skills they mastered. Although obviously their very important learning, they do not consider recommending it to others (No. 16 with just 30,77 %, or 24 recommendations on this list). The opposite situation is with "Teamwork", a skill they would definitely recommend (57,69 %, No. 3 on a list). It is not high on the list of skills they mastered, only on 12th position. Probably because of online environment.

4.2.3 Student's comments and impressions about the Class

Survey had two open questions from which student's attitudes towards the Course can be elaborated.

The first one was already mentioned, a question about what they learn, they found useful. Even some (4) "Nothing" answers had explanations, like this one saying "*… not really nothing, but most of the stuff I already know …*", and one very valuable insight – "*My task was very simple, so I did not learn much*", explaining that some students (because of online environment) just did "their" homework and never joined a team, helping others, learning from them, and their tasks, as well. This is a typical downside of online learning environment, difficulty to get all students on the same page.

Of course, there were many more "Yes" elaborations (40), just to mention typical ones: "*Better organization in work and division of work in the team*", "*I learned some important things regarding not only checking information but also how to use various information in general. This is exactly what I consider important for personal development and as something that will help me in life*". And a favorite one, right to the point: "***I learned to think differently***".

At the end of the survey, there were much fewer comments (**about anything they wanted**), but the general message was they were sad the Class was not conducted live. It is obvious they almost could feel the benefits of interaction. And if needed to quote one, the most representative wrap-up would be; "*I honestly don't remember the details of the course, and those two years during the pandemic are a black hole in my memory, however, I have some key moments that have remained imprinted on me from your lectures. I think the way of working on the course is great. From teamwork, the random selection of teams, answering questions, and giving our own opinions at the lecture as if we were on a "friendly coffee", but instructive. You break our comfort zone and allow students to get to know each other and socialize with people they might never communicate with*".

4.3 Comparation of "pre-pandemic" and "pandemic" surveys

Despite pandemia and earthquake in Zagreb, online way of work for a class which was designed as highly interactive, official student satisfaction survey results are even higher. Although results of the 2018/19 generation are not published in the previous article, the author extracted results from University records, to be

able to compare sat of three generations – the first three were conducted regularly, as designed, and the last three were educated on-line.

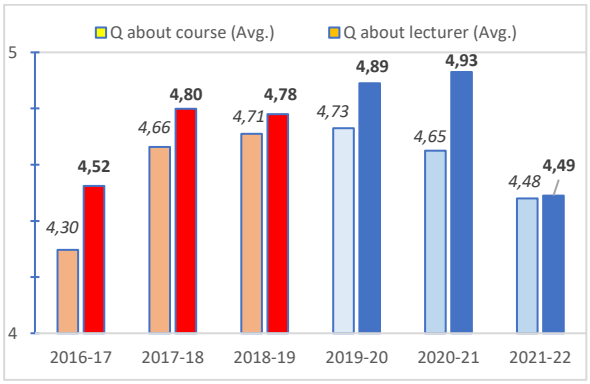


Figure 5. Comparison of Course and Lecturer grades

During pre-pandemic years Course average was 4,56, and improved to 4,62. Lecturer grade shows similar improvement: from 4,70 to 4,77. So, despite pandemic circumstances, both Course and Lecturer got better grades from students.

Out of comparison of the pre-pandemic and pandemic improvement skill lists, shown in Fig. 6, it is obvious that there was less enhancement/improvement during pandemia (overall average dropped from 4,84 to 4,38). Apart from two top skills ("Taking the stance/attitude" and "Critical evaluation of stance (own incl.)" pandemic generation also significantly improved in "Responsibility", "Result orientation" and "Argumentation" which is different from their improvement in pre-pandemic environment, probably because of new circumstances.

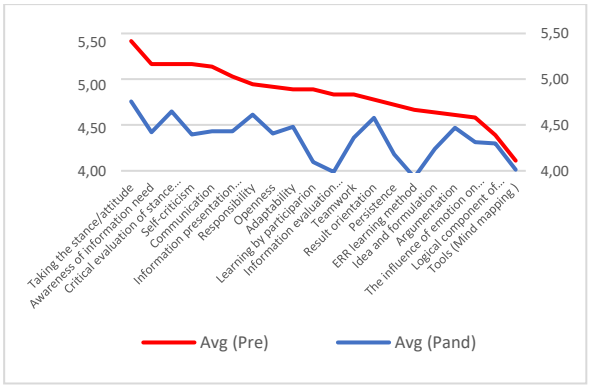


Figure 6. Comparison of skills improvement

On a list of the most valuable acquired skills (Table 2.), top five post-pandemic skills are also in top five of the pre-pandemic survey: 1. "Responsibility" (pre-pandemic rank as 5.), 2. "Taking the stance/attitude" (2.), 3. "Teamwork" (1.), 3. "Critical evaluation of stance" (2.) and 3. "Argumentation" (5.).

5 Discussion

As this Class is about developing skills for future ICT experts, it is important to discuss anything that can improve it, but also to raise some general questions/ recommendations author experienced in effort to teach students to think on a higher cognitive level (meta-thinking).

5.1 Results

The overall result of the Institutional survey showed that students liked (graded well) both the Course and the Lecturer as well, despite the pandemic sort of depression and online environment. Their grades are even better than in the pre-pandemic period.

The authors skills improvement survey well-explained the pandemic situation they went through, with their mostly positive coping with pandemia and earthquake, but kind of "depression" in the third year. Despite that they learned a lot and liked the Course.

In the same survey they confirmed the valuable contribution of the Class in developing/enhancing as many as 20 skills, which proves not only enhancing Critical thinking and/or Information literacy but a wide area of skills improvement. Obviously, they did not enhance those skills to full potential like their predecessors who had "a classic", interactive classes.

So, the obvious consequence is that although a lecturer can conduct a successful course in on-line, pandemic environment, for obtaining better, deeper improvement of skills face-to-face classroom is much better.

5.2 Online improvements

Upon writing the article, students are released back to classrooms, much more interested in open communication with us, online education went back to normal usage (if any), but as we never know what the future brings or some of us will combine classroom and online teaching. The author fully agree with a quote from Nature magazine: "Paths for greater access and opportunities to online education have now been forged, and there is a clear route for the next generation of adopters of online education" (Lockee, 2021). For anyone eager to plan any online education, this short paper offers 19 relevant references to start from.

The author, who was never fond of online/LMS educational methods, is definitely encouraged by this experience, so is considering at least usage of flipped classroom approach. It is ideal for the introduction of basic Information Literacy and Critical Thinking concepts and definitions as described in detail in (Divjak, 2022). By doing just that part, at the very beginning of a Class, even greater student engagement is expected.

5.3 Area of skills

Area of skills is slightly different from the list of ten most important skills Croatian ICT employers require from HE (Aničić & Bušelić, 2020), but not very far from that either. Especially "Ability to work in teams", "Analytical thinking" and "Taking responsibility". That means that by carefully designing topics, themes, and especially assignments, the lecturer can with more or less success select and improve a very specific set of skills. Not one skill, per se, but the whole family of them.

At the very end of the general points for further discussion, the author found the inspirational and valuable concept of Four-dimensional education for sustainable societies (Fadel & Groff, 2019) confirming that this Class is **adaptive, flexible, and balanced** as recommended. "While the curriculum must able to adapt to a rapidly changing world, it also must be able to be flexible to individual learner interests, needs, and goals, as well as local needs at the classroom and community level". And balanced as well, covering topics of "... humanities, science, technology, engineering, and mathematics (STEM)", not falling as "victim to a mindset of false choices, such as: "Which is better?"—teaching knowledge, or teaching skills?".

We can, and we should teach them both, as early as we got them in their educational journey!

References

Alfino, M., Pajer, M., Pierce, L., & Jenks, K. O. B. (2008). Advancing critical thinking and information literacy skills in first year college students. *College and Undergraduate Libraries, 15(1–2),* 81–98.

Aničić, K. P., & Bušelić, V. (2020). Importance of Generic Skills of ICT Graduates—Employers, Teaching Staff, and Students Perspective. *IEEE Transactions on Education, 64(3),* 245-252.

Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., ... & Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of educational research, 74(3),* 379-439.

Bušelić, V. (2019). Information Literacy and Critical Thinking Freshman Course Experience. *42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO),* 920-925.

Choy, S. C., & Cheah, P. K. (2009). Teacher perceptions of critical thinking among students and its influence on higher education. *International Journal of teaching and learning in Higher Education, 20(2),* 198-206.

Divjak, B., Rienties, B., Iniesto, F., Vondra, P., & Žižak, M. (2022). Flipped classrooms in higher education during the COVID-19 pandemic: findings and future research recommendations. *International Journal of Educational Technology in Higher Education*, 19(1), 1-24.

Fadel, C., Bialik, M., & Trilling, B. (2015). Four-dimensional education. *Center for Curriculum Redesign.*

Fadel, C., & Groff, J. S. (2019). Four-dimensional education for sustainable societies. *Sustainability, human well-being, and the future of education* (pp. 269-281). Palgrave Macmillan, Cham.

Grandzol, J. & Grandzol, C. (2006). Best Practices for Online Business Education. *International Review of Research in Open and Distributed Learning, 7(1),* 1–18.

Harley, B. (2001). Freshmen, information literacy, critical thinking and values. *Reference Services Review, 29(4),* 301–306.

Keengwe, J., & Kidd, T. T. (2010). Towards best practices in online learning and teaching in higher education. *MERLOT Journal of Online Learning and Teaching*, *6*(2), 533-541.

Lauer, T. (2005). Teaching Critical-Thinking Skills Using Course Content Material. *Journal of College Science Teaching, 34*, 34-44.

Lockee, B.B. Online education in the post-COVID era. *Nat Electron* **4,** 5–6 (2021). https://doi.org/10.1038/s41928-020-00534-0

Mandernach, B. J., Forrest, K. D., Babutzke, J. L., & Manker, L. R. (2009). The role of instructor interactivity in promoting critical thinking in online and face-to-face classrooms. MERLOT *Journal of online Learning and Teaching, 5(1),* 49-62.

Palvia, S., Aeron, P., Gupta, P., Mahapatra, D., Parida, R., Rosner, R., & Sindhi, S. (2018). Online education: Worldwide status, challenges, trends, and implications. *Journal of Global Information Technology Management, 21(4),* 233-241.

Rezaee, M., Farahian, M., & Ahmadi, A. M. (2012). Critical thinking in higher education: Unfulfilled expectations. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, *3*(2), 64-73.

Weiner, J. (2011). Is there a difference between critical thinking and information literacy? A systematic review 2000-2009. *Journal of information literacy*, 5(2), 81-92.

Wilson, S. D. (2018). Leading edge online classroom education: Incorporating best practices beyond technology. *American Journal of Business Education (AJBE), 11*(3), 41-48.

World Economic Forum. (2015). New Vision for Education: Unlocking the Potential of Technology. *Vancouver, BC: British Columbia Teachers' Federation*.

World Economic Forum. (2016). The Future of Jobs: Employment, Skills and Workforce Strategy for the Fourth Industrial Revolution. *Global Challenge Insight Report*.

Insights into understanding the perspective of digital platform providers

Katarina Tomičić-Pupek, Ana Kutnjak, Vjeran Strahonja
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{katarina.tomicic, ana.kutnjak, vjeran.strahonja}@foi.unizg.hr

Abstract. *Organizational transformation is an integral part of implementing, joining, or adopting digital platforms in platform-based ecosystems. Insights into the perspective of digital platform providers presented in this paper aim to contribute to understanding the digital platform providers' intention to implement their solution for supporting platform-based business models and operations. This research is motivated by the scarce existing literature on this matter and based on the application of qualitative data-extraction and assessment methods for identifying concepts that define the role of DP providers. Ten factors describing providers perspective were identified and which are related to previous factors describing other stakeholders' intentions as well. By evaluating their capabilities across the conceptual model presented in this paper, providers can assess their readiness to implement DP and thereby evaluate perceived values and risks impacting their intention to implement a DP.*

Keywords. Digital platform, provider, conceptual model

1 Introduction

The increasing use of digital technologies in recent years has led to the evolution and development of the digital ecosystem concept. The platform for their gathering is digital. Digital ecosystems vary in how dominant some partners, products and services are, in the openness, flexibility and adaptability of the structure and process to changes in the environment, in the space in which they operate in scalability and sustainability and other factors. The objectives of digital ecosystems are synergy through partnership, integration of business processes and supply chains, etc., with the aim of creating a shared value of the digital ecosystem. Cooperation within the digital ecosystem and the design of appropriate business models is the result of the dynamic capabilities of organizations (Helfat & Raubitschek, 2018) whose connectivity is based on digital infrastructure, most often digital platforms (DPs). The complexity of digital platforms is seen in their intertwining with different institutions, markets and technologies (de Reuver et al., 2018). Their value has been identified in a number of industrial areas, as digital platforms emphasize strengthening cooperation (Hein et al., 2020) and open innovation (Bonollo & Poopuu, n.d.), facilitate information sharing (Sedera et al., 2016; Sutherland & Jarrahi, 2018) and focus on meeting the needs of different stakeholders within the digital ecosystem.

„*Digital platforms are often multi-sided, providing interfaces with and among two or more groups of economic actors on different 'sides' of the platform, including providers of complementary assets*" (Helfat & Raubitschek, 2018). Platforms design business and organizational models that focus on joint value creation (Karhu et al., 2018) for end users of products or services, stakeholders within the digital ecosystem, and platform owners. In doing so, platform owners enable other stakeholders to create products or services on the platform, which leads to the strengthening of the entire digital ecosystem and the creation of a diversified advantage for the ecosystem (Valdez-De-Leon, 2019). The platform architecture enables the centralization and integration of common features into core modules (Cenamor et al., 2019), thus facilitating the use of shared resources and the integration of knowledge from different fields. In this way, new features of business models were initiated that put information management and the network of relationships at the centre of the model.

Platforms differ on several key grounds: the interactions they allow, the speed of growth and development, the level of openness, competitive strategies, the way of problem solving, represented monetization methods and subsidies (Web Content Hunt Platform, 2018). Based on the analysis of more than 170 platform businesses with different characteristics, the authors identified the following types of platforms: Technology Platforms; Computing Platforms; Utility Platforms; Interaction Networks; Marketplaces; On-demand Service Platforms; Content Crowdsourcing Platforms; Data Harvesting Platforms; Content Distribution Platforms.

Platforms that provide building blocks or services for reuse by third-party developers with the goal of embedding them in their own products are called *Technology Platforms* (Web Content Hunt Platform, 2018). Such platforms are usually invisible to end users, are cashed in by selling services to developers, and are not based on the interaction between supply and demand (do not connect platform participants) (Web Content Hunt Platform, 2018). The mentioned characteristics are favourable to third parties (Cardoso, 2017), as they reduce the budget required for hardware procurement and data centre maintenance, which further leads to greater development agility (Kejariwal, 2013).

Computing Platforms are based on the interaction between third-party developers and users of the platform, where the platform "owns" users (Tran et al., 2017; Web Content Hunt Platform, 2018). These platforms can be expanded with new uses, generating more value for the platform based on the development of two-way network effects in which the following scenario is represented: users attract developers, developers create applications, applications attract users, users developers, etc.

Providing a free service is a basic feature of *Utility Platforms* (Web Content Hunt Platform, 2018), which, after identifying a critical mass of users, can be opened to other participants (Manur et al., 2018) such as advertisers and carriers.

Interaction Networks are designed to facilitate the interaction and communication of different participants in the platform (Web Content Hunt Platform, 2018), with users attracting other users, who further attract even more users, thus creating a network effect. Such one-sided platform enables the exchange of messages, photos, videos and other content for which a mass of users has gathered on the platform (Feezell & Ortiz, 2021).

Marketplaces platforms connect participants on the supply and demand side (Web Content Hunt Platform, 2018). Suppliers on the supply side offer their products, while customers appear on the demand side (Zervas et al., 2021), and the platforms allow them to make mutual (monetary) transactions.

Platforms that connect waiting time sensitive customers with independent service providers (Taylor, 2018) are called *On-demand Service Platforms.* These platforms combine functionalities such as discovery, ordering, payment, certification, etc., with the platform defining the price, the required quality standards, and the processes to meet them. For On-demand Service Platforms it is characteristic that customers have very little influence in defining the way the service will be delivered (Web Content Hunt Platform, 2018).

Collecting content (reviews, videos, etc.) from a subset of users and sharing it with a wider set of users (Linares-Bahillo et al., 2019), is the core functionality of *Content Crowdsourcing Platforms.* The more content the platform offers, the more users will join it (Web Content Hunt Platform, 2018), which will consequently affect its growth and popularity.

Platforms that generate their value based on collected data (Malandrino et al., 2018) are called *Data Harvesting Platforms*. The condition for joining the platforms is the generation of data by their users (Web Content Hunt Platform, 2018).

Content Distribution Platforms are the link between owners of user touch-points (web sites, mobile apps) and content owners who want to deliver content to interested users (Web Content Hunt Platform, 2018). Such platforms use functionalities such as content recommendation algorithms to deliver interesting and personalized content to content users (Xiao et al., 2021).

Extending functionalities through adding services based on business requirement blurs the distinctive primary type of classification of a digital platform. A typical example here fore are e-learning platforms, initially designed mostly as content distribution platforms. E-learning platforms have embraced other services among others, specific for network platforms like supporting social interaction between platforms participants. This widening of the pool of desired functionalities, was especially triggered by recent disruptive events, motivating more organizational and business transformation, and demonstrating the possibilities of digital transformation stronger than any other enabler prior investigated. Therefore, the evolution of DPs bears potentials yet to be discovered and harvested.

The paper is structured in 6 chapters. After introduction and defining basic concepts of the research, perception factors impacting producer's or customer's intention to join or adopt a DP are considered. The research methodology follows, after which it is shown extended conceptual model of factors impacting the decision to join, adopt, or implement a DP. Evaluation of perceived values and risks in assessing readiness to implement DP is the next chapter, and the paper ends by presenting research conclusions.

2 Perception factors impacting producer's or customer's intention to join or adopt a DP

Previous research (Kadoic et al., 2020, 2021; Tomicic-Pupek et al., 2020) was focused on identifying perception factors impacting producer's or customer's intention to join or adopt a DP. Perception in this context refers to a subjective assessment of the objective properties of an object of interest (Tomicic-Pupek et al., 2020). For each of these actors, i.e., customers and producers, 10 perception factors have been identified that determine the perception of customers and producers regarding the utility of a platform for the distribution of agricultural products.

The assumption is that these 10 factors impact perceived value or perceived risk gained through DP usage, regardless of the real value of these factors measured objectively.

Generalizing these factors, initially recognized by exploring utility of platforms for the distribution of agricultural products, opens research possibilities for investigating their appropriateness in a broader context of any DP. Also, it enables further investigation of other actor roles in platform-based ecosystems and analysing their perspectives regarding their intention to interact in such digital environments.

From the customer's perspective, the balance between perceived value and risk determines his/her intention of joining the platform in order to buy products, consume content, use services, interact in virtual communities or use other digital services offered as platform-provided functionalities. The generalized 10 customer factors are listed below with some examples of possible perception elements.

- Zero negative impact: Eco-friendliness, Zero waste, Low resource consumption;
- Delivery location & time: short delivery chains, delivery duration, reliability, delay management;
- Relationship history with the producer: existing experiences, response time and response quality, flexibility and commitment to build sustainable relations;
- Payment options: price, price volatility, payment currencies and security, hidden costs;
- Comfort & convenience: remote usefulness and enjoyment, suggestions based on forecasted behaviour, 24/7 availability;
- Recommendations (C2C2C): confidentiality of community reviews, social media-based platforms as a communication channel, vortex similar behaviour in recommendation availability;
- Community support: supporting sustainable and resilient local economy, community affiliation management, social responsibility;
- Producer's reliability: value for money, availability to quality ratio, cost saving; producer's social impact;
- Trust & traceability: ability to track all phases of product or service generation through the value chain including preparation, production process, harvesting, delivery, trust in social interactions;
- Product or Service safety: facing disruption challenges, experiencing perception of safety and "no harm has been done" activities through the value chain including preparation, production process, harvesting, delivery to product or service consumption;

From the producer's perspective a set of other factors determines the balance between perceived value and risk which impacts their intention to adopt platforms in order to sell products, distribute content, provide services or use other digital services offered as platform-provided functionalities for this type of actors. The generalized 10 producer factors are listed below with some examples of possible perception elements.

- Sales channels: customer-preferred sales channels and producer-preferred channels, omnichannel process alignment;
- Product or Service safety: dealing with disruption challenges, designing, creating and delivering positive perception of product or service safety and "no harm has been done" activities through the value chain including preparation, production process, harvesting, delivery to product or service consumption;
- Production or Service delivery technologies: efficient and effective production process with near zero negative impact supported by appropriate technologies, growing potential for new feasible value creation models;
- Product or Service Quality: assuring expected or behaviour- designed quality product or service through the value chain including preparation, production process, harvesting, delivery to product or service consumption;
- Resources: implementing digital technologies to empower, support or replace missing human labour, building corporate culture of continuous investment in new skills, knowledge and capacity, managing supply chain disruptions;
- Inbound logistics / Supply Chain: managing supply chain disruptions, assuring substitute raw materials, supply channels, vendors or technologies;
- Innovations: building innovation potentials, design and adoption of new value creation paradigms, setting or adjusting to new trends;
- Outbound logistics / Distribution chain: managing distribution chain disruptions; assuring substitute distribution channels, and delivery collaborating actors;
- Incentives and sustainability: building and supporting readiness to adopt new technologies, raising sustainability and resilience of business models, availability of initiatives for providing incentives;
- Regulatory compliance: dealing with uncertainty in meeting regulatory and legal frameworks; accountability in relation to customers and other stakeholders, best practice in managing risks and changing conditions in operating markets.

The generalized perception factors impact various perspectives on DP's business model because it encompasses various customer segments, value propositions and channels as well as value creation generation opportunities, efficient cost structures and revenue streams for producers. The digital business model must be supported by a DP, which has the potential of contributing to the organizational success. Therein is the importance of the DP provider's role: in the design, development, and implementation phases. Therefore, the presented generalized perception factors

should be considered when a provider decides on his intention to implement a digital platform.

3 Research method

Judging by the published scientific papers, digital platform providers have been the subject of scientific research for twenty years. A search of Web of Science (WoS) database can serve as an indicator of scientific production in this area. Search with the key term "platform provider" in the title, singled out 10 articles, an average of one paper since 2013. If abstracts are searched, the number of articles in WoS is 108. The aim of this study is gaining insight, synthesizing and consolidating knowledge from previous research in order to construct a set of factors that describe the role of digital platform providers. Therefore, a very focused literature review was conducted. Searching the Scopus database based on the combination of keyword "digital AND platform AND provider" in the document title resulted in a total of 12 articles, which were studied. Although this literature review cannot be seen as exhaustive, it is still indicative for understanding the initial role of DP providers: certain issues related to the potential of DP-supported value creation still exists, regardless of numerous papers tackling with different aspects of platform-based business, published in general. All papers were reviewed to summarize, compare and synthesize existing research topics and research gaps by applying qualitative data-extraction and assessment methods. By identifying and comparing various concepts defining the role of DP providers through concepts DP providers need to consider when implementing digital platforms.

4 Extended conceptual model of factors impacting the decision to join, adopt, or implement a digital platform

The literature gathered through a focused literature review was analysed respectively based on quality criteria for literature analysis were as follows: (1) Is the publication relevant to our research goal of understanding the role of digital platform providers? (2) Does the paper cover enough data on similar issues related to DP providers? (3) Which findings could be relevant for constructing a set of factors describing the role of digital platform providers? (4) Do the factors relate to existing producer or customer perception factors impacting their intention to join or adopt a DP? Three data sets were extracted: (a) Leading factor of provider's role impacting his intention to implement a digital platform, (b) Data extracted and coded from the literature supporting the relevance of the factor, (c) Relation to producer or customer perception factors. The data was finally compiled a table synthesizing the extracted information.

Table 1. Factors provider's role impacting his intention to implement a digital platform

Factors	Findings supporting the relevance of the factor	Relation to producer or customer perception factors
Resource dependencies & bottlenecks	• DPs are dependent on enabling technologies due to the possibility that technological and strategic bottlenecks impact access to DP's services (Ojala et al., 2018) • DPs contribute to organizational capability to harness potential of digital technologies (Saarikko, 2016) • Integration of supply chain participants of ecommerce together (Song et al., 2021) • Providing easy and convenient access to the multiple services through different devices (Ojala et al., 2020) • Importance of Dynamic Capabilities: Sensing (identifying and developing opportunities) & seizing (exploiting technological and market opportunities) by mobilizing organizational resources (Riefle et al., 2021; Teece, 2007)	*Producer:* Incentives and sustainability, Resources *Customer:* Comfort & convenience, Delivery location & time
Purpose and functionality dependencies	• DPs are dependent on the available content (Ojala et al., 2018) comments in the active content-based activity (Kim, 2018) • Providers apply tipping and coring approaches (Saarikko, 2016) • Marketplace-type of digital platforms rely strongly on marketplace promotions and recommendations aimed at attracting service providers and their integration (Vakeel et al., 2021)	*Producer:* Sales Channels, Resources *Customer:*

	• Balancing advantages and challenges of digital platform's engagement in digitalisation initiatives in retailing industry (Hardaker, 2022) • Importance of individual and group interaction (Anonymous, 2018) • DP's core open to extension of functionalities with modular services (Riefle et al., 2021) according to (Baldwin & Woodard, 2008; Tiwana et al., 2010) • Production/ Innovation/ Transaction logic-industry platforms (Saarikko, 2016) according to (Gawer, 2009)	Product or Service Quality, Comfort & convenience
Platform multi-layer architecture	• Multi-layered modular architecture: Frontend including (i) a content layer, (ii) a service layer, and backend including (iii) a network layer, and (iv) a device layer (Saarikko, 2016) and (Ojala et al., 2018) according to (Yoo et al., 2010) • Balancing front-end based business-related requirements and backend dependant technical possibilities (Saarikko, 2016) • Centralized and decentralized DPs (Kazan et al., 2014) • Modular-architecture processes (Song et al., 2021) • Bargaining options with stakeholders claiming their operator role (Hardaker, 2022) • Stable core with modular services (Riefle et al., 2021) according to (Baldwin & Woodard, 2008; Tiwana et al., 2010)	*Producer:* Resources, Innovations *Customer:* Product or Service Quality, Delivery location & time
Value creation options	• "Quest for new value-added services that could be converted into sustainable revenue streams or increased openness in innovation" (Kazan et al., 2014) • Value creation through selling targeted advertisements (Farshchian & Vilarinho, 2017) • Hybrid business ventures available offline and online (Ojala et al., 2020) • New sources of data for growth opportunities (Hardaker, 2022) • Adopting a service ecosystem perspective leading to integration of complementary services (Riefle et al., 2021) • Different DP actors are sharing value, enabling thereby DPs to evolve and grow (Anonymous, 2018)	*Producer:* Innovations, Production or Service delivery technologies *Customer:* Product or Service Quality, Producer's reliability
Co-provider relations across layers	• Centralized inclusive API and decentralized non-discriminatory API management in relation to third-party developers (Farshchian & Vilarinho, 2017; Kazan et al., 2014) • Customer loyalty to DP through consuming services from different vendors (Vakeel et al., 2021) • Big data about platform stakeholders and their behaviour (Hardaker, 2022) • Introducing innovations by complementary services or products (Wan et al., 2017) yet not all will benefit equally (Vakeel et al., 2021) • Social and relational capabilities are impacting the development of cooperation between trusted partners (Riefle et al., 2021)	*Producer:* Inbound logistics / Supply Chain, Outbound logistics / Distribution chain *Customer:* Zero negative impact, Payment options
Access control & limitations	• Prevention or limitation to DP resources which provider controls (Ojala et al., 2018) • Difference in managing access control & limitations at centralized and decentralized DPs (Kazan et al., 2014) • Policies regarding data sharing (Farshchian & Vilarinho, 2017) • Provider's data ownership as a leverage to favour some service providers over others (Vakeel et al., 2021) • Limitations (e.g. based on resource bottlenecks or legal regulations (Ojala et al., 2020)) impacting even access to markets (Anonymous, 2018)	*Producer:* Resources, Regulatory compliance *Customer:* Trust & traceability
Efforts through platform	• Theoretical concepts, activities, and empirical examples through different phases of DP lifecycle: Establishment, Entry, Commercialization, and Globalization (Ojala et al., 2018)	*Producer:*

evolution phases	• Technical bottlenecks are mostly appearing in the in the early entry phase (Ojala et al., 2018) • Users' activities and interaction on the platform are a critical element for DP's success (Anonymous, 2018; Kim, 2018), especially for the diffusion and activation the early stages (Anonymous, 2018) • Facilitating and managing uploaded contents and frequency of interactions (Kim, 2018) • Supporting internationalization phases of ventures: "Early internationalization of the digital service; Extension of digital service functionalities for global markets; Global Technology alignment" (Ojala et al., 2020) • In the globalization phase content providers, delivery channels, and users should be able to easily join and exit DP's service pool (Ojala et al., 2018)	Sales channels, Product or Service Quality *Customer:* Recommendations (C2C2C), Community support
Incentives for active interaction	• Collective actions and active interaction among trusted community members (Song et al., 2021) • Data about user relationships has more potential value than user contents (Kim, 2018) • DPs are providing availability of services at home and within communities (Farshchian & Vilarinho, 2017) • Digital platforms which are based on search and social media data can provide targeted access to specific groups of users (Vakeel et al., 2021) • Social relationships and establishing a social communities of users can be critical to platform success even more than content quality especially at the early phases (Anonymous, 2018; Wan et al., 2017)	*Producer:* Sales channels, Incentives and sustainability *Customer:* Recommendations (C2C2C), Community support
Behaviour modelling and experience design	• In the commercialization phase, finding feasible pairs of service-markets is of immense importance for building a sizable customer base covering service strategic bottlenecks (Ojala et al., 2018) • Network transaction frequency, (re)visit rate, uploading contents, and interaction activities over time describe customer behaviour and provide the base of future experience design (Kim, 2018; Vakeel et al., 2021) • Reprogrammable & editable services on DPs (re)designed to meet, trigger or engage existing or new designed customers' preferences and needs (Ojala et al., 2020)	*Producer:* Innovations, Product or Service Quality *Customer:* Recommendations (C2C2C), Relationship history with the producer, Comfort & convenience
Accountability, regulatory and legal compliance	• Criticism of avoidance of accountability, failing to address properly labour protections and other regulatory frameworks (Hardaker, 2022) • Agreement with fact that "some kind of regulation on digital platforms is necessary due to their dominant market position and tax avoidance" (Riefle et al., 2021) • Regulators are facing the challenge to coordinate regulatory and legal compliance issues (Wan et al., 2017)	*Producer:* Regulatory compliance *Customer:* Product or Service safety, Producer's reliability, Trust & traceability

Figure 1 visualizes the extension of previous research (Tomicic-Pupek et al., 2020) by presenting generalized customer and producer factors and the new dimension of identified DP provider factors that impact the provider's intention to implement a DP. Factors are influencing platform actors' perceived balance of value or risk based on their role: for DP provider's intention to implement a DP (coloured green), but also through generalized factors customer's intention to join and producer's intention to adopt the DP (grey and yellow coloured respectively).

The identified factors shown in the extended conceptual model can imply convergent or divergent relationships between various stakeholders, influencing thereby their perceived balance of value or risk and the overall decision on joining, adopting, or implementing a DP.

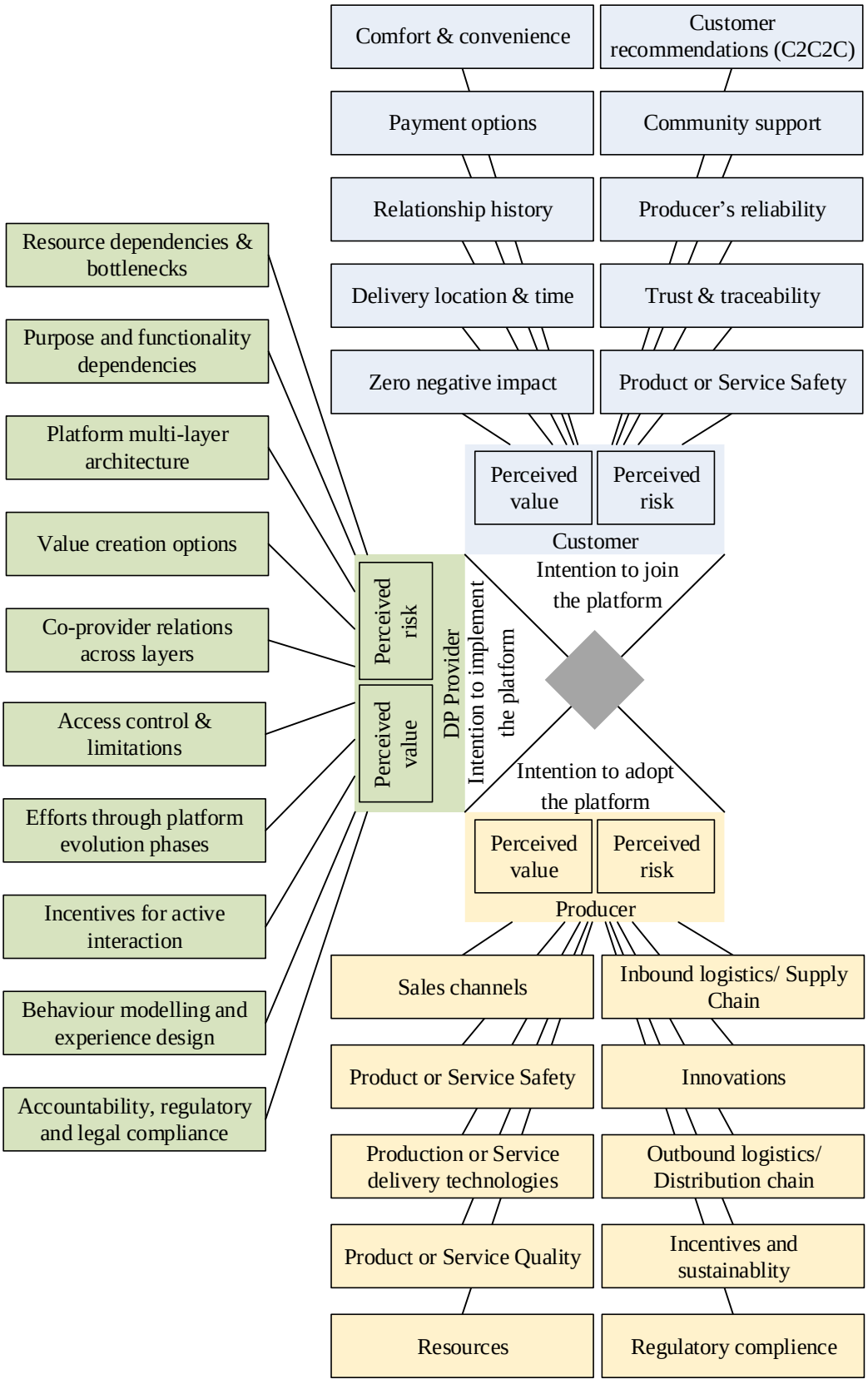


Figure 1. Extended conceptual model of factors impacting the decision to join, adopt, or implement a digital platform (extension to (Tomicic-Pupek et al., 2020))

5 Evaluation of perceived values and risks in assessing readiness to implement DP

As part of the feasibility assessment, it is desirable that potential platform providers assess their readiness to implement a DP. This assessment is based on revising their capabilities in relation to identified factors, and thereby evaluating perceived values and risks impacting their intention to implement DP:

- Technical, human, financial and other kind of resources at our disposal or do we have access to those resources for successful resolving dependencies and handling all kind of bottlenecks (Farshchian & Vilarinho, 2017; Ojala et al., 2018).
- Purpose and functionality which can be successfully replicated for global markets (Ojala et al., 2018) should provide "anticipated outcomes despite unanticipated circumstances" (Saarikko, 2016) contributing to product or service quality, as well as comfort & convenience experienced by customers.
- Designed or developed multi-layer architecture with different approaches in coupling platform layers to configure centralized (tightly coupling layers) or decentralized (decoupling layers) platforms (Kazan et al., 2014). Due to vendor-specific relations across various the multi-layered architecture technology-driven innovations pose a challenge (Farshchian & Vilarinho, 2017).
- Technical possibilities should enable replication of open modular platforms with high level of editability and reprograming based on functional requirements (Ojala et al., 2020) to ensure feasible value creation options for the DP provider.
- Co-provider relations across layers impacts the mutual agreement on what constitutes value creation opportunities (Saarikko, 2016) with possibly unequal potential for benefits (Vakeel et al., 2021).
- Appropriate access control and limitation management for different types of DPs (Kazan et al., 2014) may affect co-provider relations across layers (Vakeel et al., 2021) but also impact efforts during evolution phases on different markets (Riefle et al., 2021). Trust and traceability concerns at the customer side can be outvoted by incitive motivation to join a DP, yet minimum accountability, regulatory and legal compliance must be met.
- Efforts through platform evolution phases must not exceed potential revenue streams especially, for international new ventures (Ojala et al., 2018, 2020). Also, some organizations struggle with the transformation of an industrial corporation into a DP provider and seek support through different enablers of change (e.g., some authors suggest 7 enablers with actionable guidelines (Riefle et al., 2021)).
- Incentives for active interaction should be accessible for DP providers to boost social relations through the platform building thereby digital communities with strong user relationships, by some authors even argued to be "more important than user contents" (Kim, 2018).
- Platform use behaviour modelling and design of new experience opportunities can be crucial for adding new services at the service layer impacting thereby existing and future performance (Wan et al., 2017). Every step of the customer journey from customer engagement and onboarding to retention needs to be designed and developed for aiming DP's value creation propositions.
- Accountability, regulatory and legal compliance have an influence on the perceived product or service safety, producer's reliability, trust & traceability of other stakeholders. Failing to meet the requirements regarding these characteristics of a DP, raises issues for DP providers but also for regulators which contribute more to the perceived risks of implementing a platform on a global scale (Hardaker, 2022; Wan et al., 2017). Therefore, appropriate efforts in early phases of DP evolution phases must be performed regarding this factor.

6 Conclusion

Disruptive events, witnessed recently, impacted and still impact normal flows in global operations across industries. Therefore, understanding the role and factors influencing DP providers' intention to implement their solution for providing platform-based business models and operations needs to be explored further. Specific and more in-depth research is needed regarding the interaction with other stakeholders who are just adding to the already high complexity of operations. Although the research is based on a literature review of a small number of scientific articles that are narrowly focused on DP providers, which is a limitation of this research, it indicates that certain factors influencing the role and behaviour of DP providers exist regardless of diversity platforms described in various papers. To the best of our knowledge, there is still a lack of papers dealing with transformational changes of organizations when implementing, joining, or adopting DPs. Therefore, this research is intended to shed some insight into understanding what and how influences different stakeholders to consider platform-based interactions.

Acknowledgments

The current research is partially supported by project Digital Platform Enterprise - DEMO, funded by EU, ERASMUS+ programme, action type KA220-HED-

Cooperation partnerships in higher education, Call 2021, Round 1.

References

Anonymous. (2018, May 14). Rising above the competition: How market entry strategy can secure competitive advantage for digital platform providers. *Strategic Direction*, *34*(5), 12–13.

Baldwin, C. Y., & Woodard, C. J. (2008). The Architecture of Platforms: A Unified View. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1265155

Bonollo, N., & Poopuu, P. (n.d.). *Strategic Entrepreneurship, Digital Business (The impact of digital platforms on roles and responsibilities in value creation among stakeholders of an ecosystem)*. 102.

Cardoso, J. (2017). Cloud Reliability: Decreasing Outage Frequency Using Fault Injection. In A. Romanovsky & E. A. Troubitsyna (Eds.), *Software Engineering for Resilient Systems, Serene 2017* (Vol. 10479, p. IX–XI). Springer International Publishing Ag. https://www.webofscience.com/wos/woscc/full-record/WOS:000452448200013

Cenamor, J., Parida, V., & Wincent, J. (2019). How entrepreneurial SMEs compete through digital platforms: The roles of digital platform capability, network capability and ambidexterity. *Journal of Business Research*, *100*, 196–206. https://doi.org/10.1016/j.jbusres.2019.03.035

de Reuver, M., Sørensen, C., & Basole, R. C. (2018). The Digital Platform: A Research Agenda. *Journal of Information Technology*, *33*(2), 124–135. https://doi.org/10.1057/s41265-016-0033-3

Farshchian, B. A., & Vilarinho, T. (2017). Which Mobile Health Toolkit Should a Service Provider Choose? A Comparative Evaluation of Apple HealthKit, Google Fit, and Samsung Digital Health Platform. In A. Braun, R. Wichert, & A. Maña (Eds.), *Ambient Intelligence* (Vol. 10217, pp. 152–158). Springer International Publishing. https://doi.org/10.1007/978-3-319-56997-0_12

Feezell, J. T., & Ortiz, B. (2021). "I saw it on Facebook": An experimental analysis of political learning through social media. *Information Communication & Society*, *24*(9), 1283–1302. https://doi.org/10.1080/1369118X.2019.1697340

Gawer, A. (2009). Platform Dynamics and Strategies: From Products to Services. In *Chapters*. Edward Elgar Publishing. https://ideas.repec.org/h/elg/eechap/13257_3.html

Hardaker, S. (2022). More Than Infrastructure Providers – Digital Platforms' Role and Power in Retail Digitalisation in Germany. *Tijdschrift Voor Economische En Sociale Geografie*, tesg.12511. https://doi.org/10.1111/tesg.12511

Hein, A., Schreieck, M., Riasanow, T., Setzke, D. S., Wiesche, M., Böhm, M., & Krcmar, H. (2020). Digital platform ecosystems. *Electronic Markets*, *30*(1), 87–98. https://doi.org/10.1007/s12525-019-00377-4

Helfat, C. E., & Raubitschek, R. S. (2018). Dynamic and integrative capabilities for profiting from innovation in digital platform-based ecosystems. *Research Policy*, *47*(8), 1391–1399. https://doi.org/10.1016/j.respol.2018.01.019

Kadoic, N., Tomicic-Pupek, K., & Vrcek, N. (2020). Decision making on Digital Platforms in Agriculture. *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, 1457–1462. https://doi.org/10.23919/MIPRO48935.2020.9245236

Kadoic, N., Tomicic-Pupek, K., & Vrcek, N. (2021). Prioritisation of Factors that Influence the Digital Platform Selection in Agriculture. *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, 1318–1323. https://doi.org/10.23919/MIPRO52101.2021.9596656

Karhu, K., Gustafsson, R., & Lyytinen, K. (2018). Exploiting and Defending Open Digital Platforms with Boundary Resources: Android's Five Platform Forks. *Information Systems Research*, *29*(2), 479–497. https://doi.org/10.1287/isre.2018.0786

Kazan, E., Tan, C.-W., & Lim, E. T. K. (2014). *Towards a framework of digital platform disruption: A comparative study of centralized & decentralized digital payment providers*. Proceedings of the 25th Australasian Conference on Information Systems, ACIS 2014. Scopus.

Kejariwal, A. (2013). A Tool for Practical Garbage Collection Analysis in the Cloud. In R. Campbell, H. Lei, & V. Markl (Eds.), *Proceedings of the 2013 Ieee International Conference on Cloud Engineering (ic2e 2013)* (pp. 46–53). Ieee. https://doi.org/10.1109/IC2E.2013.13

Kim, J. (2018). Market entry strategy for a digital platform provider. *Baltic Journal of Management*, *13*(3), 390–406. https://doi.org/10.1108/BJM-07-2017-0228

Linares-Bahillo, E., Aristegui-Fradua, I., & Beloki-Maranon, U. (2019). YouTube, a platform for (in)formation, relationship, communication, fun, and (gender) identities within the context of digital

natives. *Revista Mediterranea Comunicacion-Journal of Communication*, *10*(1), 55–70. https://doi.org/10.14198/MEDCOM2019.10.1.18

Malandrino, F., Chiasserini, C.-F., & Kirkpatrick, S. (2018). Cellular Network Traces Towards 5G: Usage, Analysis and Generation. *Ieee Transactions on Mobile Computing*, *17*(3), 529–542. https://doi.org/10.1109/TMC.2017.2737011

Manur, A., Venkataramanan, G., & Sehloff, D. (2018). Simple electric utility platform: A hardware/software solution for operating emergent microgrids. *Applied Energy*, *210*, 748–763. https://doi.org/10.1016/j.apenergy.2017.07.073

Ojala, A., Evers, N., & Rialp, A. (2018). Extending the international new venture phenomenon to digital platform providers: A longitudinal case study. *Journal of World Business*, *53*(5), 725–739. https://doi.org/10.1016/j.jwb.2018.05.001

Ojala, A., Rollins, M., Fraccastoro, S., & Gabrielsson, M. (2020). *The internationalization of B2B digital platform providers: The role of cross-national distance and digital characteristics. 2020-January*, 4652–4661. Scopus.

Riefle, L., Eisold, M., & Benz, C. (2021). Industrial Corporation's Transformation into a Digital Platform Provider: A Case Study on Enablers. *2021 IEEE 23rd Conference on Business Informatics (CBI)*, 131–140. https://doi.org/10.1109/CBI52690.2021.00024

Saarikko, T. (2016). Platform Provider by Accident: A Case Study of Digital Platform Coring. *Business & Information Systems Engineering*, *58*(3), 177–191. https://doi.org/10.1007/s12599-016-0426-4

Sedera, D., Lokuge, S., Grover, V., Sarker, S., & Sarker, S. (2016). Innovating with enterprise systems and digital platforms: A contingent resource-based theory view. *Information & Management*, *53*(3), 366–379. https://doi.org/10.1016/j.im.2016.01.001

Song, H., Li, M., & Yu, K. (2021). Big data analytics in digital platforms: How do financial service providers customise supply chain finance? *International Journal of Operations & Production Management*, *41*(4), 410–435. https://doi.org/10.1108/IJOPM-07-2020-0485

Sutherland, W., & Jarrahi, M. H. (2018). The sharing economy and digital platforms: A review and research agenda. *International Journal of Information Management*, *43*, 328–341. https://doi.org/10.1016/j.ijinfomgt.2018.07.004

Taylor, T. A. (2018). On-Demand Service Platforms. *M&som-Manufacturing & Service Operations Management*, *20*(4), 704–720. https://doi.org/10.1287/msom.2017.0678

Teece, D. J. (2007). Explicating dynamic capabilities: The nature and microfoundations of (sustainable) enterprise performance. *Strategic Management Journal*, *28*(13), 1319–1350. https://doi.org/10.1002/smj.640

Tiwana, A., Konsynski, B., & Bush, A. A. (2010). Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics. *Information Systems Research*, *21*(4), 675–687. https://doi.org/10.1287/isre.1100.0323

Tomicic-Pupek, K., Pihir, I., & Tomicic Furjan, M. (2020). The Role of Perception in the Adoption of Digital Platforms in Agriculture. *2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO)*, 1429–1434. https://doi.org/10.23919/MIPRO48935.2020.9245084

Tran, T. X., Hajisami, A., Pandey, P., & Pompili, D. (2017). Collaborative Mobile Edge Computing in 5G Networks: New Paradigms, Scenarios, and Challenges. *Ieee Communications Magazine*, *55*(4), 54–61. https://doi.org/10.1109/MCOM.2017.1600863

Vakeel, K. A., Malthouse, E. C., & Yang, A. (2021). Impact of network effects on service provider performance in digital business platforms. *Journal of Service Management*, *32*(4), 461–482. https://doi.org/10.1108/JOSM-04-2020-0120

Valdez-De-Leon, O. (2019). How to Develop a Digital Ecosystem – a Practical Framework. *Technology Innovation Management Review*, *9*(8), 43–54. https://doi.org/10.22215/timreview/1260

Wan, X., Cenamor, J., & Chen, J. (2017). Exploring Performance Determinants of China's Cable Operators and OTT Service Providers in the Era of Digital Convergence—From the Perspective of an Industry Platform. *Sustainability*, *9*(12), 2247. https://doi.org/10.3390/su9122247

Web Content Hunt Platform. (2018, December 30). *The 9 Types of Software Platforms*. Platform Hunt. https://medium.com/platform-hunt/the-8-types-of-software-platforms-473c74f4536a

Xiao, W., Zhao, H., Pan, H., Song, Y., Zheng, V. W., & Yang, Q. (2021). Social explorative attention based recommendation for content distribution platforms. *Data Mining and Knowledge Discovery*, *35*(2), 533–567. https://doi.org/10.1007/s10618-020-00729-1

Yoo, Y., Henfridsson, O., & Lyytinen, K. (2010). The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research. *Information Systems Research*, *21*(4), 724–735. https://doi.org/10.1287/isre.1100.0322

Zervas, G., Proserpio, D., & Byers, J. W. (2021). A first look at online reputation on Airbnb, where every stay is above average. *Marketing Letters*, *32*(1), 1–16. https://doi.org/10.1007/s11002-020-09546-4

An instrument for measuring needs of vulnerable groups in terms of digital inclusion

Nikolina Žajdela Hrustek, Diana Šimić
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, Varaždin
{nzajdela, dsimic}@foi.unizg.hr

Žarko Čižmar
Telecentar
Kikićeva 13, Zagreb
zarko@telecentar.hr

Abstract. *Nowadays, mostly due to the COVID-19 pandemic, digital technology and the Internet have taken a very important role and become an integral part of the life of almost every individual. What happens to those who, for any reason, are still unable to participate in a society that actively uses digital technology and the Internet, especially those who belong to vulnerable groups of the population? Numerous studies have shown that these circumstances lead to social exclusion. The aim of this paper is to present a methodology for development of an instrument for measuring needs of vulnerable groups in terms of digital inclusion. The instrument was created as a part of the project Digitalna.hr, implemented by the Croatian Digital Literacy Network. It can be used to collect data on the state of digital inclusion of vulnerable groups in a population. Such data are necessary for formulating recommendations and measures for building public policies and strategies, as well as their implementation and enforcement by regional and local government, in order to include this segment of the population in the digital society.*

Keywords. digital literacy, vulnerable groups, digital inclusion, digital society

1 Introduction

Development of digital society and promoting accessibility and use of digital technologies is a priority in almost all countries today (Lyons, Kass-Hanna, Greenlee, 2020, Lyons et al., 2019, Palmeiro, Pereda Herrero, Aires, 2019, Chetty et al., 2018, ITU, 2018 a, b, Serrano-Santoyo, Rojas-Mendizabal, 2017, Real, at al., 2015). Notwithstanding numerous attempts by public authorities, non-governmental associations, and private organisations to empower citizens and include them in the digital society, the digital gap is still present (Tomczyk et.al. 2019). Research shows that every individual cannot equally take advantage of all opportunities offered by modern digital technology. Simultaneously, it has been confirmed that digital gap does not affect only economic variables, such as income per capita and costs of access and use. It has also become a sociological problem, meaning that lack of access to digital technologies can lead to social exclusion (Helsper, 2008, Helsper, Galácz, 2009, Helsper 2012). Similar perspective can be noticed in research by Van Dyke (2020), who defines the digital gap as a social and political problem, and generally a problem of unequal distribution of goods, but not as a technological problem. He also emphasises the importance of factors, such as motivation, skills, talents, desire to use, literacy level, lack of interest, lack of literacy, difficulties with access, etc. that affect the digital exclusion. Social exclusion is partly emphasised due to digital exclusion if an individual belongs to a vulnerable group of population (Pérez-Escolar, Canet, 2022). According to Mechanic and Tanner (2007), some of the most important "sources of vulnerability" are poverty and race, lack of social networks and social support, personal and physical limitations. Thus, Lyons and Kass-Hanna (2020) assert that vulnerable society groups include, but are not limited to, the poor, women, the young, the elderly, the disabled, ethnic minorities, people living in barely accessible rural and island areas, migrants, and refugees.

Access to and use of digital technologies still represent a challenge for vulnerable groups. Even though some areas have remarkable Internet access infrastructure, large parts of population, especially those with lower socio-economical status or living in rural areas, still have a limited or no access to digital devices or platforms, or available digital services do not satisfy their accessibility needs. On the other hand, there are multiple benefits of digital technologies and services for the vulnerable groups. Various researchers identified following advantages: saving money and time, space and time flexibility while searching for information, (searching for) work from home, supporting activities important for autonomy, normal and independent life, reducing practical difficulties caused by disability, easier, safer and more open communication, networking, more possibilities for fun,

opportunities for developing new hobbies and activities, discovering new horizons of online studying, increasing accessibility and quality of health services and health information (Tirado-Morueta et al. 2021, Tsatsou 2021, a, b, Tsatsou, 2020, Alabi, Mutula, 2020, Yu et al. 2019, Matthews, Nazroo, Marshall, 2019, Betts, Hill, Gardner, 2019, Tsatsou, Higgs, Stafford, 2013).

Solving the problem of digital gap has become a critical developmental challenge on the global level, exacerbated by the COVID-19 pandemic. Advocates of digital inclusion/e-inclusion maintain that digital gap can be completely bridged or reduced to the least possible measure by promoting and enabling digital inclusion. Digital inclusion creates many social benefits, e.g. it promotes democracy, mutual understanding, empowering of socially endangered groups, etc. (Mancinelli, 2008). According to the European consulting group, digital inclusion/e-inclusion relates to enabling participation of individuals, community and economy in all dimensions of knowledge based society that uses ICTs in a way that removes access and accessibility barriers, and takes advantage of social benefits of digital technologies and services (EC, 2006). Žajdela Hrustek (2015) proposes a definition of e-inclusion as "inclusion of individuals/groups/communities in the information society, where everyone has equal chance to access and use ICTs and Internet, with the purpose of improving quality of life and enabling active participation in creating new social and technological systems."

From the very beginning, the groups of individuals considered to have potentially higher risk to be socially and digitally excluded, such as the disabled, the elderly or women were identified as interest groups in strategies connected to the term digital inclusion/e-inclusion. The goals of further strategies were directed to the society as a whole, and all individuals who could, for any reason, be excluded from the information society, or could not completely use its advantages. Performing the process of digital inclusion/e-inclusion depends considerably on politics conducted by the government in a specific country regarding realising goals as follows: inclusion of all groups and individuals in digital society, efficient and transparent service providing to citizens, empowering people via access to all relevant information, efficient management of information for citizens, promoting awareness among citizens on the importance of inclusion in digital society, creating social and cultural consensus (Žajdela Hrustek, Šimić, Vrček, 2017, Žajdela Hrustek, Prosser, Dušak, 2016). Despite the pressing need for digital inclusion among the most vulnerable groups, there is still a lack of public awareness, a lack of mention of it in public policies, strategies, and operational plans, a lack of media and digital literacy among vulnerable groups, and a poor utilization of e-Citizen services. On the other hand, the European Union has set ambitious digital transformation goals. By 2030, 100% of key public services should be available online, 100% of citizens should have access to digital health records, 80% of citizens should be using digital ID, and at least 80% of citizens should have basic digital skills (EC, 2021). To facilitate acquiring of digital skills and improving digital literacy, the Joint Research Centre of the European Commission in 2022 has developed the digital competence framework for citizens DigComp 2.2 (Vuorikari, Kluzer, Punie, 2022).

This research was conducted as a part of the project Digitalna.hr implemented by the Croatian Digital Literacy Network. The project's major goal is to create a model for civil society activities that promote evidence-based advocacy for digital literacy. Its first objective, development of digital citizenship, required assessment of the needs of vulnerable groups of population in Croatia in terms of digital inclusion. The vulnerable groups targeted by the project were the retirees, the disabled, and the inhabitants of rural and island areas. However, search of research literature did not discover a valid and relevant instrument for measuring these needs.

The rest of the paper describes the development of the measurement instrument, data collection, and main characteristics of the targeted vulnerable groups.

2 Methodology of developing the measurement instrument

The development of the measurement instrument took place from February to September 2021. The first phase started by defining the research problem, questions, and objectives, and the second phase concluded with development of the measurement instrument.

2.1 Defining the research problem

Based on the literature referenced in the Introduction, the following difficulties faced by vulnerable groups were singled out:

- lack of access to and use of digital technologies and the Internet,
- low digital, media, and information literacy,
- poor offering of safe and accessible digital services.

Research questions and objectives were developed to support assessment of the needs of Croatia's vulnerable groups in order to address these difficulties. The research questions were:

1. Do vulnerable groups use digital technologies and the Internet?
2. How do vulnerable groups access the Internet?
3. Which obstacles to using digital technologies and the Internet do individuals in vulnerable groups face?
4. To what degree are digital public services used by vulnerable groups?

5. What is the user experience while using digital public services?
6. What is the level of media literacy of individuals in vulnerable groups?

The objectives were:

1. To develop a measurement instrument for assessing the needs of vulnerable target groups in terms of digital inclusion,
2. To conduct a survey on target groups , and
3. To draft policy recommendations for improving digital inclusion of vulnerable target groups, based on the results of empirical research and relevant literature.

2.2 Creating the initial item set

The initial item set was sourced from Žajdela Hrustek's (2015) research on digital inclusion. It comprised 341 items. Six team members of the working group Digital Citizens did the card sorting exercise. Evaluators assigned each item to the research question it addressed. Inter-evaluator agreement was assessed by Krippendorff's alfa (Krippendorff, 1980) using the statistical environment R (R Core Team, 2021). Only 255 items assigned by at least four evaluators to the same question were retained. Krippendorff alfa for the retained items was relatively low (0.424), however this was just the selection of the initial item set, that was further analysed for clarity and content validity.

2.3 Item clarity and content validity

Content validity and clarity of the selected item set were assessed using an expert group of researchers (six) and practitioners (four), as well as a small group of members of the target vulnerable groups (Milas, 2005, Tkalac Verčić et al., 2011).

Criteria for selecting the experts included competences and experience related to digital inclusion, relevant publications and projects, and teaching and training activities related to digital inclusion and vulnerable populations. Participation in the research was voluntary.

In the first step, the experts were asked to independently:

1. assess whether the items address the assigned research questions,
2. evaluate item's relevance for the assigned research question, using the four level scale: 1–Item is mandatory, 2–Item is desirable, 3–Item is not necessary, it should be removed, 0–I cannot determine the item's relevance,
3. assess the item clarity, using the three level scale: 1-Item is clear, 2-Item is neither clear nor unintelligible, 3-Item is unintelligible,
4. suggest to change an existing item, add a new item, combine or split items, or assign an item to a different research question.

This evaluation also referred to checking the grammatical and spelling accuracy, clarity, and unambiguity of the proposed items. In the follow-up, experts participated in two focus groups, one with researchers, and one with practitioners, where suggested changes and additions to the item sets were discussed, and a consensus was reached.

Content validity was assessed quantitatively using the Lawshe's Content Validity Ratio (CVR) (Lawshe, 1975). There were 158 items with CVR above the recommended treshold value. These items were used in the next stage with representatives of the target vulnerable groups.

The instrument was piloted in four small groups of participants (two to four) from target vulnerable groups. Participants for this stage of research were identified and contacted through representative associations, e.g. association of pensioners, people with disabilities, and rural inhabitants. The main objective of this stage was to make sure that items were understandable, and format of the questionnaire, in both paper and online form, was accessible to members of the target vulnerable groups. Participants were asked to fill in the questionnaire in paper or online form, and comment on all aspects of its form and content (including e.g. understandability, font size, organization, etc.).

The final measurement instrument comprised 158 items, after the phrasing of some items was modified to satisfy the needs of the target vulnerable groups. The questionnaire was organised in four subsection 1) Digital inclusion of vulnerable groups - access, use, skills, and barriers to the use of digital technologies; 2) Use of digital public services (e-Citizens) and 3) Media literacy, and 4) Demographic data.

3 Data collection

In order to assess the needs of target vulnerable groups in terms of digital inclusion, a survey was conducted on a snowball sample of 540 participants from Zagreb, Varaždin, and Primorsko-goranska counties, and the city of Zagreb. The snowball sampling was used, because it is particularly well suited for reaching out to hard-to reach and vulnerable groups (Ellard-Gray et al., 2015). The data collection took place from September to December 2021, when it was specially difficult to contact members of the elderly and people with disabilities, due to high COVID-19 risk and epidemiological measures. The initial set of participants for snowball sampling was identified and contacted through associations that facilitated the piloting of the instrument.

Descriptive statistics was done in IBM® SPSS® Statistics.

4 Results

Due to page restrictions we can present here only the most interesting results of the survey.

4.1 Demographic data

The sample comprised 50.5% pensioners, 32% participants with some form of disability, 8.1% inhabitants of islands, and 29.8% of inhabitants of rural areas. The majority of examinees were women (63.9%). The age group 65-74 (21.5%) was the largest, while the age group 16-24 was the smallest (4.3%). Participants living in households of two adults without (28.0%) or with children (25.6%) were the majority.

The majority of participants completed secondary trade schools, lasting four years or longer (27.1%). There was a very small proportion that did not finish any school (0.2%). In their opinion and according to their assets, more than 50% of participants do not live better or worse than the most, while a small proportion (2.8%) consider themselves living better than the most.

4.2 Access to digital technologies and the Internet

As visible in Figure 1, the majority of participants most often access the Internet by a smartphone, a personal computer, or a laptop (over 84% combined). An e-book reader (0.20%) and a games console (0.00%) are devices the least often used to access the Internet.

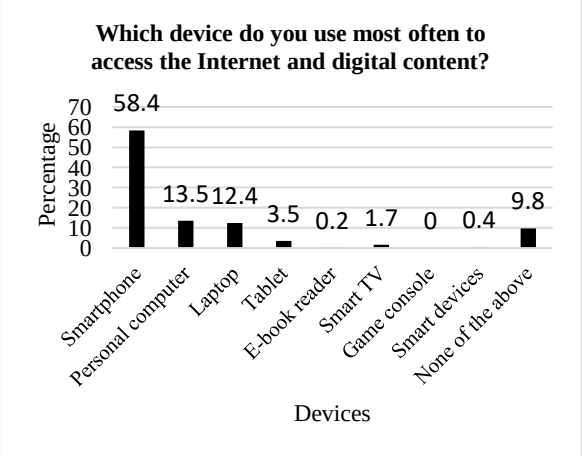


Figure 1. Access to digital technologies

Regarding the type of access to the Internet, the majority of participants use mobile or fixed access at home (82.2%), while the least use public places where they are charged for access and use. The results also show that 10.7% of participants do not use the Internet.

4.3 Reasons for use and attitudes toward digital technologies

The most important reasons why participants use or want to use digital technologies and the Internet are "better and simpler communication with family members, friends or relatives" (66.9%), "better level of information and access to desired information on various topics" (69.5%) and "greater possibilities for fun and entertainment" (39.4%), while the least said their reasons were linked to their employment. Also, 37.0% agree or agree completely that the Internet costs are "unacceptable", 36.1% are indecisive and only 26.0% disagree. Over 65% examinees agree or strongly agree that the costs of purchasing devices and equipment are too high (Figure 2). Regarding the satisfaction with the internet connection, 41.9% claim the slow Internet connection is an aggravating circumstance for quality of use, 29.5% disagree and 28.7% are indecisive.

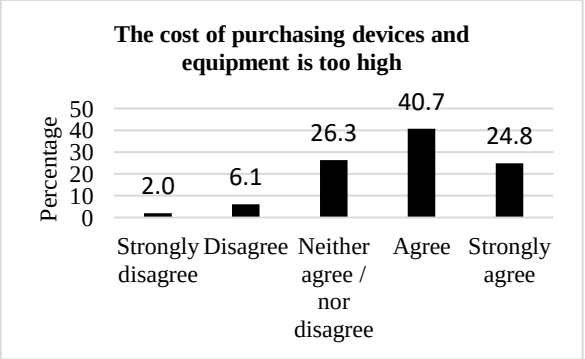


Figure 2. Cost of purchasing devices and equipment

Attitudes toward simplicity of using digital technologies and the Internet are positive. One fifth of participants consider using digital technologies complicated (over 20%), while more than 50% do not.

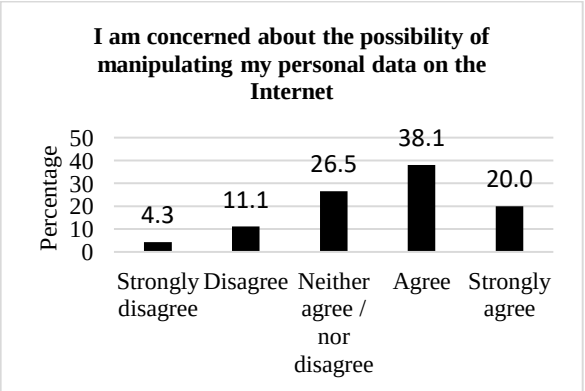


Figure 3. Concerns due to possibility of manipulating respondents' personal data on the Internet

Over 58% of participants are concerned about the possibility of manipulating personal data online, 26.5% are indecisive (Figure 3). Similar percentage of

participants (59.0%) express concern about privacy and confidentiality of personal data online, over 24% are indecisive and 15.0% are not concerned at all. Majority (57%) are concerned about to possible identity thefts online, over 17% are not concerned and 24% are indecisive.

4.4 Knowledge and skills in digital technologies

Over 27% participants agree they do not have enough knowledge and skills to use digital technologies and the Internet, 24.6% are indecisive, just over 48% disagree. Over 70% express interest in digital technologies, only 13.3% are not interested. The majority of participants (62.9%) declared having acquired digital skills on their own at home, 39.5% did it at work, individually or with the help of their colleagues. It is surprising that only 4.4% took organised and paid courses. 42% often use email for the purpose of correspondence, 22% use it sometimes, and 17.2% do not use it, because they do not think it is necessary, while only 18% help others use it. Over 49% use social networks completely on their own or they help others, 29.1% do not use social networks, and somewhat over 20% need help. 66% use or help others use Internet search engines, 13.5% do not use them, and do not consider them necessary, while 19% need help.

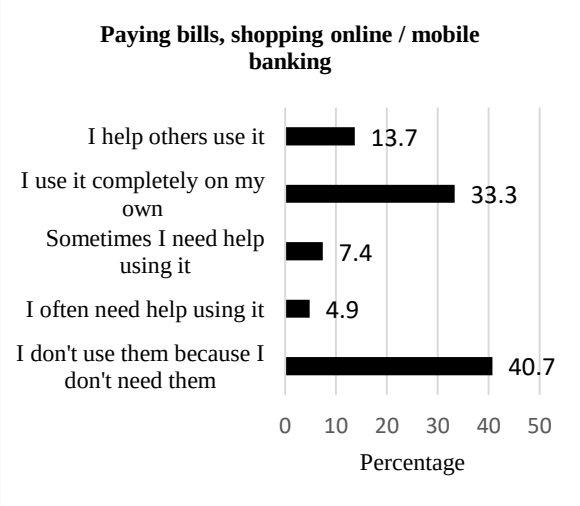


Figure 4. Bill payment, online shopping/mobile banking (Internet banking, web shop)

Figure 4 shows That 45.6% of participants do not use Internet for paying bills, shopping online/mobile banking or they often need help.,, 7.4% said they sometimes need help, and less than 50% stated they perform this activity on their own or help others.

4.5 Digital public services

Over 50% of participants said they do not use or need digital public services, while 45.9% said they use them.

From Figure 5, it is evident that 35.6% of participants do not consider the public service e-Citizen necessary, while 21.3% sometimes and often need help, and 43.2% use them on their own or help others.

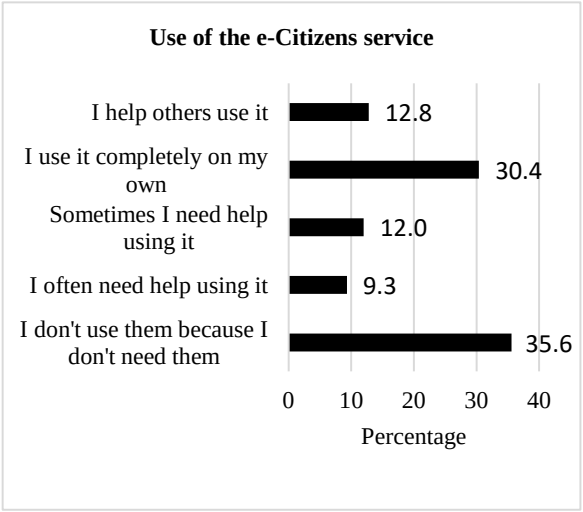


Figure 5. Use of the e-Citizens service

As a reason not to use digital services, the majority said they "do not have enough knowledge and skills", followed by "someone else did it for me" and they "have not got enough information on services that can be used online", while the smallest percentage claimed they "do not know of online public services (e-Citizen)". Among the participants who identified themselves as users of digital public services, 36.7% stated that they used digital public services of downloading official forms (via portal e-Citizen) and 33.7% downloaded certificates and statements from official records (COVID-19 certificate, certificates of no criminal record/residence/whereabouts), 23.9% used them to order health services and acquire information (via email or portal e-citizen), 4.3% used them to borrow digital books. When it comes to the user experience when using digital public services, the results show that more than 43% of respondents declared that the e-Citizens services are simple to use, and more than 50% of them agree with the statements that the e-Citizens services are fully adapted in such a way that all the desired information can be easily accessed and that they have no trouble performing all or most of the required steps during use.

4.6 Media Literacy

Related to the area Media Literacy, 65.2% participants agreed they easily find necessary information online, 20.9% were indecisive, and 13.9% disagreed. More than 47% of participants disagreed with the claim "I

check information online in order to shape my political attitudes.", 30% were indecisive, and 22.4% agreed. More than 20% of participants agreed they check information online in order to shape their life attitudes, 47% disagreed, while 31.4% did not respond. As presented in Figure 6, less than half of examinees (44.4%) check the credibility of information found online, 27.6% did not express their attitude, while the remaining 28% do not check it. Over 61% confirmed that they "critically consider the accuracy of information prior to sharing them in order to avoid spreading false news", 20.2% did not respond, while more than 18% disagreed.

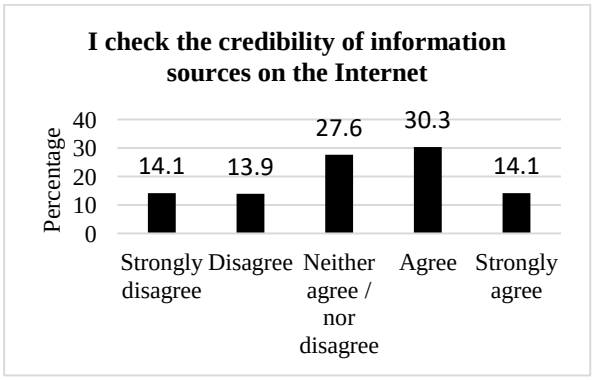


Figure 6. Checking the credibility of information found online

Over 47% agreed they know how to share and comment an online article posted on social networks, 18.3% were indecisive, and 33% disagreed.

Over 32% agreed with the statement "I know how to subscribe to and pay for online services (e.g. music service, e-books, etc.)", 50% disagreed, and 17.8% did not respond. Majority of examinees are aware their online activities are being watched, since more than 63% agreed with the claim "I know my online activities are being watched and analysed for the purpose of adjusting advertising messages.", 18.9% were indecisive, and 17.8% disagreed.

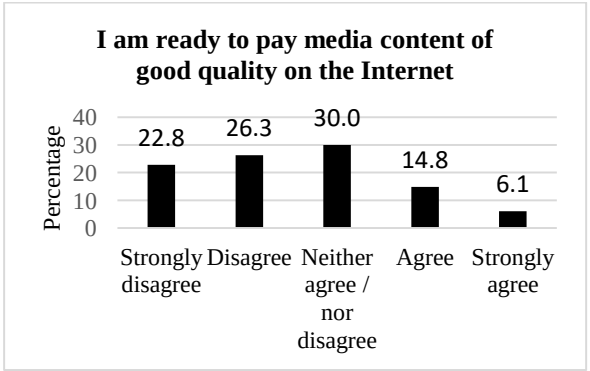


Figure 7. Willingness for payment media content of good quality on the Internet

Only 20.9% of participants are ready to pay for online media content of good quality, 30% did not provide their opinion, while over 49% disagreed. Related to data protection, 60.8% agreed they "know that protection of personal data and Internet privacy are regulated by General Data Protection Regulation in European Union (GDPR)", 23.3% did not express their opinion, while only 15.9% disagreed.

5 Discussion and conclusions

This paper presented methodology for developing an instrument for measuring needs of vulnerable groups in terms of digital inclusion, focusing on access to the Internet, digital competences, and usage of e-services. The proposed methodology ensures content validity of the instrument in relation to the goals defined by the project, and accessibility and understandability of the questionnaire by the target population.

Results of the survey with 540 participants illustrate the use of the instrument. To the best of authors' knowledge, there is no comparable study with similar target group(s) and a questionnaire. Results of the survey provided answers to research questions. They show that vulnerable groups do use digital technologies and the Internet, only 10.7% declared not using the Internet at all. Vulnerable groups most often access the Internet by mobile connection (using a smartphone) or fixed connection at home. Cost of access and devices, slow Internet connection, and concerns regarding privacy and identity theft were the most significant obstacles to using digital technologies and the Internet. More than 50% of participants do not use digital public services at all, or do not think they need them. Related to Media Literacy, 65.2% participants easily find necessary information online. Less than half of participants (44.4%) check the credibility of information found online. Over 61% critically consider the accuracy of information prior to sharing them in order to avoid spreading false news. However, only 20.9% of participants are ready to pay for online media content of good quality.

The presented research results can be of great help to the creators of policy and strategy at the local and the national level, because they provide insight into a broad spectrum of aspects of digital inclusion of vulnerable groups. Recommendations that could be derived from these results are that the first goal of a strategy should be to raise public awareness about the need for digital inclusion of vulnerable groups in the context of the digital transformation of public administration, and the gradual migration of public services to the Internet. Traditional media, which are still the main source of information for senior citizens who do not use the Internet, should play an important role in information campaigns. Digital inclusion strategies and operational plans should envisage the development of digital inclusion centres in which citizens, who do not know or cannot use online

services, can receive assistance, advice, and basic training. The existing network of public libraries could partly ensure the needed infrastructure. By providing additional services for the development of digital and media literacy, libraries could also increase their own sustainability in the 21st century. Citizens' associations that bring together and provide services to senior citizens and people with disabilities should be assisted in building capacity to provide digital literacy services to their members. Digital Competence Framework for Citizens (DigComp) could be used as a reference document for the development of training content. This would simplify the exchange of best practices between Croatian and European digital literacy training providers. Digital inclusion strategies and operational plans should ensure resources for the training of trainers who will be conducting training for vulnerable groups. Intergenerational learning, organized within the civic education of high school students, should be considered as an option to increase digital literacy of elderly. The implementation of the strategies should be evaluated on a regular basis, in order to achieve the EU's digital transformation goals by 2030.

The main limitation of this research is the use of snowball sampling in a limited number of Croatian counties. This limits generalizability of results to the national level. While the choice of counties included in the survey was dictated by the project goals, the snowball sampling, even though a non-probabilistic method, was a method of choice, due to the fact that target vulnerable groups are hard-to reach, and may not be willing to participate in the research, unless contacted by a person of trust. This was particularly relevant, since the survey was conducted during the period of high COVID-19 incidence.

Future research could extend the survey to include other counties in the Republic of Croatia. Further data analyses will include break-down by vulnerable group and county, as well as multivariate analyses. Finally, results will be used to formulate policy recommendations at the local, regional and national levels.

It is important to note that the developed measurement instrument can be used not only in Croatia, but also in other countries at both national and regional levels.

Acknowledgments

This work has been fully supported by European Social Fund (ESF) under the project “Digitalna.hr”, implemented by the Croatian Digital Literacy Network (UP.04.2.1.06).

References

Alabi, A. O., Mutula, S.M. (2020). Digital inclusion for visually impaired students through assistive technologies in academic libraries. *Library Hi Tech News.* 37(2), 14–17.

Betts, L. R., Hill, R., & Gardner, S. E. (2019). “There’s not enough knowledge out there”: Examining older adults’ perceptions of digital technology use and digital inclusion classes. *Journal of Applied Gerontology*. 38(8), 1147-1166.

Chetty, K., Aneja, U., Mishra, V., Gcora, N., & Josie, J. (2018). Bridging the digital divide in the G20: skills for the new age. *Economics Discussion Papers*, No 2017-68, Kiel Institute for the World Economy. Retrieved from http://www.economics-ejournal.org/economics/discussionpapers/2017-68

Ellard-Gray, A., Jeffrey, N. K., Choubak, M., & Crann, S. E. (2015). Finding the hidden participant: Solutions for recruiting hidden, hard-to-reach, and vulnerable populations. *International Journal of Qualitative Methods*, *14*(5), 1609406915621420.

European Commission (2006). Ministerial Declaration, 11. June 2006, Riga, Latvia. Retrieved from http://ec.europa.eu/information_society/events/ict_riga_2006/doc/declaration_riga.pdf.

European Commission (2021) 2030 Digital Compass: the European way for the Digital Decade. COM(2021) 118 final. Retrieved from https://ec.europa.eu/info/sites/default/files/communication-digital-compass-2030_en.pdf

Helsper, E. (2008). *Digital inclusion: an analysis of social disadvantage and the information society*. Department for Communities and Local Government.

Helsper, E. J., & Galácz, A. (2009). Understanding the links between social and digital exclusion in Europe. *World wide internet: Changing societies, economies and cultures.* (pp.146-178).

Helsper, E. J. (2012). A corresponding fields model for the links between social and digital exclusion. *Communication theory*, 22(4), 403-426.

International Telecommunication Union (ITU). Measuring the Information Society Report: Volume 1. Geneva, Switzerland: ITU. (2018a) Retrieved from https://www.itu.int/en/ITUD/Statistics/Documents/publications/misr2018/MISR-2018-Vol-1-E.pdf.

International Telecommunication Union (ITU). Digital Skills Toolkit. Geneva, Switzerland: ITU. (2018b) Retrieved from https://www.itu.int/en/ITU-D/Digital-

Inclusion/Documents/ITU%20Digital%20Skills%20Toolkit.pdf

Krippendorff, K. (1980). Content analysis: An introduction to its methodology. Beverly Hills, CA: Sage.

Lyons, A., Zucchetti, A., Kass-Hanna, J., & Cobo, C. Bridging the gap between digital skills and employability for vulnerable populations. (2019). Retrieved from https://t20japan.org/wp-content/uploads/2019/03/t20-japan-tf7-9-bridging-gap-between-digital-skills-employability.pdf.

Lyons, A. C., & Kass-Hanna, J. A human development approach to measuring and improving the digital livelihoods of vulnerable populations. T20 Policy Brief. (2020). Retrieved from https://www.g20-insights.org/wp-content/uploads/2020/11/a-human-development-approach-to-measuring-and-improving-the-digital-livelihoods-of-vulnerable-populations-2-1606240348.pdf.

Lyons, A., Kass-Hanna, J., Greenlee, A. Impacts of Financial and Digital Inclusion on Poverty in South Asia and Sub-Saharan Africa. (2020). Retrieved from https://ssrn.com/abstract=3684265 or http://dx.doi.org/10.2139/ssrn.3684265Available at SSRN 3684265.

Mechanic, D., Tanner, J. (2007). Vulnerable people, groups, and populations: societal view. *Health Affairs*, 26(5), 1220-1230.

Mancinelli, E. e-Inclusion in the Information Society, Information Society from Theory to Political Practice. (2008). Retrieved from http://www.ittk.hu/netis/doc/NETIS_Course_Book_English.pdf.

Matthews, K., Nazroo, J., & Marshall, A. (2019). Digital inclusion in later life: cohort changes in internet use over a ten-year period in England. *Ageing & society*, 39(9), 1914-1932.

Milas, G. (2005). *Istraživačke metode u psihologiji i drugim društvenim znanostima.* Jastrebarsko/Hrvatska: Naklada Slap.

Palmeiro, R., Pereda Herrero, M.V., Aires, L. (2019). Digital inclusion programs: the case of the Basque Country. *Revista Lusófona de Educação*, 45, 73–88.

Pérez-Escolar, M., & Canet, F. (2022). Research on vulnerable people and digital inclusion: toward a consolidated taxonomical framework. *Universal Access in the Information Society*, 1-14.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Real, B., McDermott, A.J., Bertot, J.C., Jaeger, P.T. (2015). Digital inclusion and the affordable care act: public libraries, politics, policy, and enrollment in "Obamacare." *Public Library Quarterly*, 34(1), 1–22.

Serrano-Santoyo, A., Rojas-Mendizabal, V. (2017). Exploring a complexity framework for digital inclusion interventions. *CENTERIS – International Conference on ENTERprise Information Systems (ed Procedia computer science, 121)*, Barcelona, Spain, 8–10 November 2017, (pp. 212–217).

Tkalac Verčič, A. Sinčić Ćorić, D. Pološki Vokić, N. (2011). *Priručnik za metodologiju istraživanja u društvenim djelatnostima*, Zagreb/Hrvatska: M.E.P. d.o.o.

Tomczyk, Ł., Eliseo, M. A., Costas, V., Sánchez, G., Silveira, I. F., Barros, M. J., ... & Oyelere, S. S. (2019). Digital Divide in Latin America and Europe: Main characteristics in selected countries. *In 2019 14th Iberian Conference on Information Systems and Technologies (CISTI)* (pp. 1-6). IEEE.

Tsatsou, P., Higgs, G., & Stafford, I. (2013). Digital inclusion and connectivity of minority communities: the case of communities of disabled people. *Management of Broadband Technology and Innovation*, 146-166.

Tsatsou, P. (2020). Digital inclusion of people with disabilities: a qualitative study of intra-disability diversity in the digital realm. *Behaviour & Information Technology*, 39(9), 995-1010.

Tsatsou, P. (2021a). Vulnerable people's digital inclusion: intersectionality patterns and associated lessons. *Information, Communication & Society*, 1-20.

Tsatsou, P. (2021b). Is digital inclusion fighting disability stigma? Opportunities, barriers, and recommendations. Disability & Society, 36(5), 702-729.

Van Dijk, J. (2020). *The digital divide.* John Wiley & Sons.

Vuorikari, R., Kluzer, S. and Punie, Y. (2022) DigComp 2.2 The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes, EUR 31006 EN, Publications Office of the European Union, Luxembourg, doi:10.2760/490274, JRC128415

Yu, H., Goggin, G., Fisher, K., Li, B. (2019). Introduction: disability participation in the digital economy. *Information, Communication & Society*. 22(4), 467-473.

Žajdela Hrustek, N. Multidimensional and multiperspective approach for monitoring e-Inclusion. University of Zagreb, Faculty of

Organization and Informatics. (2015). Retrieved from https://urn.nsk.hr/urn:nbn:hr:211:404534

Žajdela Hrustek, N., Prosser, A., Dušak, V. (2016). Multidimensional Model of E-inclusion and its Implications. *Proceedings of the Central and Eastern European eDem and eGov Days 2016*. In A. Balthasar, B. Golob, H. Hansen, R. Muller Torok, A. Nemeslak. J. Pichler, A.Prosser (Eds.), Austrian Computer Society, (p.p. 435-447). Budapest, Hungary.

Žajdela Hrustek, N., Šimić, D., Vrček, N. (2017). Electronic services with impact on quality of life - a survey of usage and attitudes of north-western Croatian. *Proceedings of the Central and Eastern European eDem and eGov Days 2017*. In H. Hansen, R. Müller-Török, A. Nemeslaki, J. Pichler, A. Prosser, D. Scola (Eds.), Computer Gesellschaft, (p.p. 29-41). Wien: Osterreichische.

Regulation of the legal protection of individual in digital environment within the Croatian educational system

Regulacija zaštite pojedinca u digitalnom okruženju unutar hrvatskog odgojno – obrazovnog sustava

Azra Staščik
Primary School of Dobriša Cesarić Osijek
Osnovna škola "Dobriša Cesarić" Osijek
Neretvanska ulica 10, Osijek
{azra.stascik@skole.hr}

Amanda Glavaš
Faculty of Humanities and Social Sciences
Josip Juraj Strossmayer University of Osijek
Filozofski fakultet Osijek
Sveučilište Josipa Jurja Strossmayera u Osijeku
Ul. Lorenza Jägera 9, Osijek
{aglavas@ffos.hr}

Filip Glavaš
Filip Glavaš, Attorney at Law
Odvjetnik Filip Glavaš
Lučki prilaz 2, Osijek
{filip@odvjetnik-glavas.hr}

Abstract. *The COVID-19 pandemic put the educational system in front of a big challenge, which required an immediate transition to distance learning and digitalization of teaching. Teaching and learning were transferred to virtual classrooms, which required a complete adjustment of the educational system. The aim of research is to examine the legal regulation of individual protection in the digital (educational) environment, at macro and micro level. The examination at the macro level implies a qualitative analysis of the relevant legislation. At the micro level on a representative sample of primary and secondary schools (N=303) documentation analysis was conducted to examined whether schools have regulated this matter internally. The results indicate insufficient legal and internal (only 5%) regulation of safe and responsible use of ICT.*

Keywords. Digitally advanced schools, information and communication technologies (ICT), legal protection of an individual, educational system

Sažetak. *COVID – 19 pandemija stavila je odgojno obrazovni sustav pred veliki izazov, koji je zahtijevao trenutni prijelaz na nastavu na daljinu te digitalizaciju nastave. Nastava je premještena u virtualne učionice što je zahtijevalo cjelovitu prilagodbu odgojno obrazovnog sustava. Cilj je istraživanja ispitati pravnu regulaciju zaštite pojedinca u digitalnom (odgojno-obrazovnom) okruženju, na makro i mikro razini. Ispitivanje na makro razini podrazumijeva kvalitativnu analizu odgovarajućih regulativa. Na mikro razini se na reprezentativnom uzorku osnovnih i srednjih škola (N=303) provela analiza dokumentacije kojom se ispitalo jesu li i u kojoj mjeri škole interno regulirale navedeno. Rezultati upućuju na nedovoljno zakonsko i interno (samo 5%) reguliranje sigurne i odgovorne upotrebe IKT.*

Ključne riječi. Digitalno zrele škole, informacijsko-komunikacijske tehnologije, pravna zaštita pojedinca, odgojno-obrazovni sustav

1 Introduction

Progressive development of technology affects changes in the modern world in every aspect, from

1 Uvod

Progresivni razvoj tehnologije utječe na promjene u suvremenom svijetu u svim njegovim aspektima, od industrijsko-proizvodnih, poslovno-

industrial-production, business-communication, educational, etc. In other words, it can be stated that digital technology has entered all parts of everyday life (NN 22/2018). Information and communication technologies (ICT) provide new ways of communicating, transferring, searching, storing, and exchanging information, and the increased use of ICT is very present in the education system as well, especially school education. The digital competence has been declared by the European Parliament and the Council of the European Union (2006) as one of the eight key competences for lifelong learning, which was accepted by Croatian education policy with emerging of the National Curriculum Framework (NOK) (Fuchs et al., 2011). Additionally, digital competence is the ability to utilize technology effectively, appropriately, securely, critically, creatively, autonomously, and ethically in order to execute activities, solve problems, communicate, manage information, collaborate, and create and distribute material. (Skov, 2016).

Given the scope of the concept of digital competence, various terms and definitions can be found in the literature, which acknowledge existence of various types of literacy in relation to digital competences. In the NOK (Fuchs et al., 2011, p. 17) digital competence is defined as "the ability to safely and critically use information and communication technology for work in personal and social life as well as in communication." For the context of this paper, it is crucial to identify determinants of digital competence: effective and responsible communication and cooperation in the digital environment, understanding and responsible application of security recommendations to protect the individual and taking into consideration legal guidelines when using digital technology (NN 22/2018).

One of the preconditions for the development of digital competence in students is the evolution of schools into digitally developed schools. Digitally developed schools are those schools that have a high level of ICT integration, as well as a systematic approach to the use of ICT in school management, but also within other educational processes (Begičević Ređep et al., 2018).

The Strategy of Education, Science and Technology (NN 124/2014) sets development of informatization of the educational process and the school management process as one of the main aims, i.e., the development of digitally advanced schools (e-schools). The importance of developing digitally advanced schools was particularly highlighted by the emergence of the COVID-19 pandemic when teaching and learning was abruptly moved to virtual classrooms via digital devices. Coincidently, the number of users in digital learning environment rapidly increased, which required quick adaptation of the educational system throughout all (different)

komunikacijskih, odgojno-obrazovnih itd. Drugim riječima, može se reći da je digitalna tehnologija ušla u sve pore svakodnevnog života (NN, 22/2018). Informacijsko-komunikacijske tehnologije (dalje: IKT) donose nove načine komuniciranja, prenošenja, traženja, pohrane i razmjene informacija, a povećana uporaba IKT-a prisutna je i sustavu odgoja i obrazovanja, posebice školskog. Digitalnu kompetenciju Europski parlament i Vijeće (EC 2006/962) proglasilo je za jednu od osam ključnih kompetencija za cjeloživotno obrazovanje, a što je hrvatska obrazovna politika i prihvatila Nacionalnim okvirnim kurikulumom (dalje: NOK) (Fuchs i sur., 2011). Osim toga digitalna kompetencija može se definirati i kao kombinacija znanja, vještina i stavova u vidu korištenja tehnologije sa svrhom obavljanja nekog zadatka, rješavanja problema, komunikacije, upravljanja informacijama, suradnje, kao i kreiranja i dijeljenja digitalnog sadržaja na učinkovit, siguran, kreativan, primjeren, kritički, neovisan i etički način (Skov, 2016).

Obzirom na širinu koncepta digitalne kompetencije u literaturi se mogu pronaći različiti pojmovi te definicije koje uključuju i postojanje različitih vrsta pismenosti vezanih uz digitalne kompetencije. U NOK-u (Fuchs i sur., 2011, str. 17) digitalna kompetencija je definirana kao „osposobljenost za sigurnu i kritičku upotrebu informacijsko-komunikacijske tehnologije za rad u osobnomu i društvenomu životu te u komunikaciji". U kontekstu ovog rada vrlo bitne su odrednice digitalne kompetentnosti: učinkovito i odgovorno komuniciranje i surađivanje u digitalnome okruženju, razumijevanje i odgovorno primjenjivanje sigurnosnih preporuka s ciljem zaštite pojedinca te poštivanje pravnih odrednica pri korištenju digitalne tehnologije (NN 22/2018).

Jedna od pretpostavki razvijanja digitalne kompetencije kod učenika su razvijene digitalno zrele škole. Digitalno zrelim školama smatraju se one škole koje imaju visoku razinu integracije IKT-a, kao i ususstavljen pristup korištenja IKT-a pri upravljanju školom, ali i unutar drugih odgojno-obrazovnih procesa (Begičević Ređep i sur., 2018). Strategijom obrazovanja, znanosti i tehnologije (NN 124/2014) kao jedan od ciljeva postavlja se razvijanje informatizacije odgojno-obrazovnog procesa i procesa upravljanja školom, tj. razvijanje digitalno zrelih škola (e-škola). Važnost razvijanja digitalno zrelih škola posebno se istaknula pojavom COVID-19 pandemije kada se poučavanje i učenje odvijalo putem digitalnih uređaja, a nastava je naglo premještena u virtualne učionice. Tada dolazi do ekspanzije broja korisnika u školskom digitalnom okruženju, što je zahtijevalo brzu prilagodbu odgojno-obrazovnog sustava u svim (različitim) aspektima. Ako već ranije nije bilo regulirano, bilo je potrebno regulirati pitanja odgovornog i prihvatljivog načina komuniciranja i suradnje između učenika, učenika i nastavnika, nastavnika, a

aspects. If it had not been regulated before, now, it was necessary to regulate the matter of responsible and acceptable manner of communication and collaboration among students, as well as among teachers. It was necessary to provide regulations in regard of mutual respect and appreciation and the use of netiquette as protection against abuse. Moreover, it was necessary to choose a secure digital platform, i.e., a secure digital environment in terms of user data protection and online security for virtual teaching and learning to take place. Furthermore, it was necessary to regulate the protection of copyright for the works of teachers (as well as students) from unauthorized publication, sharing of data and content, with respect of sources as well as protection of privacy rights of students and teachers. In other words, it was necessary to create conditions for the protection of the individual in the digital learning environment, i.e., to define acceptable and unacceptable behaviors and to prescribe measures in case of breaching the regulations. Due to the latter, it is necessary to ensure that the acts regulating the protection of the individual in the digital environment are coordinated with the relevant and applicable legislation.

2 Related work

2.1 Croatian legal framework

In order for an individual to be able to protect himself in a digital learning environment, one must have a legal foundation for it. The Constitution of the Republic of Croatia (NN 5/2014), as the legal act with the greatest legal force, defines the highest values of the constitutional order, including respect for human rights (Art. 3), which in the context of this paper would be: the right to personal security, the right to privacy; the right to protection of moral and material interests arising from cultural and scientific creation. Additionally, copyright is regulated by the Law on Copyright and Related Rights (NN 111/2021). Privacy and its protection are regulated by the General Data Protection Regulation (EU 2016/679) and the Law on the Implementation of the General Data Protection Regulation (NN 42/2018). In addition, the issue of online security is based on the Law on Information Security (NN 79/2007). Issues of safety and protection of children within the school environment are specifically regulated by the Law on Education in Primary and Secondary Schools (NN 64/2020). Finally, the protection and well-being of children (even in the field of upbringing and education) is the duty of parents, whose obligations are regulated by the Family Law (NN 98/2019).

s posebnim naglaskom na međusobno poštovanje i uvažavanje te primjenu netiquettea, kao zaštite od zlostavljanja pojedinca. Osim toga, bilo je potrebno odabrati sigurnu digitalnu platformu, odnosno sigurno digitalno okruženje putem kojeg će se odvijati virtualna nastava, u pogledu zaštite podataka korisnika i sigurnosti na mreži. Nadalje, bilo je potrebno regulirati zaštitu autorskih prava radova nastavnika (ali i učenika) od neovlaštenog objavljivanja, dijeljenja podataka i sadržaja, uz poštivanje izvora te zaštitu prava privatnosti učenika i nastavnika. Drugim riječima, trebalo je ostvariti uvjete za zaštitu pojedinca u digitalnom školskom okruženju, tj. definirati prihvatljive i neprihvatljive načine ponašanja te propisati mjere u slučaju nepridržavanja. Zbog potonjeg, potrebno je voditi računa o tome da akti kojima se regulira(la) zaštita pojedinca u digitalnom okruženju budu usklađeni sa odgovarajućim važećim zakonskim regulativama.

2 Teorijska razmatranja

2.1 Hrvatski pravni okvir

Da bi se pojedinac mogao zaštititi u (školskom) digitalnom okruženju mora za to imati zakonsku podlogu. Ustavom Republike Hrvatske (NN 5/2014) zakonskim aktom s najvećom pravnom snagom definirane su najviše vrednote ustavnog poretka unutar kojih je i poštivanje prava čovjeka (čl. 3.), a koja bi u kontekstu ovog rada bila: pravo na osobnu sigurnost, pravo na zaštitu privatnosti, pravo na zaštitu moralnih i materijalnih interesa koji proizlaze iz kulturnog i znanstvenog stvaralaštva. Dodatno, autorska prava regulirana su Zakonom o autorskom pravu i srodnim pravima (NN 111/2021). Privatnost i njezina zaštita uređeni su Općom uredbom o zaštiti osobnih podataka (EU 2016/679) i Zakonom o provedbi Opće uredbe o zaštiti podataka (NN 42/2018). Osim toga, pitanje sigurnosti na mreži podlogu nalazi u Zakonu o informacijskoj sigurnosti (NN 79/2007). Pitanja sigurnosti i zaštite djece unutar školskog okruženja, posebno su regulirana još Zakonom o odgoju i obrazovanju u osnovnoj i srednjoj školi (NN 64/2020). Konačno, zaštita i dobrobit djece (pa i u domeni odgoja i obrazovanja) dužnost je roditelja, a čije obveze su regulirane Obiteljskim zakonom (NN 98/2019).

2.2 Školsko digitalno okruženje

Tradicionalna nastava podrazumijeva nastavnika koji poučava, učenike koji uče, individualno ili unutar grupe, obično koristeći dobro poznata nastavna pomagala: papir, olovku, kredu i ploču. Tradicionalno poučavanje i učenje izdržalo je test vremena i bez obzira na promjene paradigme učenja i poučavanja i dalje je najrasprostranjenije diljem svijeta (Dash i sur., 2021). Uvođenjem tehnologije u

2.2 Digital learning environment

Traditionally, learning environment presumes teacher teaching, students learning, individually or within group, usually using well known tools, paper, pen, chalk and a blackboard. Traditional teaching and learning stood the test of time and is worldwide accepted. (Dash et al., 2021) By introducing technology into education traditional learning environment has transformed. Emerging of technology did not change learning environment just in regard of infrastructure but it produced a new learning environment, beyond physical dimension (Barana & Marchiso, 2021). The learning environment reachable everywhere via Internet using tools such as computers, tablets, or even mobile devices (Yong, 2007, Barana & Marchiso, 2021). Except mere environment and learner, digital devices are necessary to reach learning process, therefore "Digital Learning Environment" (DLE) is name used to describe such learning environment. (Barana & Marchiso, 2021).

There are many definitions of DLE, but the common denominator is of course use of Internet and "Learning Management System" (LMS). In this paper we chose to use term "Digital Learning Environment" as well as Barana and Marchiso definition (2021, p. 498) by which DLE is "a learning ecosystem in which teaching, learning, and the development of competence are fostered in classroom-based, online, or blended settings". Classroom based is a face-to-face modality where students work individually or with peers using digital tools and devices. Online modality presumes the DLE as the only learning environment in online courses (e.g., during COVID – 19 pandemic or professional development courses). Blended modality includes students' learning activities in classroom or computer lab as well as online integration e.g., teacher assigning homework online. The third modality is one frequently used and stayed in practice even after the pandemic (Barana & Marchiso, 2021).

2.3 Analysis of used digital learning environments during the COVID-19 pandemic

Croatian education policy defines the development of digitally mature schools as one of the goals of education development (Strategija obrazovanja, znanosti i tehnologije, 2014). The ownership of digital competences and skills, ensuring the availability of digital infrastructure and equipment proved to be very important with the appearance of the COVID-19 pandemic. The European Commission paid great attention to the new circumstances and challenges of education, creating

obrazovanje transformira se tradicionalna nastava. Pojava tehnologije nije promijenila nastavu i proces poučavanja samo u pogledu infrastrukture, već je proizvela novo okruženje za učenje, izvan fizičke dimenzije (Barana & Marchiso, 2021). Nastava postaje dostupna svuda „i svakome", putem interneta i pomoću uređaja kao što su računala, tableti ili čak mobilni uređaji (Yong, 2007; Barana & Marchiso, 2021). U tom kontekstu, osim okruženja za učenje i učenika, digitalni uređaji su nužni za ostvarenje procesa učenja, stoga je školsko digitalno okruženje (DLE) naziv koji se koristi za opisivanje takvog okruženja za učenje (Barana & Marchiso, 2021).

Postoje mnoge definicije školskog digitalnog okruženja, ali zajednički nazivnik je upravo korištenje interneta i sustava upravljanja učenjem (LMS). U ovom radu koristit će se definicija školskog digitalnog okruženja prema Barana i Marchiso (2021, str. 498) prema kojoj je ono "ekosustav učenja u kojem se potiče poučavanje, učenje i razvoj kompetencija u učionici, online ili u kombiniranom okruženju". DLE u učionici može se koristiti u tri oblika. Način korištenja *licem u lice* je oblik pri kojem učenici rade pojedinačno ili sa svojim vršnjacima koristeći digitalne alate i uređaje. Online način pretpostavlja školsko digitalno okruženje kao jedino okruženje za učenje u online nastavi (npr. tijekom pandemije COVID-19 ili tečajeva za profesionalni razvoj). Školsko digitalno okruženje u kombiniranom načinu korištenja uključuje aktivnosti učenja učenika u učionici ili računalnim učionicama, ali i integraciju online aktivnosti, npr. učitelj zadaje zadaću koja će se riješiti online. Posljednje navedena opcija često se koristi, a njezina uporaba ostala je u praksi i nakon završetka pandemije (Barana & Marchiso, 2021).

2.3 Analiza korištenih školskih digitalnih okruženja tijekom pandemije COVID-19

Hrvatska obrazovna politika definira razvoj digitalno zrelih škola kao jedan od ciljeva razvoja obrazovanja (Strategija obrazovanja, znanosti i tehnologije, 2014), a posjedovanje digitalnih kompetencija i vještina te osiguravanje dostupnosti digitalne infrastrukture i opreme, vrlo važnim se pokazalo pojavom COVID-19 pandemije. Novonastalim okolnostima i izazovima obrazovanja, veliku pozornost dala je Europska Komisija, izrađujući akcijski plan za digitalno obrazovanje u razdoblju od 2021 do 2027 godine. (European Education Area, 2021).

Školska digitalna okruženja u Hrvatskoj postaju sve zrelija pojavom COVID – 19 pandemije te prelaskom cjelokupnog školskog sustava na nastavu na daljinu u ožujku 2020. godine. U to vrijeme odgojno-obrazovne ustanove nisu dobile jedinstvenu i standardnu uputu oko izbora alata za održavanje online nastave, posljedično, digitalna rješenja koja su

an action plan for digital education in the period from 2021 to 2027 (European Education Area, 2021).

Digital learning environments in Croatia are becoming more and more advanced with the emergence of the COVID-19 pandemic and the transition of the entire educational system to distance learning in March 2020. At that time, educational institutions. i.e., schools did not receive universal or standardized instruction on the choice of tools for online teaching, consequently, the digital solutions they used differed greatly in terms of regulating the protection of the individual in the digital environment. Although standardized guidelines have not been regulated or provided at the national level (or european), there is a legal framework and foundation for the protection of individuals in the digital environment and it is to be expected that educational institutions would be guided by them when choosing digital solutions. Besides that, the European Commission made efforts to support the transition of educational systems from physical to the digital learning environment, which are visible, by creating various educational initiatives and policies that could potentially guide them in decision-making. (European Educational Area, n.d.)

By including the principle of *"privacy by design"* in the General Data Protection Regulation (EU 2016/679), it becomes a legal obligation of all creators of digital and technical solutions both at the level of the European Union and in the Republic of Croatia. However, the analysis of digital solutions proposed for use in educational purposes, during the COVID-19 pandemic, revealed that some of these solutions did not fully comply with this principle aforementioned (Politiscope, 2021b). At that time, the Ministry of Science and Education issued a number of guidelines, including a document entitled Action Plan for the Implementation of Distance Learning (MZO, 2020). This document provides recommendations for use of technical solutions such as Moodle, Teams, Yammer, Google Classroom and Edmondo. At the same time CARNET publishes a document entitled Online systems for organizing and implementing distance learning (CARNET, 2020) and proposes technical solutions such as Google Classroom and Google Meet, Microsoft Teams, Yammer, CARNET Loomen and Zoom. In addition to recommendations for choosing a technical solution, CARNET in its document lists the advantages and disadvantages of each of its recommendations, and based on this, schools were supposed to make their own judgment on the appropriateness of using the solutions recommended (CARNET, 2020).

The Politscope analysis (2021a) indicates proposed solutions does not fully protect the privacy of its users, most often teachers and students. By analyzing the most commonly used platforms and applications for distance learning, Cohney et al. (2021) in their research draw attention to the koristile uvelike se razlikuju po pitanju regulacije zaštite pojedinca u digitalnom okruženju. Iako standardizirane upute nisu donesene na nacionalnoj razini (pa ni na europskoj), pravni okvir za zaštitu pojedinaca u digitalnom okruženju postoji te je za očekivati da će se odgojno – obrazovne ustanove istima voditi pri odabiru digitalnih rješenja. Osim toga, vidljiva su nastojanja Europske Komisije za davanje potpore tranziciji ,odgojno obrazovnim sustavima, iz fizičkog u školsko digitalno okruženje, kreiranjem raznih odgojno obrazovnih inicijativa i politika koje su ih potencijalno mogle usmjeriti u donošenju odluka. (European Educational Area, n.d.)

Kada je riječ o nastanku novih digitalnih alata i aplikacija *"privacy by design"* načelo, pravna je obveza svih kreatora digitalnih rješenja u Republici Hrvatskoj koju nalaže Opća uredba o zaštiti osobnih podatka (EU 2016/679). Ipak, analizom digitalnih rješenja predloženih za korištenje u odgojno – obrazovne svrhe tijekom COVID-19 pandemije utvrđeno je kako neka od tih rješenja nisu u potpunosti ispoštovala navedeno načelo (Politiscope, 2021b). U to vrijeme Ministarstvo znanosti i obrazovanja izdaje niz smjernica, između ostalog i dokument s nazivom Akcijski plan za provedbu nastave na daljinu (MZO, 2020). Tim dokumentom preporučuje se korištenje tehničkih rješenja poput Moodle, Teams, Yammer, Google Classroom i Edmondo. Ujedno CARNET objavljuje dokument pod nazivom Online sustavi za organizaciju i provođenje nastave na daljinu (CARNET, 2020) te predlaže tehnička rješenja poput Google Classroom i Google Meet, Microsoft Teams, Yammer, CARNET Loomen i Zoom. Osim preporuka za odabir tehničkog rješenja, CARNET u svom dokumentu navodi prednosti i mane za svaku od svojih preporuka te su na temelju toga škole mogle donijeti vlastiti sud primjerenosti upotrebe predloženih rješenja (CARNET, 2020).

Analiza Politscope-a (2021a) ukazuje na to da korištenje predloženih rješenja ne štiti u potpunosti privatnost svojih korisnika, na**j**češće nastavnika i učenika. Analizom najčešće korištenih platformi i aplikacija Cohney i suradnici (2021) upozoravaju na specifičnosti digitalnih rješenja koja zaobilaze zabrane uređaja na kojima se nalaze kako bi prikupljale podatke o lokaciji korisnika. Proučavanjem politika privatnosti digitalnih rješenja (Cohney i sur., 2021; Politiscope, 2021a) također su uočeni brojni primjeri razmjene podataka o korisnicima s trećim stranama u marketinške svrhe. Uočeno je i kako Google Classroom te MS Teams prikupljaju velike količine osobnih podataka s kojima se potencijalno stvara profil korisnika, a razlozi prikupljanja tih podataka nisu jasno definirani. Također, MS Teams nudi samo enkripciju u prijenosu, ali ne nudi end-to-end enkripciju (E2EE). Naglim porastom broja korisnika Zoom aplikacije, eksperti za zaštitu privatnosti korisnika upućuju na mnogobrojne probleme koje se nalaze u

specifics of most used digital solutions that bypass the prohibitions installed on the devices in order to collect user location data. Research on the privacy policies of digital solutions (Cohney et al., 2021; Politiscope, 2021a) identified numerous cases of the exchange of customer data with third parties for marketing purposes. It was also noticed that Google Classroom and MS Teams collect large amounts of personal data of their users that can potentially be used to create a user profile, and the reasons for collecting this kind and amount of data are not clearly specified. Also, MS Teams only offers encryption while transferring data among users, but does not offer end-to-end encryption (E2EE). With the large increase of the number of Zoom users, user privacy experts pointed out many problems found in the default settings of Zoom regarding user privacy (Dash et al., 2021). In the foreseeable future, Zoom has greatly improved its privacy shortcomings. Looking into Loomen privacy and authentication setting it is evident that Loomen also collects large amounts of personal data, however, the risk of sharing this information with third parties is considered much lower, as Loomen is an official, national technical solution (Politiscope, 2021a).

Similar challenges in protecting students and teachers' privacy in the digital learning environment can be seen in analyzes of practices around the world, including Nepal and the United States of America (Lamsal, 2021; Cohney et al., 2021). Aforementioned research contains many tips and guidelines for successful implementation of distance learning and safe digital learning environments. One of the key recommendations is the use of open – code tools that would allow installation of this tools onto schools' own servers. In addition, of great importance is considered educating teachers on protection of their work as well as on the protection of privacy during online teaching. Transferring this knowledge to students would mean an original acquaintance with their own digital rights, raising the level of digital competencies and realistic and quality preparation for life in a modern and digital environment (Cohney et al., 2021).

3 Methodology

3.1 The aim and research questions

The aim of this research is to examine the legal order, i.e., regulation of the legal protection of an individual in the digital environment, with emphasis on the educational dimension at the macro (national) and micro (local) level. The examination at the macro level implies a qualitative analysis of primary data sources (legislation) of the Republic of Croatia relevant to this problem. At the micro level it examines whether and to what extent did primary and secondary schools regulate this matter in zadanim postavkama (Dash i sur., 2021). U dogledno vrijeme Zoom je uvelike unaprijedio nedostatke u pogledu zaštite privatnosti, a iako Loomen također prikuplja velike količine osobnih podataka, smatra se da je rizik od dijeljenja tih informacija s trećom stranom puno manji, s obzirom da je Loomen službeno, nacionalno tehničko rješenje (Politiscope, 2021a).

Slične izazove u zaštiti učenika i nastavnika u digitalnom školskom okruženju možemo vidjeti u analizama praksi diljem svijeta pa tako i u Nepalu i SAD – u (Lamsal, 2021; Cohney i sur., 2021). U navedenim istraživanjima sadržani su i mnogi savjeti i smjernice za budućnost, a jedna od ključnih je korištenje alata s otvorenim kodom koji omogućuju da se ti alati postave na vlastite servere škola. Osim toga, ključnim se također smatra, edukacija nastavnika o zaštiti svog autorskog djela, ali i o zaštiti privatnosti tijekom online nastave. Prijenos tih znanja na učenike značilo bi izvorno upoznavanje s vlastitim digitalnim pravima, podizanje razina digitalnih kompetencija te realnu i kvalitetnu pripremu za život u suvremenom i digitalnom okruženju (Cohney i sur., 2021).

3 Metodologija istraživanja

3.1 Cilj i istraživačka pitanja

Cilj je ovog istraživanja ispitati pravnu uređenost, tj. regulaciju zaštite pojedinca u digitalnom okruženju, s naglaskom na odgojno-obrazovnu dimenziju na makro (nacionalnoj) i mikro (lokalnoj) razini. Ispitivanje na makro razini podrazumijeva kvalitativnu analizu primarnih izvora podataka (zakonskih regulativa) u Republici Hrvatskoj relevantnih za navedeni problem, a na mikro razini se ispituje jesu li i u kojoj mjeri osnovne i srednje škole samostalno, internim posebnim aktima, regulirale navedeno.

U skladu s time formulirana su sljedeća istraživačka pitanja:

Q1: Je li regulirana zaštita pojedinca u digitalnom školskom okruženju zakonskim aktima i podaktima?

Q2: Jesu li osnovne i srednje škole samostalno, internim aktima regulirale problem zaštite pojedinca u školskom digitalnom okruženju?

Q2.1. Koji su aspekti zaštite pojedinca u digitalnom školskom okruženju obuhvaćeni pravilnicima?

3.2 Uzorak i postupak

Kako bi se odgovorilo na prvo istraživačko pitanje, analizirali su se sljedeći relevantni zakonski akti:

- Ustav Republike Hrvatske (NN 5/2014)
- Zakon o odgoju i u obrazovanju u osnovnoj i srednjoj školi (NN 64/2020)

independent matter by providing users with their internal regulations specified for use of ICT.

Accordingly, the following research questions were formulated:

Q1: Is the protection of the individual in the digital learning environment regulated by legal acts and sub-acts?

Q2: Did primary and secondary schools independently, by internal regulations, regulate the problem of individual protection in the digital learning environment?

Q2.1. What aspects of individual protection in digital learning environment are covered by regulations?

3.2 Sample and procedure

In order to answer the first research question, following, relevant, legal acts were analyzed:

- Constitution of the Republic of Croatia (NN 5/2014)
- Primary and secondary education act (NN 64/2020)
- National pedagogical standard of the primary school education system (NN 63/2008)
- National pedagogical standard of the secondary education system (NN 63/2008)
- Policy on criteria for imposing pedagogical measures (NN 3/2017)
- General Data Protection Regulation (EU 2016/679)
- Law on the Implementation of the General Regulation on Data Protection (NN 42/2018)
- Law on Information Security (NN 79/2007)
- Copyright and Related Rights Act (NN 111/2021)
- Family Law (NN 98/2019)

In January 2022, the European Commission adopted a draft declaration on European digital rights and principles (European Commission, 2022). The principles are grouped into 6 thematic areas: placing people at the center of the digital revolution, solidarity and inclusion, freedom of choice, engaging in public digital space, empowerment and security, and sustainability. Through these values, the significance of upholding democracy, safeguarding human rights, encouraging responsible and safe behavior among all stakeholders, encouraging unity rather than division, maintaining a peaceful online environment free of objectionable or threatening material, participating in democratic processes, and the existence of digital devices that support sustainability and the green transition are all brought to light (European Commission, 2022). As a result, the following criteria were the main focus of the document analysis:

- Copyright protection (C1)
- Državni pedagoški standard osnovnoškolskog sustava odgoja i obrazovanja (NN 63/2008)
- Državni pedagoški standard srednjoškolskog sustava odgoja i obrazovanja (NN 63/2008)
- Pravilnik o kriterijima za izricanje pedagoških mjera (NN 3/2017)
- Opća uredba o zaštiti osobnih podataka (EU 2016/679)
- Zakon o provedbi Opće uredbe o zaštiti podataka (NN 42/2018)
- Zakon o informacijskoj sigurnosti (NN 79/2007)
- Zakon o autorskim i srodnim pravima (NN 111/2021)
- Obiteljski zakon (NN 98/2019)

Europska Komisija, u siječnju 2022. godine, donijela je nacrt deklaracije o europskim digitalnim pravima i načelima (Europska Komisija, 2022). Načela su grupirana 6 tematskih područja; ljudi u središtu digitalne transformacije, solidarnost i inkluzija, sloboda izbora, sudjelovanje u javnom digitalnom prostoru, sigurnost, zaštita i osnaživanje i održivost. Kroz ta načela provlači se važnost zaštite ljudskih prava, podrška demokracije, osiguranja da se svi dionici ponašaju odgovorno i sigurno, promicanja jedinstva, a ne podjele, mirno digitalno okruženje, bez ilegalnog i ugrožavajućeg sadržaja, uključivanja u demokratske procese kao i važnost kontrole nad vlastitim podatcima, postojanje digitalnih uređaja koji promiču održivost i potiču zelenu tranziciju (Europska Komisija, 2022). Sukladno navedenom, analiza dokumenata usmjerila se na sljedeće kriterije:

- Zaštita autorskog prava (C1)
- Reguliranje (prihvatljivog) ponašanja u digitalnom školskom okruženju (C2)
- Zaštita privatnosti u digitalnom školskom okruženju (C3)
- Sigurnost na mreži (C4)

Da bi se odgovorilo na drugo istraživačko pitanje provelo se kvalitativno istraživanje javno dostupne i službene dokumentacije škola.

Opću populaciju čini 1 315 redovnih matičnih osnovnih i srednjih škola[1], od kojih je metodom slučajnog odabira (uz pomoć *random number generator*) odabrano 308 škola, a zadržano 303, što čini reprezentativan uzorak u skladu s pristupom određivanja veličine slučajnog uzorka prema Krejcie i Morgan (1970, prema Cohen i sur., 2007). Škole koje su izostavljene iz istraživanja pripadaju: alternativnim školama, školama nacionalnih manjina čije internetske stranice nisu dostupne na hrvatskom ili engleskom jeziku te centri za odgoj i obrazovanje.

U ukupnom uzorku od 303 škole, 68% čine osnovne škole (206), a 32% srednje škole (97).

Kroz razdoblje od 1.6. do 10.6.2022. godine s ciljem povećanja valjanosti, dva nezavisna

[1] Podatci o tome prikupljeni s web aplikacije Ministarstva znanosti i obrazovanja, tzv. ŠeR-u.

- Regulating (acceptable) behavior in the digital learning environment (C2)
- Privacy protection in the digital learning environment (C3)
- Network security (C4)

In order to answer the second research question, a qualitative research of publicly available and official school documentation was conducted. The general population consists of 1,315 regular primary and secondary schools, out of which 308 schools were randomly selected and 303 held onto, which is a representative sample in accordance with the determination of random sample size approach according to Krejcie and Morgan (1970, according to Cohen et al., 2007). Schools excluded from the survey are identified as: alternative schools, schools for national minorities whose websites are not available in Croatian or English, and centers for education. In the total sample of 303 schools, 68% are primary schools (206) and 32% secondary schools (97). During the period from 1.6. to 10.6.2022., in order to increase the validity, two independent researchers reviewed and collected data available on the official websites of the schools. Among the available documentation on the school websites, the researchers noted whether there were internal regulations (provided by schools) on the use of ICT, and then analyzed them according to the criteria set in advance.

4 Results and discussion

4.1 Q1: Is the protection of the individual in the digital learning environment regulated by legal acts and sub-acts?

The qualitative analysis of legal regulations is to be commenced with a core legal act in the Republic of Croatia, i.e., the Constitution. It is aforementioned that the Constitution defines human rights as the highest values, and Article 37. clarifies and precises the protection of personal data, stating that everyone is "guaranteed the safety and confidentiality of personal data" and that only under statutory conditions can personal data be collected and processed and used without the consent of the individual. Article 69. of the Constitution explicitly guarantees the protection of moral and material rights deriving from scientific, cultural, artistic, intellectual and other creativity.

The protection of personal data is regulated by the General Regulation on Data Protection adopted by the European Parliament and the Council of the European Union (EU 2016/679), and which applies to all EU citizens. It considers the protection of the individual with regard to personal data processing to be a fundamental human right and puts the issue of

istraživača su pregledavala i prikupljala podatke dostupne na službenim web stranicama škola. Istraživači su među dostupnom dokumentacijom na stranicama škola bilježili postoje li interno doneseni akti o korištenju IKT, a potom ih analizirali obzirom na postavljene kriterije analize.

4 Rezultati i diskusija

4.1 Q1: Je li regulirana zaštita pojedinca u digitalnom školskom okruženju zakonskim aktima i podaktima?

Kvalitativnu analizu zakonskih regulativa započet ćemo temeljnim pravnim aktom u Republici Hrvatskoj, odnosno Ustavom. Već ranije je rečeno kako se njime prava čovjeka definiraju kao najveće vrednote, a člankom 37. jasnije se precizira zaštita osobnih podataka navodeći da se svakome „jamči sigurnost i tajnost osobnih podataka" te da se jedino u zakonima donesenim uvjetima osobni podatci mogu prikupljati, obrađivati i koristiti bez privole ispitanika. Ustavom se izrijekom člankom 69. jamči i zaštita moralnih i materijalnih prava koje proistječu iz znanstvenog, kulturnog, umjetničkog, intelektualnog i drugog stvaralaštva.

Zaštita osobnih podataka regulira se Općom uredbom o zaštiti osobnih podataka koju je donio Europski parlament i Vijeće (EU 2016/679) i koja se odnosi na sve građane EU. Njome se zaštita pojedinca s obzirom na obradu podataka smatra temeljnim pravom čovjeka, a pitanje zaštite osobnih podataka stavlja u suvremeni kontekst, odnosno kontekst brzog tehnološkog razvoja i globalizacije koji donose nove izazove u zaštiti osobnih podataka. Upravo radi toga navodi se u Uredbi da je potrebno stvoriti pravnu i praktičnu sigurnost pojedinaca. Iako se nigdje decidirano ne govori o regulaciji zaštite podataka unutar obrazovnog sustava, članak 38. govori da je potrebno posebno voditi računa o zaštiti podataka djece te da odgovarajuće nadzorno tijelo (što je u ovom slučaju Agencija za zaštitu osobnih podataka (dalje: AZOP)) posebno treba davati pozornost aktivnostima namijenjenim djeci u smislu promicanja javne svijesti o rizicima, pravilima te zaštitnim mjerama i pravima u vezi s obradom podataka. Tako je AZOP donio preporuke za primjenu Uredbe u školskim ustanovama (AZOP, 2019), a koja govore o potrebi imenovanja školskih službenika za zaštitu podataka te informiranje. Nadalje preporuke primjene Uredbe za školske ustanove govore o dužnostima škole u provođenju odgovarajućih tehničkih i organizacijskih mjera zaštite (čl. 25 i čl. 32) od npr. neovlaštenog pristupa ili raspolaganja osobnim podatcima, tehničke opreme, kao i osiguravanja sigurnosti mreže i informacijskih sustava subjekata koje svoje usluge nude putem mreže, kao što bi bio slučaj za vrijeme

personal data protection in a modern context, i.e., the context of rapid technological development and globalization that brings new challenges in personal data protection. That is why the General Regulation on Data Protection states that it is necessary to create legal and practical security for individuals. Although there is nowhere to be found any definite regulation of data protection within the education system, Article 38. states that special attention should be paid to the protection of children's data and that the appropriate competent authorities (in this case the Croatian Personal Data Protection Agency) (AZOP) should pay special attention to activities aimed at children in terms of promoting public awareness of risks, rules and safeguards and rights related to data processing.

Thus, AZOP made recommendations for the application of the Regulation in school institutions (AZOP, 2019), which advocate the need to appoint school protection data and information officer. Furthermore, the recommendations for the application of the General Regulation on Data Protection, for school institutions, states school duties in implementing appropriate technical and organizational measures for protection (Art. 25 and Art. 32) from e.g., unauthorized access or disposal of personal data, technical equipment and security of network and information systems which offer their services through the network, as would be the case during the implementation of distance learning (Art. 39.). During the complete transition to distance learning during the pandemic, AZOP addressed some of the basic issues related to the collection and processing of personal data (AZOP, 2021), but did not provide any specific guidelines regarding privacy and data protection of students in digital environments. Thus, it can be concluded that the General Regulation on Data Protection does not explicitly regulate the protection of individuals data in the digital learning environment, and what is provided are non-binding recommendations. None of the issues related to the protection of the individual in the digital environment is regulated by the Law on the Implementation of the General Regulation on Data Protection (NN 42/2018).

Regarding the regulation of online safety, the Information Security Act establishes security measures and standards of safety (in the areas of security clearance, physical security, data security, information system and business cooperation) which are applicable to all legal entities with public authority, which means school facilities as well (NN 79/2007).

The Family Law (NN 98/2019) regulates the rights of the child, which implies the right to parental care, which consists of responsibilities, duties, and rights of parents for the purpose of protecting and promoting the child's well-being, which is not exclusive to the digital school environment. Among other things, the scope of parental care includes the provedbe nastave na daljinu (čl.39). Za vrijeme potpunog prelaska nastave na daljinu za vrijeme pandemije, AZOP je obradio neka od osnovnih pitanja vezanih za prikupljanje i obradu osobnih podataka (AZOP, 2021), no nije dao konkretne zaštitne mjere u vezi privatnosti i zaštite podataka učenika u digitalnim okruženjima. Dakle, može se zaključiti da Općom uredbom nije eksplicitno regulirana zaštita podataka pojedinca u digitalnom školskom okruženju, a ono što je dano ostaje na neobvezujućim preporukama. Nijedno od pitanja vezanih za zaštitu pojedinca u digitalnom okruženju nije regulirano ni Zakonom o provedbi Opće uredbe o zaštiti podataka (NN 42/2018).

Po pitanju regulacije sigurnosti na mreži, Zakonom o informacijskoj sigurnosti utvrđuju se mjere i standardi sigurnosti (u područjima sigurnosne provjere, fizičke sigurnosti, sigurnosti podataka, informacijskog sustava i poslovne suradnje) koja se primjenjuju na sve pravne osobe s javnim ovlastima što znači onda i na školske ustanove (NN 79/2007).

Obiteljskim zakonom (NN 98/2019) regulirana su prava djeteta, što podrazumijeva pravo na roditeljsku skrb, koju čine odgovornosti i dužnosti i prava roditelja u svrhu zaštite i promocije djetetove dobrobiti, a što nije isključujuće za digitalno školsko okruženje. Između ostalog u sadržaj roditeljske skrbi ulazi pravo i dužnost zaštite prava djeteta na odgoj i obrazovanje, koje se regulira člankom 94. navedenog zakona. (NN 98/2019). Tim člankom navedena je i obveza roditelja da odgajaju djecu u skladu s njihovom dobi i zrelošću te ga štite od manipulirajućih i ponižavajućih postupaka drugih osoba. Iako se ovim zakonom izričito ne regulira zaštita djece u digitalnom okruženju, može se reći kako je pravna obveza primjenjiva kako u fizičkom tako i u digitalnom okruženju.

U nastavi na daljinu, nastavnici su uglavnom samostalno kreirali digitalne obrazovne sadržaje koje su dijelili i koji su bili namijenjeni učenicima kojima su predavali. Pri tome se povećala razina rizika i prijetnji povredama autorskih i srodnih prava te intelektualnog vlasništva. Ponajprije se to odnosi na neovlašteno dijeljenje i korištenje digitalnih materijala, predstavljanje tuđih radova kao svoje ili korištenje tuđih materijala bez citiranja, navođenja izvora itd. Autorska i srodna prava regulirana su Zakonom o autorskom i srodnim pravima (NN 111/2021), a obuhvaćaju moralna i imovinska prava koja pripadaju autoru (čl. 26). Imovinska prava autora obuhvaćaju pravo umnažanja, distribuiranja, činjenje dostupnim javnosti itd., dok se moralna prava odnose na pravo prve objave, poštovanja časti i ugleda autora te poštovanje autorskog djela. U vrijeme pojave pandemije na snazi je bio stari zakon iz 2018. godine (NN, 96/2018) kojim se navodilo da je dopušteno reproduciranje i korištenje (uz pravilno citiranje i navođenje izvora) zakonito objavljenih autorskih djela, ako su namijenjena isključivo nastavi, osim ako to autor izričito ne zabrani ili ako

right and duty to protect the child's right for education, which is regulated by Article 94 of the aforementioned law (NN 98/2019). That article also states parents' obligation to raise children in accordance with their age and maturity and to protect them from the manipulative and humiliating actions of other people. Although this law does not specifically regulate the protection of children in the digital environment, it can be concluded that the legal obligation exceeds physical environment and is applicable in the digital environment.

While distance teaching, teachers mostly created their own digital educational content that they shared and that were intended for the students they taught. At the same time, the level of risk and threat of copyright, related rights and intellectual property infringement increased. This primarily refers to the unauthorized distribution and use of digital materials, presenting other people's works as one's own or using other people's materials without quoting or citing sources, etc. Copyright and related rights are regulated by the Copyright and Related Rights Act (NN 111/2021) and include moral and property rights of the author (Art. 26.). The property rights of the author include the right to reproduce, distribute, make available to the public, etc., while the moral rights relate to the right of first publication, respect for the honor and reputation of the author and respect for the author's work. At the time of the pandemic, the old law from 2018 (NN 96/2018) was in force, which stated that the reproduction and use (with proper quotation and citation of sources) of legally published copyright works was allowed, if they were intended exclusively for teaching, except if the author does not explicitly forbid it or if the manner of use would endanger the honor and reputation of the author (Art. 85 and Art. 88). Following the pandemic, the aforementioned Law was revisited (NN 111/2021), adding the Article 199. *Use in digital and cross-border teaching*, which means "without the approval of the right holder and without payment of compensation, reproduction and communication to the public, including making it publicly available, copyrighted works and related rights for their digital use and for the purpose of setting an example in teaching practice" and under the condition that such use takes place within the competence of the educational institution, on its premises or in other facilities or through a secure electronic environment accessible only to pupils or students and teaching staff of that educational institution ", alongside proper quotation, citation of sources and authors. In particular, Paragraph 2 of this Article states that this applies to digital and online teaching, distance learning, cross-border teaching and across all levels of education. Therefore, this Law best regulates the issue of protection of the individual in the digital learning environment, given the criteria of copyright protection.

bi se načinom korištenja ugrozila čast i ugled autora (čl. 85 i čl. 88). Nakon pojave pandemije, dolazi do izmjena spomenutog Zakona (NN, 111/2021), kada se dodaje i članak 199. *Korištenje u digitalnoj i prekograničnoj nastavi* što podrazumijeva „bez odobrenja nositelja prava i bez plaćanja naknade, umnožavanje i priopćavanje javnosti, uključujući činjenje dostupnim javnosti, autorskih djela i predmeta srodnih prava radi njihova digitalnog korištenja u svrhu davanja primjera u nastavi" te pod uvjetom da se "takvo korištenje odvija u okviru nadležnosti obrazovne ustanove, u njezinim prostorijama ili u drugim objektima ili putem sigurnog elektroničkog okružja kojemu mogu pristupiti samo učenici ili studenti i nastavno osoblje te obrazovne ustanove", a uz pravilno citiranje, navođenje izvora i autora. Posebno je naznačeno stavkom 2. ovog članka da se to odnosi na digitalnu i internetsku nastavu, učenje na daljinu, prekograničnu nastavu te na sve obrazovne razine. Dakle, ovim Zakonom najbolje je uređeno pitanje zaštite pojedinca u digitalnom školskom okruženju, obzirom na kriterij zaštite autorskih prava.

Unutar krovnog zakona koji regulira djelatnost osnovnog i srednjeg obrazovanja Zakona o odgoju i obrazovanju u osnovnoj i srednjoj školi (dalje: ZOO) člankom 67. definira se sigurnost u školskim ustanovama, prema kojem su školske ustanove dužne stvoriti uvjete za zdrav mentalni i fizički razvoj te socijalnu dobrobit; voditi evidenciju, ali i sprječavati oblike neprihvatljivog ponašanja učenika; brinuti se o sigurnosti učenika; i pratiti socijalne probleme učenika te poduzimati adekvatne mjere za otklanjanje njihovog uzroka i posljedica. Osim toga, reguliraju se obveze učenika koje podrazumijevaju poštivanje pravila kućnog reda kao i drugih pravnih propisa (čl.61) što znači da pravilnici koji bi regulirali sigurnu i odgovornu primjenu IKT-a i regulirali zaštitu pojedinca u digitalnom školskom okruženju imaju uporište i opravdanje u zakonu. U slučaju povreda dužnosti, neprihvatljivog i nasilnog ponašanja te neispunjavanja obveza učenicima se izriču pedagoške mjere prema Pravilniku o kriterijima za izricanje pedagoških mjera (NN 3/2017).

Ovim Pravilnikom obuhvaćena su neka neprihvatljiva ponašanja koja se izrijekom odnose na ponašanja u digitalnom okruženju, ili su vezana za sigurnu i odgovornu upotrebu IKT-a poput: nedopuštenog korištenja informacijsko-komunikacijskih uređaja tijekom odgojno-obrazovnog rada, objavljivanje materijala elektroničkim ili drugim putem, a koji za posljedicu imaju povredu ugleda, časti i dostojanstva druge osobe ili oštećivanje imovine u prostorima škole. No, iako za druga neprihvatljiva ponašanja nije izrijekom rečeno da se ona odnose i na ponašanja u virtualnim učionicama, odnosno u digitalnom okruženju, ona su primjenjiva. Primjeri za to su: korištenje nedopuštenih izvora podataka u svrhu prepisivanja (u

Within the supreme law regulating the activity of primary and secondary education, the Primary and Secondary Education Act (ZOO) by Article 67. defines safety in school institutions, according to which school institutions are obliged to create conditions for healthy mental and physical development and social welfare; keep evidence, but also prevent different unacceptable student behaviour; take care of student safety; monitor the social problems of students and take adequate measures to eliminate their causes and consequences. Furthermore, the obligations of students are regulated, which include compliance with house rules and other legal regulations (Art. 61.), which means that regulations that would regulate the safe and responsible use of ICT and regulate the protection of individuals in the digital learning environment have a basis and justification in law. In case of breaches of duty, unacceptable and violent behavior and non-fulfillment of obligations, students are imposed pedagogical measures according to the Policy on criteria for imposing pedagogical measures (NN 3/2017).

This Policy covers some unacceptable behaviors that are explicitly related to behaviors in the digital environment or are related to the safe and responsible use of ICT, such as: unpermitted use of information and communication devices during teaching, publishing materials electronically or otherwise, and which result in damage to the reputation, honor and dignity of another person, or damage to property on school premises. However, although other unacceptable behaviors are not explicitly stated to apply to behaviors in virtual classrooms or digital environments, they are applicable. Examples for that are: using unauthorized data sources for the purpose of copying (in the form of copyright infringement), violating the dignity of another person by belittling, insulting, or spreading untruths and rumors about another student or school worker, using or abusing another student's data from pedagogical documentation or encouraging group hate speech.

Furthermore, by the ZOO is explicitly stated (Art. 58.) that school boards are obliged to specify house rules, but not a similar document that would regulate safe and responsible so as acceptable behavior in virtual classrooms. However, this does not exclude the possibility of regulating the same within house rules. Also, Article 49 of the ZOO states that schools are obliged to provide conditions for the realization of their pedagogical and public functions, which also leaves the legal possibility for schools to independently regulate the safe and responsible use of ICT by internal regulations. Regarding the issue of personal data protection, the ZOO states that they are regulated in accordance with the provisions of regulations governing data protection (NN 64/2020).

The National pedagogical standards of the primary and secondary school education system (NN 63/2008) do not regulate any of the four set criteria.

vidu povrede autorskih prava), povreda dostojanstva druge osobe omalovažavanjem, vrijeđanjem ili širenjem neistina i glasina o drugome učeniku ili radniku škole, korištenje ili zlouporaba podataka drugog učenika iz pedagoške dokumentacije ili poticanje grupnoga govora mržnje.

Nadalje, u ZOO izrijekom je jedino navedeno (čl. 58.) da su školski odbori dužni donijeti pravila kućnoga reda, no ne i sličan dokument kojim bi se regulirala sigurna i odgovorna te prihvatljiva ponašanja u virtualnim učionicama. Ipak, to ne isključuje mogućnost regulacije istog kućnim redom. Isto tako, u članku 49. ZOO-a navodi se da su škole dužne osigurati uvjete za ostvarenje svoje pedagoške i javne funkcije što također ostavlja pravnu mogućnost školama da samostalno internim pravilnicima reguliraju sigurnu i odgovornu upotrebu IKT-a. Što se tiče pitanja zaštite podataka u ZOO se navodi da se ona reguliraju u skladu s odredbama propisa koji uređuju zaštitu podataka (NN 64/2020).

Državnim pedagoškim standardima osnovnoškolskog i srednjoškolskog sustava odgoja i obrazovanja (NN 63/2008) ne regulira se nijedan od četiri kriterija. Iako postoji standard koji predviđa prostorne uvjete za izvođenje informatike, no ne govori ništa, primjerice o standardima sigurnosti na mreži i sl.

Konačno, može se reći da je tek u jednom zakonu eksplicitno regulirana zaštita pojedinca u digitalnom (školskom) okruženju i to s aspekta zaštite autorskih i srodnih prava te o djelomičnoj reguliranosti prihvatljive komunikacije i oblike ponašanja pri korištenju IKT, samo jednim pravilnikom. Ipak, primjenom različitih zakonskih regulativa na školske ustanove, može se reći da je u određenoj mjeri regulirana zaštita pojedinca u digitalnom okruženju, a u najvećoj mjeri s obzirom na sigurnost na mreži. Pitanje zaštite privatnosti i podataka te regulacije prihvatljivog ponašanja i komunikacije u školskom digitalnom okruženju ostaje nedovoljno pokriveno. Nakon svega, može se zaključiti da navedeni pravni akti zasebno nedovoljno reguliraju pitanje zaštite pojedinca u (školskom) digitalnom okruženju, ali u dovoljnoj mjeri daju pravnu podlogu i lako su primjenjivi za kreiranje zasebnog pravilnika o sigurnom i odgovornom korištenju IKT, po uzoru na prijedlog istoimenog pravilnika koji je nastao u sklopu projekta e-Škole, a koji će posljedično imati zaštitu pojedinca u digitalnom (školskom) okruženju. Kako bi školske ustanove dobile pravnu sigurnost u regulaciji istoga, ali se istaknula i važnost reguliranja navedenog, autori predlažu da se jedan takav pravilnik donese na (makro) nacionalnoj razini.

4.2 Q2: Jesu li osnovne i srednje škole samostalno, internim aktima

Although there is a standard that assumes spatial conditions for teaching informatics, it does not say anything, for example about online security standards etc.

To conclude, it can be said that only one law explicitly regulates the protection of individuals in the digital (learning) environment, which is from the aspect of copyright and related rights protection. There is also partial regulation of acceptable communication and behavior when using ICT, only by one policy. However, by applying various legal regulations to school institutions, it can be said that the protection of the individual in the digital environment is regulated to a certain extent, and to the greatest extent with regard to online security. The issue of privacy and data protection and the regulation of acceptable behavior and communication in the school digital environment remains insufficiently covered. Finally, it can be concluded that these legal acts do not sufficiently regulate the protection of individuals in digital (learning) environment but provide a sufficient legal foundation and could be easily applied in creation of a separate guidelines or policy on safe and responsible use of ICT. It could be modeled based on the proposal of the guidelines, under the same name, created as part of the e-School project, which will consequently have the role of protection of the individual in the digital (learning) environment. In order for school institutions to gain legal certainty in the regulation of that matter, as well as to emphasize the importance of regulating the above mentioned, the authors suggest that such policy or regulations should be provided at the (macro) national level.

4.2 Q2: Did primary and secondary schools independently, by internal regulations, regulate the problem of individual protection in the digital learning environment?

Qualitative analysis of data available on the websites of primary and secondary schools showed that only 15 schools out of a total of 303 observed internally regulated the safe and responsible use of ICT and the protection of individuals in the digital learning environment by some form of an individual act. Of the total sample, 5% did so, mostly primary schools, 12 out of 15 (other data are given in Table 1). Podbojec and Mekovec (2021) conducted research on data protection at universities, which coincides with criterion C3 of this research. Their results also indicate insufficient awareness of educational systems about the need to coordinate internal regulations with valid and current legal acts that define data protection.

Most schools have created their own regulations according to the proposed CARNET (2017) Ordinance on safe and responsible use of information

regulirale problem zaštite pojedinca u školskom digitalnom okruženju?

Kvalitativnom analizom podataka dostupnih na web stranicama osnovnih i srednjih škola pokazalo se da je samo 15 škola od ukupno promatranih 303 interno nekim oblikom zasebnog akta reguliralo sigurnu i odgovornu upotrebu IKT i zaštitu pojedinca u digitalnom školskom okruženju. Od ukupnog uzorka to je učinilo 5% škola, a uglavnom su to osnovne škole, njih 12 od 15 (ostali podatci dani su u Tablici 1). Podbojec i Mekovec (2021) proveli su istraživanje o zaštiti podataka na visokim učilištima što se poklapa s kriterijem C3 ovog istraživanja. Njihovi rezultati također ukazuju na nedovoljnu osviještenost obrazovnih sustava o potrebi usklađivanja internih regulativa s važećim zakonskim aktima koji definiraju zaštitu podataka.

Većina škola kreirala je svoje pravilnike prema danom prijedlogu CARNET-a (2017) Pravilnika o sigurnoj i odgovornoj upotrebi informacijsko-komunikacijske tehnologije u školi, a koji je donesen u sklopu provedbe projekta e-Škole: Uspostava sustava razvoja digitalno zrelih škola (dalje: e-Škole), njih 66,7%. Ostale škole kreirale su vlastite interne akte u obliku odluka o (prihvatljivom) korištenju interneta, računala i mobitela u školi i izvan nje, etičkih kodeksa o sigurnosti učenika i djelatnika na internetu te odluka o računalnoj sigurnosti i sigurnosti informacijskih sustava.

Tablica 1. Distribucija rezultata obzirom na razinu školskog obrazovanja

	Ima Pravilnik	Nema Pravilnik	Ukupno	Postotak (“ima”)
Osnovne škole	12	194	206	5,8%
Srednje škole	3	94	97	3,1%
Ukupno	15	288	303	5%

4.2.1 Q2.1: Koji su aspekti zaštite pojedinca u digitalnom školskom okruženju obuhvaćeni pravilnicima?

Kvalitativnom analizom 15 pronađenih internih regulativa škola vezanih za zaštitu pojedinca u digitalnom školskom okruženju, utvrđeno je da se one u određenoj mjeri sadržajno razlikuju, obzirom na obuhvaćenost postavljenih kriterija analize (C1–C4). U najvećoj mjeri, osam škola (oko 53%) obuhvatilo je sve aspekte zaštite pojedinca u digitalnom okruženju, u kontekstu ovog rada. Šest škola (40%) reguliralo je 3 od pretpostavljena 4 aspekta, a u svim slučajevima izostavljen je kriterij zaštite privatnosti u školskom digitalnom okruženju, dok je samo jedna škola (oko 7%) koja je regulirala samo dva aspekta, u kontekstu ovog rada (samo C2 i

and communication technology in schools, which was developed as part of the e-School project: Establishment of a system for developing digitally advanced schools (e-Schools), 66.7% of them. Other schools have created their own internal acts in the form of decisions on (acceptable) use of the Internet, computers, and mobile phones in and out of school, ethical codex on the safety of students and staff on the Internet and resolutions on computer security and information systems security.

Table 1. Distribution of results according to the level of school education

	Have Ordinance	Don't have Ordinance	∑	% ("have")
Primary schools	12	194	206	5,8%
Secondary schools	3	94	97	3,1%
Sum	15	288	303	5%

4.2.1 Q2.1: What aspects of individual protection in digital learning environment are covered by regulations?

Through a qualitative analysis of 15 found internal regulations of schools, related to the protection of individuals in the digital learning environment, it was concluded that they differ in content to a certain extent, considering the scope of the set analysis criteria (C1–C4). To the greatest extent, eight schools (about 53%) covered all aspects of individual protection in the digital environment, in the context of this paper. Six schools (40%) regulated 3 of 4 proposed aspects, and in each of those cases the criterion of privacy protection in the digital learning environment was overlooked, while only one school (about 7%) regulated only two aspects, in the context of this paper (only C2 and C3). In other words, the analysis showed that in all acts (100%), (un)acceptable forms of behaviour in the digital learning environment are regulated and defined, and in 93.3% of the acts, general safety provisions for digital equipment and how to achieve online security were adopted. To a large extent, i.e., in 86.7% of acts, the protection of the copyright of participants in the digital learning environment is regulated, while to the smallest extent (73.3%) the issue of protecting individual privacy is regulated. The analysis of the mentioned acts exhibits insufficient elaboration and narrow description of the protection of individual privacy, furthermore, they do not consider, include, or regulate new challenges that have emerged in this area due to the onset of the pandemic and transition to distance learning (e.g., distance teaching and learning via Zoom and Big Blue Button applications C3). Odnosno, analiza je pokazala da su u svim aktima (100%) regulirani i definirani (ne)prihvatljivi oblici ponašanja u digitalnom školskom okruženju, a u 93,3% akata donesene su opće sigurnosne odredbe digitalne opreme te kako se ostvaruje sigurnost na mreži. U velikoj mjeri, odnosno u 86,7% akata regulirana je zaštita autorskih prava sudionika u digitalnom školskom okruženju, dok je u najmanjoj mjeri (73,3%) regulirano pitanje zaštite privatnosti pojedinca. Analizom spomenutih akata ustvrđeno je kako je zaštita privatnosti pojedinca nedovoljno elaborirana te minimalno regulirana, a novi izazovi koji su se u tom području nametnuli pojavom pandemije i prelaskom na nastavu na daljinu nisu razmotreni, uključeni niti regulirani navedenim aktima (npr. izvođenje nastave na daljinu putem Zoom i Big Blue Button aplikacije ili dijeljenje privatnih podataka učenika poput ocjena na polu – javnim platformama kao što su Ms Teams i Google Classroom).

Većina propisa donesena je kroz razdoblje od 2017. do 2021. godine, iznimno jedan pravilnik je donesen 2013. godine, a za jedan nije poznato. U tablici 2 dana je distribucija godina donošenja propisa. Ove rezultate zanimljivo je promatrati u odnosu vrijeme pojave COVID-19 pandemije. Prije pojave pandemije i prelaska na nastavu na daljinu, internim aktima pitanje regulacije sigurnog digitalnog okruženja i zaštite pojedinca u istom imalo je 13 od 15 škola koje uopće imaju regulirano ovo pitanje, tj. 4,3% škola od ukupnog uzorka. Drugim riječima, nakon pojave pandemije i za vrijeme iste, manje od 1% škola je regulirano navedeni problem. Budući da je istraživanje rađeno na reprezentativnom uzorku, rezultati se mogu generalizirati na opću populaciju redovnih matičnih škola u Hrvatskoj (Cohen i sur., 2007).

Tablica 2. Prikaz distribucije pravilnika obzirom na godinu donošenja

Godina donošenja	Broj škola
2013	1
2017	1
2018	6
2019	2
2020	3
2021	1
nepoznato	1

Osim toga, rezultate je zanimljivo promatrati u odnosu na provedbu projekta e-Škole jer je u sklopu njega jedna od preporuka bilo kreiranje Pravilnika o sigurnoj i odgovornoj upotrebi IKT-a u školi kao dijela sigurnosne politike digitalno zrele škole. Od ukupno 15 škola koje su donijele slične propise i time regulirale zaštitu pojedinca u digitalnom okruženju, njih 6 (40%) je bilo sudionikom projekta.

or sharing personal students' data on semi – public platforms such as Ms Teams or Google Classroom). In accordance with the set criteria of qualitative analysis of legal regulations related to the protection of individuals in the digital learning environment, a qualitative analysis of 15 school acts was conducted. All acts (100%) regulate and define (un)acceptable behavior in the digital learning environment, and 93.3% of acts adopt general security provisions for digital equipment and how online security is achieved. The protection of copyrights of participants in the digital learning environment is largely regulated, i.e., in 86.7% of acts, while the issue of protection of individual privacy is regulated to the smallest extent (73.3%).

Most of the regulations were adopted during the period from 2017 to 2021, exceptionally one ordinance was adopted in 2013, and for one this data is unknown. Table 2 shows the distribution of the years of enactment. These results are interesting to observe in relation to the timing of the COVID-19 pandemic. Before the pandemic and the transition to distance learning, 13 out of the 15 schools (which generally have internally regulated safe digital learning environment and protection of the individual in digital learning environment) had some form of regulation, i.e., 4.3% of the total sample. In other words, after the onset of a pandemic and during it, less than 1% of schools have regulated this problem. Since the research was done on a representative sample, the results can be generalized to the general population of regular main schools in Croatia (Cohen et al., 2007).

Table 2. Distribution of regulations according to the year of adoption

Year	Num. of schools
2013	1
2017	1
2018	6
2019	2
2020	3
2021	1
unknown	1

Moreover, it is interesting to observe the results in relation to the implementation of the e-School project because one of its recommendations was to develop Ordinance on safe and responsible use of ICT in school as part of the security policy of digitally advanced schools. Out of a total of 15 schools that adopted similar regulations and thus regulated the protection of the individual in the digital environment, 6 (40%) were participants in the project.

4.3 Ograničenja istraživanja i preporuke za buduća istraživanja

Ograničenja ovog istraživanja vezana su uz metodologiju. Iako su postavljeni kriteriji analize proizašli iz teorijskih razmatranja, zasigurno postoji još kriterija po kojima bi se mogla i trebala napraviti analiza zakonskih regulativa, kao i analiza nekih drugih dokumenata poput mišljenja koje su publicirale Agencija za zaštitu osobnih podataka ili smjernica Ministarstva znanosti i obrazovanja i sl.

Osim toga, u istraživanju se ispitivalo postoje li zasebni pravilnici koji reguliraju zaštitu pojedinca u digitalnom školskom okruženju, a autori smatraju da postoji mogućnost da su školske ustanove pravila prihvatljivog i odgovornog ponašanja te sigurne upotrebe IKT regulirale u sklopu kućnog reda, etičkog kodeksa, statuta škole ili nekog drugog internog dokumenta. Upravo u tom smjeru ide preporuka za daljnja istraživanja da se ispita jesu li škole nekim drugim internim dokumentima regulirale zaštitu pojedinca u digitalnom okruženju, tj. oblike sigurne, odgovorne i prihvatljive upotrebe IKT.

Nadalje, potencijalno ograničenje vezano je uz odabir uzorka. U ovom istraživanju uzorak su činile matične škole, ali ne i područne, kojih je još 1 115. Iako se prilikom pregleda web stranica matičnih škola moglo uočiti da područne škole uglavnom imaju zajedničku web stranicu sa matičnim školama te nemaju posebno istaknute vlastite interne dokumente nego slijede one matičnih škola, ne može se bez empirijske potvrde tvrditi da se rezultati dobiveni u ovom istraživanju odnose i na područne škole.

5 Zaključak

Digitalno školsko okruženje te digitalna zrelost škola prirodan su slijed razvoja odgojno – obrazovnog sustava koji prati tehnološke i digitalne napretke 21. stoljeća, ali ciljeva postavljenih Strategijom obrazovanja, znanosti i tehnologije (NN 124/2014). Taj proces sazrijevanja digitalno zrelih škola nedvojbeno je ubrzala pojava COVID–19 pandemije te prisilan prijelaz na nastavu na daljinu, kako diljem svijeta, tako i u Hrvatskoj. S obzirom na nagli prijelaz na online nastavu nije iznenađujuće kako je u prvim trenutcima izostala standardizirana uputa, nacionalnih tijela i agencija, usmjerena na zaštitu pojedinca u školskom digitalnom okruženju.

Analizom zakonskih regulativa utvrđeno je da postoji dobra pravna podloga za reguliranje zaštite pojedinca u (školskom) digitalnom okruženju te da su određeni aspekti zaštite (pitanje zaštite autorskih prava) dobro regulirani. Većina zakonskih regulativa koji su nadležni za pitanje zaštite pojedinca u digitalnom okruženju predstavljaju dobru podlogu za donošenje pravilnika o sigurnoj i odgovornoj

4.3 Limitations of research and recommendations for future research

The limitations of this research are related to the methodology. Although the set criteria of analysis are based on theoretical background, there are certainly other criteria by which the analysis of legal acts could and should be made. In addition, there are as well some other documents such as opinions published by Croatian Personal Data Protection Agency or guidelines of the Ministry of Science and Education which could be included in further research.

Besides, the research examined whether there are separate regulations governing the protection of individuals in the digital learning environment, and the authors believe that there is a possibility that schools have regulated rules of acceptable and responsible behavior and safe use of ICT within the house rules, ethical codex, school statute or some other internal document. Accordingly, authors for further research recommend examining whether schools have regulated the protection of the individual in the digital environment with some other internal documents, i.e., forms of safe, responsible, and acceptable use of ICT.

Furthermore, a potential limitation of research is related to sample selection. In this research, the sample consisted of main schools, but not dislocated (subordinate) schools, whose number is 1,115. Although when reviewing the websites of the main schools, we noticed that the dislocated schools mostly have a common website with the main schools and don't have their own internal documents, but follow those of the main schools, it cannot be claimed without empirical confirmation that the results obtained in this research also apply to dislocated schools.

5 Conclusion

Digital learning environment and digital advancement of schools present a natural flow of educational system development that follows technological and digital progress of 21st century as well as the goals set in the Strategija obrazovanja, znanosti i tehnologije (2014). That development process of digitally advanced schools was undoubtedly accelerated by COVID-19 pandemic and forced transition to distance teaching and learning all over the world as well as in Croatia. Taking into consideration the abrupt transition to online teaching and learning, it comes as no surprise that in the first moments there had been a lack of standardized instruction of competent authorities on the national level regarding the protection of an individual in a digital school environment.

Analyzing the law regulations, it has been determined that there is solid legal foundation for upotrebi IKT-a, s ciljem zaštite pojedinca u školskom digitalnom okruženju.

Iako ne postoji takav pravilnik na makro (nacionalnoj) razini, pojedine škole su samoinicijativno ili u sklopu CARNET-ovog projekta e-Škole donijele slične akte koji reguliraju zaštitu pojedinca u školskom digitalnom okruženju. Ipak, rezultati broja škola koje imaju donesen takav propis je više nego nezadovoljavajući i oko 5%. Pri tome, valja istaknuti kako je većina škola koje su donijele slične akte to učinila prije pojave pandemije, a čak manje od 1% škola je smatralo potrebnim to učiniti i nakon pojave pandemije, odnosno pri prelasku na nastavu na daljinu.

References

[AZOP] Agencija za zaštitu osobnih podataka (2019). Primjena Opće uredbe o zaštiti osobnih podataka u školskim ustanovama. Retrieved from: https://azop.hr/primjena-opce-uredbe-o-zastiti-podataka-u-skolskim-ustanovama

[AZOP] Agencija za zaštitu osobnih podataka (2021). Obrada osobnih podataka u kontekstu pandemije COVID-19. Retrieved from: https://azop.hr/obrada-osobnih-podataka-u-kontekstu-pandemije-covid-19/

Balaban, I. (2017). *Sigurna i odgovorna upotreba informacijsko-komunikacijske tehnologije u školi – Prijedlog sadržaja dokumenta*. CARNET: Zagreb.

Barana, A. & Marchisio, M. (2021). Analyzing Interactions in Automatic Formative Assessment Activities for Mathematics in Digital Learning Environments. In b. Caspo & J. Uhomoibhi (Eds.) *Proceedings of the 13th International Conference on Computer Supported Education (CSEDU 2021)* (pp. 497-504).

Begičević Ređep, N., Balaban, I., Klačmer Čalopa, M. & Žugec, B. (2018). *Okvir za digitalnu zrelost osnovnih i srednjih škola u Republici Hrvatskoj s pripadajućim instrumentom*. CARNET: Zagreb.

CARNET (2020). Online sustavi za organizaciju i provođenje nastave na daljinu. Retrieved from: https://www.CARNET.hr/wp-content/uploads/2020/08/Online-sustavi-za-organizaciju-i-provo%C4%91enje-nastave-na-daljinu1.pdf

Cohen, L., Manion, L. & Morrison, K. (2007). Metode istraživanja u obrazovanju. NAKLADA SLAP: Jastrebarsko.

Cohney, S., Teixeira, R., Kohlbrenner, A., Narayanan, A., Kshirsagar, M., Shvartzshnaider, Y. & Sanfilippo, M. (2021). Virtual classrooms and real harms: Remote learning at U.S.

regulating the protection of an individual in a digital (school) environment and that some of those protection aspects (e.g., copyright) are well regulated. Most regulations in charge of the protection of an individual in a digital environment present a solid base for adoption of rules of procedure for safe and responsible ICT use, all with the aim of protecting an individual in a digital learning environment.

Even though there are no such rules or guidelines on a national level, some schools took self-initiative, or took part in a CARNET project e-Schools and adopted similar acts regulating the protection of an individual in a digital learning environment. Nevertheless, the number of schools that adopted such an act is below satisfactory rate and accounts for around 5%. Most of schools that adopted those acts did it before the pandemic started. Moreover, only less than 1% schools found it necessary to adopt such regulations with the emergence of the pandemic and the transition to remote online teaching.

References

[AZOP] Agencija za zaštitu osobnih podataka (2019). Primjena Opće uredbe o zaštiti osobnih podataka u školskim ustanovama. Retrieved from: https://azop.hr/primjena-opce-uredbe-o-zastiti-podataka-u-skolskim-ustanovama

[AZOP] Agencija za zaštitu osobnih podataka (2021). Obrada osobnih podataka u kontekstu pandemije COVID-19. Retrieved from: https://azop.hr/obrada-osobnih-podataka-u-kontekstu-pandemije-covid-19/

Balaban, I. (2017). *Sigurna i odgovorna upotreba informacijsko-komunikacijske tehnologije u školi – Prijedlog sadržaja dokumenta*. CARNET: Zagreb.

Barana, A. & Marchisio, M. (2021). Analyzing Interactions in Automatic Formative Assessment Activities for Mathematics in Digital Learning Environments. In b. Caspo & J. Uhomoibhi (Eds.) *Proceedings of the 13th International Conference on Computer Supported Education (CSEDU 2021)* (pp. 497-504).

Begičević Redep, N., Balaban, I., Klačmer Čalopa, M. & Žugec, B. (2018). *Okvir za digitalnu zrelost osnovnih i srednjih škola u Republici Hrvatskoj s pripadajućim instrumentom*. CARNET: Zagreb.

CARNET (2020). Online sustavi za organizaciju i provođenje nastave na daljinu. Retrieved from: https://www.CARNET.hr/wp-content/uploads/2020/08/Online-sustavi-za-organizaciju-i-provo%C4%91enje-nastave-na-daljinu1.pdf

Cohney, S., Teixeira, R., Kohlbrenner, A., Narayanan, A., Kshirsagar, M., Shvartzshnaider, Y. & Sanfilippo, M. (2021). Virtual Classrooms and Real Harms: Remote Learning at U.S. Universities. Retrieved from: https://www.usenix.org/system/files/soups2021-cohney.pdf

Dash, S., Samadder, S., Srivastava, A., Meena, R. & Ranjan, P. (2021). Review of Online Teaching Platforms in the Current Period of COVID-19 Pandemic. *The Indian journal of surgery*, 18, 1-6. https://doi.org/10.1007/s12262-021-02962-4 .

European Education Area (2021) Digital Education Action Plan (2021-2027). Retrieved from: https://education.ec.europa.eu/focus-topics/digital-education/digital-education-action-plan

European Education Area (n.d.) Digital education initiatives. Retrieved from: https://education.ec.europa.eu/focus-topics/digital-education/about

Europska Komisija (2022) Europska deklaracija o digitalnim pravima i načelima za digitalno desetljeće. Retrieved from: https://digital-strategy.ec.europa.eu/en/library/declaration-european-digital-rights-and-principles

Europski parlament i Vijeće (2006). Recommendation 2006/962/EC of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning. *Official Journal L 394*.

Europski parlament i Vijeće (2016). Opća uredba o zaštiti osobnih podataka (EU 2016/679). Retrieved from: https://www.zakon.hr/z/3112/Op%C4%87a-uredba-o-za%C5%A1titi-podataka---Uredba-%28EU%29-2016-679-

Fuchs, R., Vican, D. & Milanović Litre, I. (2011). Nacionalni okvirni kurikulum za predškolski odgoj i obrazovanje te opće obvezno i srednjoškolsko obrazovanje. Retrieved from: http://mzos.hr/datoteke/Nacionalni_okvirni_kurikulum.pdf

Lamsal, B. (2021). Issues Surrounding Safe Digital Learning Space for the Students/Learners in Nepal. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3924407

[MZO] Ministarstvo znanosti i obrazovanja (2020). *Akcijski plan za provedbu nastave na daljinu*. Retrieved from: https://mzo.gov.hr/UserDocsImages//dokumenti/Obrazovanje/NastavaNaDaljinu//Akcijski%20plan%20za%20provedbu%20nastave%20na%20daljinu%20-%20Model%20nastave%20na%20daljinu.pdf

[NN] Official Gazette of the Republic of Croatia (2014). *Zakon o informacijskoj sigurnosti*, 79/207 [30 July 2007]

Cohen, L., Manion, L. & Morrison, K. (2007). Metode istraživanja u obrazovanju. NAKLADA SLAP: Jastrebarsko.

Cohney, S., Teixeira, R., Kohlbrenner, A., Narayanan, A., Kshirsagar, M., Shvartzshnaider, Y. & Sanfilippo, M. (2021). Virtual classrooms and real harms: Remote learning at U.S. Universities. Retrieved from: https://www.usenix.org/system/files/soups2021-cohney.pdf

Dash, S., Samadder, S., Srivastava, A., Meena, R. & Ranjan, P. (2021). Review of Online Teaching Platforms in the Current Period of COVID-19 Pandemic. *The Indian journal of surgery*, 18, 1-6. https://doi.org/10.1007/s12262-021-02962-4 .

European Education Area (2021) Digital Education Action Plan (2021-2027). Retrieved from: https://education.ec.europa.eu/focus-topics/digital-education/digital-education-action-plan

European Education Area (n.d.) Digital education initiatives. Retrieved from: https://education.ec.europa.eu/focus-topics/digital-education/about

Europska Komisija (2022) Europska deklaracija o digitalnim pravima i načelima za digitalno desetljeće. Retrieved from: https://digital-strategy.ec.europa.eu/en/library/declaration-european-digital-rights-and-principles

Europski parlament i Vijeće (2006). Recommendation 2006/962/EC of the European Parliament and of the Council of 18 December 2006 on key competences for lifelong learning. *Official Journal L 394*.

Europski parlament i Vijeće (2016). Opća uredba o zaštiti osobnih podataka (EU 2016/679). Retrieved from: https://www.zakon.hr/z/3112/Op%C4%87a-uredba-o-za%C5%A1titi-podataka---Uredba-%28EU%29-2016-679-

Fuchs, R., Vican, D. & Milanović Litre, I. (2011). Nacionalni okvirni kurikulum za predškolski odgoj i obrazovanje te opće obvezno i srednjoškolsko obrazovanje. Retrieved from: http://mzos.hr/datoteke/Nacionalni_okvirni_kurikulum.pdf

Lamsal, B. (2021). Issues Surrounding Safe Digital Learning Space for the Students/Learners in Nepal. Retrieved from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3924407

[MZO] Ministarstvo znanosti i obrazovanja (2020). *Akcijski plan za provedbu nastave na daljinu*. Retrieved from: https://mzo.gov.hr/UserDocsImages//dokumenti/

[NN] Official Gazette of the Republic of Croatia (2008). *Državni pedagoški standard osnovnoškolskog sustava odgoja i obrazovanja,* 59/1990, 26/1993, 27/1993, 29/1994., 7/1996, 59/2001, 114/2001, 76/2005, 63/2008 [2 June 2008]

[NN] Official Gazette of the Republic of Croatia (2008). *Državni pedagoški standard srednjoškolskog sustava odgoja i obrazovanja,* 19/1992, 26/1993, 27/1993, 50/1995, 59/2001, 114/2001, 81/2005, 63/2008 [2 June 2008]

[NN] Official Gazette of the Republic of Croatia (2014). *Ustav Republike Hrvatske*, 56/1990, 135/1997, 8/1998, 113/2000, 124/2000, 28/2001, 41/2001, 55/2001, 76/2010, 85/2010, 5/2014 [15 January 2014]

[NN] Official Gazette of the Republic of Croatia (2014). *Strategija obrazovanja, znanosti i tehnologije*, 124/2014 [24 October 2014]

[NN] Official Gazette of the Republic of Croatia (2018). *Odluka o donošenju kurikuluma za nastavni predmet Informatike za osnovne škole i gimnazije u Republici Hrvatskoj*, 22/2018 [6 March 2018]

[NN] Official Gazette of the Republic of Croatia (2018). *Zakon o provedbi Opće uredbe o zaštiti podataka*, 42/2018 [9 May 2018]

[NN] Official Gazette of the Republic of Croatia (2018). *Zakon o autorskom pravu i srodnim pravima*, 167/2003, 79/2007, 80/2011, 141/2013, 127/2014, 62/2017, 96/2018 [31 October 2018]

[NN] Official Gazette of the Republic of Croatia (2019). *Obiteljski zakon*, 103/2015, 98/2019 [16 October 2019]

[NN] Official Gazette of the Republic of Croatia (2020). *Zakon o odgoju i obrazovanju u osnovnoj i srednjoj školi*, 87/2008, 86/2009, 92/2010, 105/2010, 90/2011, 5/2012, 16/2012, 86/2012, 126/2012, 94/2013, 1 52/2014, 07/2017, 68/2018, 98/2019, 64/2020 [30 May 2020]

[NN] Official Gazette of the Republic of Croatia (2021). *Zakon o autorskom pravu i srodnim pravima*, 111/2021 [22 October 2021]

Podbojec, A. & Mekovec, R. (2021) Zaštita podataka na visokim učilištima u Republici Hrvatskoj: preliminarno istraživanje. CroDiM, 4(1), 69-80.

Politiscope, (2021a). Analiza aplikacija za učenje na daljinu iz perspektive zaštite privatnosti djece. Retrieved from: https://acfcroatia.hr/wp-content/uploads/2021/04/Analiza_nastava-apps.pdf

Obrazovanje/NastavaNaDaljinu//Akcijski%20plan%20za%20provedbu%20nastave%20na%20daljinu%20-%20Model%20nastave%20na%20daljinu.pdf

[NN] Official Gazette of the Republic of Croatia (2014). *Zakon o informacijskoj sigurnosti*, 79/207 [30 July 2007]

[NN] Official Gazette of the Republic of Croatia (2008). *Državni pedagoški standard osnovnoškolskog sustava odgoja i obrazovanja,* 59/1990, 26/1993, 27/1993, 29/1994., 7/1996, 59/2001, 114/2001, 76/2005, 63/2008 [2 June 2008]

[NN] Official Gazette of the Republic of Croatia (2008). *Državni pedagoški standard srednjoškolskog sustava odgoja i obrazovanja,* 19/1992, 26/1993, 27/1993, 50/1995, 59/2001, 114/2001, 81/2005, 63/2008 [2 June 2008]

[NN] Official Gazette of the Republic of Croatia (2014). *Ustav Republike Hrvatske*, 56/1990, 135/1997, 8/1998, 113/2000, 124/2000, 28/2001, 41/2001, 55/2001, 76/2010, 85/2010, 5/2014 [15 January 2014]

[NN] Official Gazette of the Republic of Croatia (2014). *Strategija obrazovanja, znanosti i tehnologije*, 124/2014 [24 October 2014]

[NN] Official Gazette of the Republic of Croatia (2018). *Odluka o donošenju kurikuluma za nastavni predmet Informatike za osnovne škole i gimnazije u Republici Hrvatskoj*, 22/2018 [6 March 2018]

[NN] Official Gazette of the Republic of Croatia (2018). *Zakon o provedbi Opće uredbe o zaštiti podataka*, 42/2018 [9 May 2018]

[NN] Official Gazette of the Republic of Croatia (2018). *Zakon o autorskom pravu i srodnim pravima*, 167/2003, 79/2007, 80/2011, 141/2013, 127/2014, 62/2017, 96/2018 [31 October 2018]

[NN] Official Gazette of the Republic of Croatia (2019). *Obiteljski zakon*, 103/2015, 98/2019 [16 October 2019]

[NN] Official Gazette of the Republic of Croatia (2020). *Zakon o odgoju i obrazovanju u osnovnoj i srednjoj školi*, 87/2008, 86/2009, 92/2010, 105/2010, 90/2011, 5/2012, 16/2012, 86/2012, 126/2012, 94/2013, 1 52/2014, 07/2017, 68/2018, 98/2019, 64/2020 [30 May 2020]

[NN] Official Gazette of the Republic of Croatia (2021). *Zakon o autorskom pravu i srodnim pravima*, 111/2021 [22 October 2021]

Podbojec, A. & Mekovec, R. (2021) Zaštita podataka na visokim učilištima u Republici

Politiscope, (2021b). Analiza praksi, prikupljanja i obrada podataka za vrijeme Covid pandemije – slučaj Hrvatske. Retrieved from: https://acfcroatia.hr/wp-content/uploads/2021/04/Analiza_Covid.pdf

Skov, A. (2016) What is Digital Competence? Center for Digital Dannelse. Published online March 2016. Retrieved from: https://digital-competence.eu/dc/front/what-is-digital-competence/

Yong, Jianming. (2007). Digital Identity Design and Privacy Preservation for e-Learning.

Hrvatskoj: preliminarno istraživanje. CroDiM, 4(1), 69-80.

Politiscope, (2021a). Analiza aplikacija za učenje na daljinu iz perspektive zaštite privatnosti djece. Retrieved from: https://acfcroatia.hr/wp-content/uploads/2021/04/Analiza_nastava-apps.pdf

Politiscope, (2021b). Analiza praksi, prikupljanja i obrada podataka za vrijeme Covid pandemije – slučaj Hrvatske. Retrieved from: https://acfcroatia.hr/wp-content/uploads/2021/04/Analiza_Covid.pdf

Skov, A. (2016) What is Digital Competence? Center for Digital Dannelse. Published online March 2016. Retrieved from: https://digital-competence.eu/dc/front/what-is-digital-competence/

Yong, Jianming. (2007). Digital Identity Design and Privacy Preservation for e-Learning.

Opportunities for a Digital Transformation of the Global Supply Chains: Case Study of the Suez Canal Blockade

Ema Novak, Ana Kutnjak, Larisa Hrustek
Faculty of Organization and Informatics,
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{ema.novak, ana.kutnjak, larisa.hrustek}@foi.unizg.hr

Abstract. *Supply chains (SC) are the heart of global trade. Shortening the routing distance, lower transport and distribution costs, the global connection between countries and partners, as well as worldwide customer supply, are just a few of the major advantages of the global supply chains. However, like every other global process, supply chains also face obstacles in their continuity. One of the recent, bigger disruptions (DISR) happened when the Ever Given container vessel got stranded in the Suez Canal causing a global blockade of the customer supply. A question arises if such an incident would have been prevented if there was less impact of a human hand and greater use of digital technology in the supply process itself. This paper aims to show how the usage of digital technologies can improve the quality and continuity of the supply chains, prevent disruptions such as the Suez Canal blockade and prove that digital transformation is the future of a high-quality distribution process.*

Keywords. Supply chains, Suez Canal blockade, digital transformation, digital technologies

1 Introduction

In the year 2020, the world faced a major crisis when the first patient of the COVID-19 virus was confirmed. The fast spread of the virus closed the world and stopped human activities such as education, healthcare, tourism, travel, production, trade and transportation. The COVID-19 measures had a significant impact on the global supply chains since the actors were connected all over the world.

Due to closures and disruptions in the supply chains, international trade has weakened but was still organized and performed. In March 2021, the supply chains faced another big crisis when the Suez Canal, one of the world's most important trading paths, got blocked by the container vessel Ever Given. This incident caused delivery delays, interruption of supply, production losses and higher costs of transportation, products and services.

Regarding such events, it is clear that supply chain disruptions cause a loss of up to half of an annual profit for an average company (Chupanova et al., 2021). Because such events are unpredictable and sudden, companies have to create, implement and use digital solutions to mitigate their impact on the business.

Digital technologies are being adopted in modern supply chains for product traceability, enabling data sharing amongst trading partners, quick availability of product data and end-to-end visibility of products (Syed et al., 2022). The most used technologies fall into the category of Internet of Things technologies which help to collect data and make it understandable thus minimizing the problems that may arise in supply chains. Looking back on recent events, it is important to support the logistics processes with the Internet of Things enabled technologies, especially in supply chains that are vulnerable to disruptions (Kazancoglu et al., 2022).

This paper aims to show how digital technologies can strengthen the processes and components of the supply chains so that they are more resistant to disruptions. For this purpose, the authors defined the following three research questions (RQ): (RQ1) Which disruptions occur in the supply chains?; (RQ2) Which technologies are most prevalent in the supply chains?; (RQ3) How can technologies strengthen the supply chains' resilience to disruptions?

Related to the mentioned, the paper consists of seven chapters. Apart from the introduction, the second chapter shows the methodology used to get the relevant information, the third chapter counts and explains the main disruptions that occur in the supply chains and the fourth chapter provides the recognized digital technologies and their impact on the disruptions in the supply chains. The following fifth chapter deals with the case study of the Ever Given vessel, together with the causes, impacts on the supply chains and technology solutions, while the sixth chapter is the discussion about the used and potential new technology in the supply chains. In the last, seventh, chapter of the paper, the authors state relevant conclusions, practical implications and research limitations.

2 Methodology

In order for the authors to get answers on the, initially set, three research questions, it was necessary to search relevant databases with the purpose of collecting secondary data for literature analysis. The research query was made in the Scopus and Web of Science (WoS) databases, and it included the keywords below: *("supply chain management" OR "supply chain") AND ("digital technologies" OR "digital transformation") AND ("disruption" OR "accidents")*. When defining a research query in WoS, no limitations were set, so all fields were covered by the search and the papers were classified according to the highest citation. Furthermore, when setting the research query in Scopus, the authors put a limitation on title-abstract-keywords, because without limitation, there was a larger set of papers (more than two thousand papers), which were not relevant to this research. As in WoS, the papers were sorted according to the highest citations. The search resulted in a total of 65 papers in Scopus and 38 papers in WoS. After the analysis of each paper, those unavailable and overlapping in the mentioned two databases were removed and ultimately, the literature review was based on the analysis of a total of 47 papers (43 papers from the Scopus and 4 from the WoS). With the detailed analysis of the mentioned papers, the authors tried to determine the answers to the RQ and thus contribute to the field of supply chain management (SCM), while looking at the impact of digital technologies and their significance in mitigating disruptions.

3 Disruptions in Supply Chain Management

In the last few years, incidents in the environment have shown a rather large impact on the usual functioning of the supply chain business. The common name of such incidents is disruptions and their unique features make a serious impact on processes, which leads to complexity and their spread through multiple stages in the supply chain, called the ripple effect (Ivanov & Dolgui, 2019; Lohmer et al., 2020; Paul & Chowdhury, 2020). In this research, disruptions in the supply chain are categorized into two key categories as follows in Figure 1: *external disruptions* and *internal disruptions*. By classification of disruptions, the authors answered the first research question (RQ1).

External disruptions are related to factors that are not under the organization's control and appear suddenly in the supply chain. Natural disruption includes incidents such as fire, flood, earthquakes or tempests (Ivanov et al., 2018, 2019; Ivanov & Dolgui, 2019; Weber-Snyman, 2021). Economic and financial disruption presents devaluation and revaluation risk, exchange rate risk, crisis, stock market crash, inflation and stagflation, while legal disruption can be non-compliance with regulations, protests or political instability (Ali & Govindan, 2021; Moretto & Caniato, 2021). The external disruptions with the most significant impact are demand disruptions which are customer bankruptcy, imbalanced demand, price fluctuation, changing customer requirements as well as low data visibility about customers (Das et al., 2019; Ivanov et al., 2018; Paul & Chowdhury, 2020), and supply disruptions present supplier bankruptcy, price fluctuations, unstable quality and quantity of inputs (Ivanov et al., 2018), imbalanced supply, delivery risk and complex mix of product specification (Paul & Chowdhury, 2020). A pandemic disruption, whose impact has left an indelible mark on the entire supply chain, is Covid-19 and one of the most famous in the last twenty years was the Ebola pandemic.

Internal disruptions are related to factors that are under the organization's control and they can be reduced or prevented. Information disruption presents a communication breakdown, while time disruption includes delays in processes and time bottlenecks. Inability to adapt to digital solutions presents digital and technology disruption. Logistics and transportation disruption imply risks such as delivery delays, packaging problems or automated vehicle accidents (Ali et al., 2021). Capacity and process disruption represents operational, infrastructural and human capabilities within the organization (Das et al., 2019).

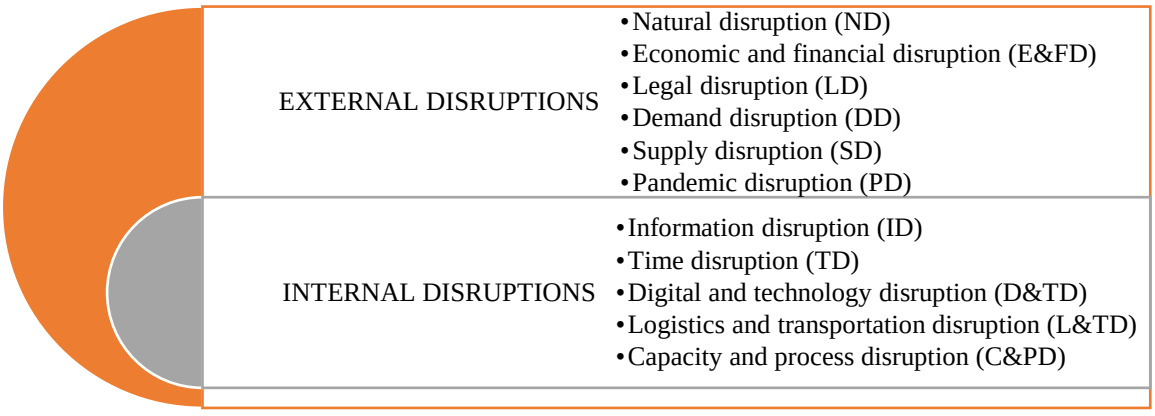


Figure 1. Identified disruptions in the supply chain

4 Digital Transformation and Supply Chain Management

The investigation process consisted of setting three research questions. In the first phase of the research, through literature analysis, the answer to the second research question (RQ2) was obtained: Which technologies are most prevalent in the supply chain? In the results in Table 1, the technologies represented in supply chain management are identified. The technologies are classified according to the authors' previous research (Hrustek et al., 2022). They are Big Data Analytics (BDA), Robotics and Autonomous Technology (R&AT), Internet of Things (IoT), Artificial Intelligence (AI), Blockchain (BC), Digital Twin (DT), Virtual and Augmented Reality (VR&AR), Cloud Computing (CC), Digital Platforms (DP), Social Media (SM), Cyber Physical System (CPS) and Mobile Technology (MT). Each of the mentioned technologies was recognized in the literature review and contributes to the efficient management of the supply chain.

Table 1. Recognized technologies per authors based on literature review

TECH	AUTHORS
BDA	(Afanasyev et al., 2019; Ali et al., 2021; Ali & Govindan, 2021; Butt, 2021; Chaouni Benabdellah et al., 2021; Chatterjee et al., 2021; Das et al., 2019; Haleem et al., 2022; Hopkins, 2021; Irfan et al., 2022; Ivanov, 2021; Ivanov et al., 2018, 2019; Kapoor et al., 2021; Martinelli et al., 2021; Menon & Shah, 2020; Obade & Gaya, 2021; Pyun & Rha, 2021; Rajesh, 2016; Reyes et al., 2021; Saleh & Awny, 2020; Shah et al., 2021; Tortorella et al., 2021)
R& AT	(Ali et al., 2021; Ali & Govindan, 2021; Butt, 2021; Chatterjee et al., 2021; Cozmiuc & Petrisor, 2018b; Fleury & Fleury, 2020; Hald & Coslugeanu, 2021; Haleem et al., 2022; Hopkins, 2021; Ivanov et al., 2018; Kapoor et al., 2021; Mavlutova et al., 2021; Obade & Gaya, 2021; Pyun & Rha, 2021; Reyes et al., 2021; Sajjad, 2021; Shah et al., 2021; Shaw & Chisholm, 2020; Sun & Zhao, 2018)
IoT	(Afanasyev et al., 2019; Ali et al., 2021; Ali & Govindan, 2021; Butt, 2021; Chaouni Benabdellah et al., 2021; Chatterjee et al., 2021; Cozmiuc & Petrisor, 2018b; Das et al., 2019; Dev et al., 2021; Fleury & Fleury, 2020; Hald & Coslugeanu, 2021; Haleem et al., 2022; Hopkins, 2021; Irfan et al., 2022; Ivanov et al., 2018, 2019; Ivanov, 2021; Kapoor et al., 2021; Kazancoglu et al., 2022; Mavlutova et al., 2021; Obade & Gaya, 2021; Pyun & Rha, 2021; Quayson et al., 2020; Rapaccini et al., 2020; Reyes et al., 2021; Sajjad, 2021; Santos et al., 2021; Sermuksnyte-Alesiuniene, 2021; Shah et al., 2021; Shaw & Chisholm, 2020; Sun & Zhao, 2018; Tortorella et al., 2021; Varriale et al., 2021),
AI	(Afanasyev et al., 2019; Belhadi et al., 2021; Chaouni Benabdellah et al., 2021; Chatterjee et al., 2021; Cozmiuc & Petrisor, 2018b; Hald & Coslugeanu, 2021; Haleem et al., 2022; Hopkins, 2021; Irfan et al., 2022; Ivanov et al., 2019; Ivanov, 2021; Kapoor et al., 2021; Mavlutova et al., 2021; Moretto & Caniato, 2021; Obade & Gaya, 2021; Reyes et al., 2021; Sajjad, 2021; Saleh & Awny, 2020; Serrano et al., 2021; Shaw & Chisholm, 2020; Sun & Zhao, 2018; Tortorella et al., 2021),
BC	(Afanasyev et al., 2019; Ali & Govindan, 2021; Chaouni Benabdellah et al., 2021; Chatterjee et al., 2021; El Idrissi et al., 2021; Hald & Coslugeanu, 2021; Haleem et al., 2022; Hopkins, 2021; Irfan et al., 2022; Ivanov, 2021; Ivanov et al., 2018, 2019; Lohmer et al., 2020; Mavlutova et al., 2021; Menon & Shah, 2020; Moretto & Caniato, 2021; Quayson et al., 2020; Rauniyar et al., 2022; Reyes et al., 2021; Sajjad, 2021; Saleh & Awny, 2020; Santos et al., 2021; Taghizadeh & Taghizadeh, 2021; Tortorella et al., 2021; Varriale et al., 2021; Zhang, 2021),
DT	(Afanasyev et al., 2019; Ivanov et al., 2019; Reyes et al., 2021; Serrano et al., 2021; Tortorella et al., 2021)
VR& AR	(Chaouni Benabdellah et al., 2021; Hald & Coslugeanu, 2021; Haleem et al., 2022; Hopkins, 2021; Pyun & Rha, 2021; Reyes et al., 2021; Sun & Zhao, 2018; Tortorella et al., 2021)
CC	(Ali et al., 2021; Ali & Govindan, 2021; Chaouni Benabdellah et al., 2021; Dev et al., 2021; Haleem et al., 2022; Martinelli et al., 2021; Menon & Shah, 2020; Reyes et al., 2021; Tortorella et al., 2021)
DP	(Irfan et al., 2022; Ivanov, 2021; Moretto & Caniato, 2021)
SM	(Hu, 2022; Sun & Zhao, 2018)
CPS	(Chaouni Benabdellah et al., 2021; Cozmiuc & Petrisor, 2018a; Tortorella et al., 2021)
MT	(Chatterjee et al., 2021; Martinelli et al., 2021)

The emergence of technologies, based on the literature review, is shown in Figure 2. The figure shows that there are leading technologies in the supply chain, namely the Internet of Things, Blockchain, Big Data Analytics, Artificial Intelligence, as well as Robotics and Autonomous Technology. The remaining technologies, such as Cloud Computing, Virtual and Augmented Reality, Digital Twin, Digital Platforms, Cyber Physical System, Mobile Technology and Social Media, appear less often in the analysed literature, which does not necessarily mean that they are not used to a greater extent, and their lower representation may be the result of the research query posed in the research.

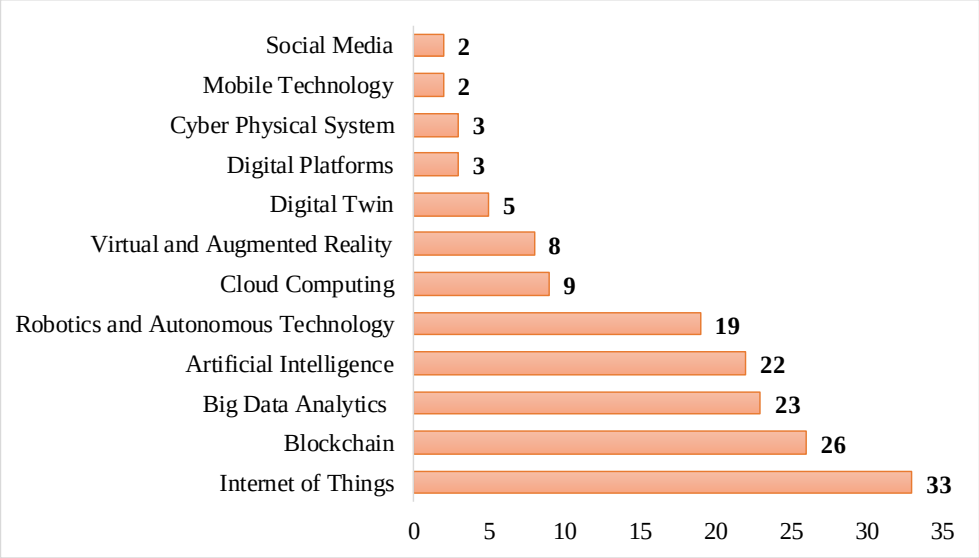


Figure 2. Emergence of technologies in the supply chain

Table two presents the answer to the third research question (RQ3) How can technologies strengthen the supply chains' resilience to disruptions? The table shows how each technology contributes to supply chain management and to which disruptions technologies can provide an adequate response and enable supply chain resilience. Also, it provides a connection between technologies and their impact on mitigating risk elements (RE) such as hazard (H), exposure (E) and vulnerability (V) (Cardona et al., 2012).

Table 2. Technologies' impact on disruptions in the supply chain

TECH	CONTRIBUTION IN SCM	DISR (RE)
SM	LinkedIn, Facebook, Instagram and YouTube are used to share information about products, services and for spreading electronic word-of-mouth communication. They strengthen cooperation between organizations in SC, focusing on information sharing and joint problem solving (Hu, 2022; Sun & Zhao, 2018).	DD (E); SD (E); PD (H); D&TD (V); L&TD (E/V); C&PD (V)
MT	Ease of installation and the ubiquity of mobile technology leads to the constant availability of information in the SC. The aforementioned technology is integrated with other systems within organizations in SC, which leads to monitoring, updating, comparison and access to information in SC from anywhere at any time (Chatterjee et al., 2021; Martinelli et al., 2021).	DD (E); SD (E); PD (H); D&TD (V); L&TD (E/V); C&PD (V)
CPS	They represent a system that unites computing entities with the physical world that are in an intensive relationship with its ongoing processes, providing and using, at the same time, data access and data processing. In the SC, such systems rely on digital technology integrated with production machinery, storage systems and transportation technology, and the characteristic of the relationship is the autonomous and intelligent exchange of information (Cozmiuc & Petrisor, 2018a; Ivanov et al., 2019; Tortorella et al., 2021).	DD (E); SD (E); ID (V); D&TD (V); L&TD (E/V); C&PD (V)
DP	The platforms represent a base for cooperation and connection of organizations with business partners in the SC. Such technology enables the SC to respond quickly to market changes and fluctuations, thus ensuring the long-term success of the SC. Optimizing the planning of production activities, the exchange of important business data that enables monitoring of current conditions, but also the simulation of future demand, support of business processes and transparency of operations, are the key advantages obtained from the use of such technologies (Das et al., 2019). Essentially, SC visibility is enabled by digital platforms (Ivanov et al., 2019), and they are crucial for the speed of data collection, analysis and knowledge creation (Irfan et al., 2022; Ivanov, 2021).	DD (E); SD (E); PD (H); ID (V); D&TD (V); L&TD (E/V); C&PD (V)

DT	A model that represents the network state of any moment in time and allows complete end-to-end SC visibility to improve resilience. It is used for complex analysis of SC risks, the development of contingency plans and efficient operational management. It supports the decision making process on the basis of data by mirroring the physical SC: transportation, inventory, demand and capacity data. Also, the simulation of the DT helps to forecast possible disruption propagation and quantify its impact and it can be transferred to the ERP system of business intelligence tool to analyse the performance impact of the disruption (Ivanov et al., 2019). DT can enable simulation and optimization methods in combination with Machine Learning, promote faster actions to reduce lead times, improve forecast accuracy, as well as promote demand forecasting, aggregate and inventory planning of the SC to be more reliable, efficient and quick (Serrano et al., 2021).	DD (E); PD (H); L&TD (E/V); C&PD (V)
VR& AR	VR enables interaction in the environment with a realistic effect in three-dimensional situations, using equipment with sensors that help to better understand the situation and make decisions, by providing a visual experience. AR describes the ability to integrate virtual 3D objects into the real environment, which helps to realize a real experience. SC recognizes the potential of this technology in training/education, product visualization and improved inventory management, while research and development are focused on the ability to create virtual work environments that enable meaningful interactions between colleagues, partners and customers, regardless of their physical location (Hopkins, 2021; Pyun & Rha, 2021).	ND (H); E&FD (H); LD (E); SD (E); ID (V); TD (E)
CC	CC refers to the use of a common master server to which all devices are connected, which allows the collection and creation of a common set of data for further processing. The joint application of technologies and CC enables connectivity of devices, data and information sharing, which requires minimal effort and limited investment. It offers optimization, predictive maintenance, real-time monitoring and customer profile data to support process adaptation sales and services (Ali et al., 2021; Ali & Govindan, 2021).	DD (E); SD (E); ID (V); TD (E); C&PD (V)
R& AT	Collaborative robots read product information from sensors and flexibly decide where to forward the identified product. The development of robots is focused on their independent learning, automatic execution of various actions and optimization of commoditized tasks, while the collected data gives the possibility of creating an ideal strategy for implementing the process (Ivanov et al., 2018; Sun & Zhao, 2018).	SD (E); ID (V); TD (E); C&PD (V)
AI	AI exploits data, expands it and enables research and analysis of it. It leads to the generation of new ideas within SC, evaluates innovative solutions, supports organizations in problem analysis and leads to rapid prototyping. SC, using the aforementioned technology, can easily process information, recognize key data for uninterrupted SC operations, and can learn from data with the aim of mitigating uncertain market conditions and ensuring SC resilience (Belhadi et al., 2021). Furthermore, AI improves SC efficiency by clearly understanding the interaction and contribution of each SC participant (Chaouni Benabdellah et al., 2021). In this way, AI provides greater precision, accuracy and speed of data collection (Sun & Zhao, 2018).	PD (H); ID (V); TD (E); C&PD (V)
BDA	It is based on the extraction of knowledge from a huge amount of data which facilitates data-driven decision-making, and in SCM it covers a wide scope of processes such as the formation of price optimization models, routing optimization, real-time traffic monitoring and proactive security management. Highlighted areas where BDA can contribute to the development of SC in the near future, include quality control in manufacturing, dynamic vehicle routing, in-transit inventory management in logistics/transportation, order picking and warehouse inventory control systems (Das et al., 2019; Haleem et al., 2022; Ivanov et al., 2018).	DD (E); SD (E); ID (V); TD (E); L&TD (E/V); C&PD (V)
BC	BC allows new solutions and creates rules to make the SC more intelligent, automated and efficient. BC is used to reduce delivery times, fasten the monitoring of goods and reduce of potential losses due to human error and unnecessary bureaucratic activities. It can improve productivity, reduce the time spent controlling processes and increase competitiveness in B2B organizations.	E&FD (H); DD (E); SD (E); ID (V); TD (E); D&TD (V)

	A BC based SC includes traceability (locates real-time movements of goods in the SC), immutability (enables secure transaction validation, purity verification, and authenticity), compliance (used to create applications that address compliance and responsibility) and sustainability (reduces product retraction, charges carbon tax, cultivates behaviour and improves the efficiency of emission trading blueprint) (Zhang, 2021). BC enhances transparency and visibility through the network, providing a secure connection between various databases and systems (Varriale et al., 2021).	
IoT	Additive manufacturing is the process of making three-dimensional products. Its application in SC is visible in the production of spare parts in logistics, which contributes to increased production flexibility, shorter delivery times, increased product individualization and reduced inventory. The tracking and tracing systems aim at the timely identification of disruptions, the analysis of deviations and the danger of disruptions that have occurred or may be caused, as well as the elaboration of control actions for recovery in the SC. Combined with radio-frequency identification and mobile devices, they are used to provide up-to-date information on process performance (Das et al., 2019; Ivanov et al., 2018).	ND (H); DD (E); SD (E); ID (V); L&TD (E/V); C&PD (V)

In the continuation of the paper, the most represented technologies and their impact on mitigating disruptions will be presented. The Internet of Things can be an adequate response to natural disruption, demand disruption, supply disruption, information disruption, logistics and transportation disruption, as well as capacity and process disruption. Economic and financial disruption, demand disruption, supply disruption, information disruption, time disruption, digital and technological disruption can be managed by Blockchain. Big Data Analytics finds a functional application in solving the following disruptions – demand disruption, supply disruption, information disruption, logistics and transportation disruption, as well as capacity and process disruption. Furthermore, the application of Artificial Intelligence mitigates the effects of pandemic disruption, information disruption, time disruption, capacity and process disruption. Also, supply disruption, information disruption, time disruption, capacity and process disruption, are covered through Robotics and Autonomous Technology.

In the next chapter, a case study of Ever Given, the largest overseas container vessel, will be presented, with an emphasis on technologies that could have prevented or mitigated the consequences of the incident known as the Suez Canal Blockade.

5 Case Study of the Suez Canal Blockade

World waterways play an important role in the global trade and shipping industry. One of the biggest and most important artificially dug waterways is the Suez Canal in Egypt, which divides the African from the Asian continent and connects the Mediterranean Sea with the Red Sea. The Canal itself is about 120 miles long and 105 meters wide, with a usage capacity of up to 19.000 ships with a net tonnage of 1.17 billion tons. All this indicates that the Suez Canal plays a crucial role in connecting the eastern and western world, as well as enabling the transport of natural gas, cargo and oil (Lee & Wong, 2021).

On the 23rd March 2021, a large container vessel named Ever Given sailed from Malaysia to the Netherlands through the Canal. The 400 meters long, weighing 20.000 tons and capable to carry around 20.000 containers, the Ever Given is one of the largest cargo vessels in the world. Although it has crossed the Suez Canal many times before, the huge vessel got stuck at the southern end of it, causing a blockade, as well as major and critical disruptions in the global supply chains.

Some of the problems that have arisen are the sudden stops of maritime transport for other vessels, which then had to find alternate routes to deliver the cargo. One of the alternative routes was the Cape of Good Hope in Africa which added several days of travel and millions of dollars in costs for services. Also, delays, congestions at terminal berths and shipyards increased the bottleneck effect in the logistics chain (How Ever Given Has Shaped the Future for Shipping Canals - Ship Technology Global, 2021). Over 300 ships, container vessels and bulk carriers were unable to use the Canal for six to seven days and the cost of disturbing the global markets was around billions of dollars (Coles, 2021). Regarding other consequences, the producers in China faced a shortage of empty containers to ship their goods, manufacturers around Europe and USA did not receive goods, components and materials for their production processes (Smet, 2021), the insurance companies suffered losses since they had to compensate victims (companies, customers, vessels) and law firms were dealing with expensive lawsuits regarding cargo interests (Lee & Wong, 2021).

Over the past months, many opinions have arisen about the reasons for the incident. One of the most likely reasons for the blockade is the weather conditions, e.g. wind and sand storms because sudden changes in the wind speed and direction make it difficult to manoeuvre vessels (Nikkanen, 2021).

Another reason for the crisis is the size of the container vessels because they are hard to shift once stuck (McLoughlin, 2021). Authors have also considered human errors, cyberattacks and lack of appropriate digital technology as possible causes of the blockade (Lee & Wong, 2021; Moulds, 2021). From the above, the authors identified three disruptions: natural, capacity and process, as well as digital and technology.

The use of outdated digital technologies could have had a significant impact on the incident. Standard used technologies in today's vessels are divided into four groups: IT bridge systems (radars, GPS, VDRs, ECDIS, AIS), communication systems (RF communication, satellite communication, integrated communication system, WLANs, voice over Internet Protocol), cargo systems (water ingress alarm, level indication, CCTV network, container- sensing system, on-board loading computers) and operational tech systems (power management, engine governor, emergency response) (Chan et al., 2021). Also, modern technologies such as Machine Learning (which is used for scheduling at ports and optimization of shipping patterns) and Blockchain (Coles, 2021) are used to control and manage the shipping process.

Although those technologies provide adequate support during the transportation process, they are still insufficient when crossing critical points such as the Suez Canal. In those situations, the use of modern, detailed and precise technologies would assure a secure and efficient passage of the vessels. Regarding the mentioned, some of the technologies which should be integrated into the logistics process are Artificial Intelligence (with wider use of Machine Learning), Digital Twins and Big Data Analytics. These technologies were already identified in Table 2, and can support functionalities such as weather quality measurement, simulation modelling, lean manufacturing and product optimization, container automation, as well as diagnostic tools and risk-based scheduling.

6 Discussion

Although IoT, BC, BDA, AI, and R&AT are recognized as the most represented, it is evident from the results that the remaining technologies for which less representation was determined, in accordance with the research query (SM, MT, DP, etc.), reduce the impact of the disruption on the SC. The use of specific technologies is related to the digital maturity of organizations, which implies their focus on the ability to apply technologies in a collaborative network, while dealing with disruptions. Problems caused by disruptions are reflected on all SC stakeholders through a ripple effect, and in order to keep it under control, it is necessary to transfer the digital culture to the entire supply chain. Finally, if digital technologies were more prevalent in maritime transport, incidents such as the blockade of the Suez Canal could be prevented on time.

7 Conclusion

Unpredictable disasters, such as health crises or wars, cause disruptions in supply chain management, and disruptions need to be responded to as early as possible. Organizations react differently to disruptions in supply chains, and digital technologies are of great help in this. Research results show that the opportunity to respond to disruptions is recognized in different digital technologies such as the Internet of Things, Blockchain, Big Data Analytics, Artificial Intelligence, Robotics and Autonomous Technology. Each of the mentioned technologies can provide resistance to disruptions based on their functionalities.

The importance of technologies is recognized in the analysed case study of the Ever Given vessel, but also in the pandemic crisis. In such cases, digital technology can help to predict the possible negative outcomes and reduce, as well as prevent their impact on the organization's business and the resilience of the SC.

This article provides several practical implications. First, a list of disruptions was identified and shows the organizations' possible threats for which they have to be prepared. Second, the authors identified which technologies can mitigate the negative impacts of the disruptions. Third, and last, the case study of the Ever Given vessel shows that the transport process of every supply chain is exposed to critical points such as the Suez Canal and digital technologies might be useful support in maritime transport.

In the very end, the authors agree there are some limitations of this research. Such defined research query has maybe excluded some relevant papers, whereby the frequency of technologies used in SC could be differently distributed. Also, the authors' recommendations are limited only to maritime transport since it was the main subject of this paper. Finally, there could be a different technology classification than the one mentioned in this paper, so the final results could be different from the ones the authors recognized.

References

Afanasyev, V. Y., Lyubimova, N. G., Ukolov, V. F., & Shayakhmetov, S. R. (2019). Digitalization of Energy Manufacture: Infrastructure, Supply Chain Strategy and Communication. *International Journal of Supply Chain Management*, *8*(4), 601–609.

Ali, I., Arslan, A., Khan, Z., & Tarba, S. Y. (2021). The Role of Industry 4.0 Technologies in Mitigating Supply Chain Disruption: Empirical Evidence From the Australian Food Processing Industry. *IEEE Transactions on Engineering Management*, 1–11. https://doi.org/10.1109/TEM.2021.3088518

Ali, I., & Govindan, K. (2021). Extenuating operational risks through digital transformation of agri-food supply chains. *Production Planning & Control*, *0*(0), 1–13. https://doi.org/10.1080/09537287.2021.1988177

Belhadi, A., Mani, V., Kamble, S. S., Khan, S. A. R., & Verma, S. (2021). Artificial intelligence-driven innovation for enhancing supply chain resilience and performance under the effect of supply chain dynamism: An empirical investigation. *Annals of Operations Research*. https://doi.org/10.1007/s10479-021-03956-x

Butt, A. S. (2021). Mitigating the effects of COVID-19: An exploratory case study of the countermeasures taken by the manufacturing industry. *Journal of Business & Industrial Marketing*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/JBIM-04-2021-0236

Cardona, O. D., Aalst, M. K. V., Birkmann, J., Fordham, M., Gregor, G. M., Rosa, P., Pulwarty, R. S., Schipper, E. L. F., Sinh, B. T., Décamps, H., Keim, M., Davis, I., Ebi, K. L., Lavell, A., Mechler, R., Murray, V., Pelling, M., Pohl, J., Smith, A. O., & Thomalla, F. (2012). Determinants of risk: Exposure and vulnerability. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change*, 65–108. https://doi.org/10.1017/CBO9781139177245.005

Chan, S., Yehuda, E., Schaefer, R., Schneuwly, A., Zicherman, S., Deutscher, S., & Klier, O. (2021, August 23). *Navigating Rising Cyber Risks in Transportation and Logistics*. BCG Global. https://www.bcg.com/publications/2021/navigating-rising-cyber-risks-in-transportation-and-logistics

Chaouni Benabdellah, A., Zekhnini, K., & Cherrafi, A. (2021). Sustainable and Resilience Improvement Through the Design for Circular Digital Supply Chain. In A. Dolgui, A. Bernard, D. Lemoine, G. von Cieminski, & D. Romero (Eds.), *Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems* (pp. 550–559). Springer International Publishing. https://doi.org/10.1007/978-3-030-85910-7_58

Chatterjee, S. R., Shukla, V. K., Wanganoo, L., & Dubey, S. (2021). Transforming Supply chain Management through Industry 4.0. *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 1–6. https://doi.org/10.1109/ICRITO51393.2021.9596392

Chupanova, K., Otrokov, O., Mosina, N., Zharov, A., & Garnik, S. (2021). Supply Chain Management Concept and Digital Economy: Digital Supply Chain Technological Innovation. *Indian Journal of Economics and Development*, 928–933. https://doi.org/10.35716/IJED/21272

Coles, T. (2021, May 22). *Machine Learning Automation Couldn't Keep the Suez Canal Unstuck*. ITPro Today: IT News, How-Tos, Trends, Case Studies, Career Tips, More. https://www.itprotoday.com/machine-learning/machine-learning-automation-couldnt-keep-suez-canal-unstuck

Cozmiuc, D., & Petrisor, I. (2018a). Industrie 4.0 by Siemens: Steps Made Next. *Journal of Cases on Information Technology*, *20*, 31–45. https://doi.org/10.4018/JCIT.2018010103

Cozmiuc, D., & Petrisor, I. (2018b). Industrie 4.0 by Siemens: Steps Made Today. *Journal of Cases on Information Technology (JCIT)*, *20*(2), 30–48. https://doi.org/10.4018/JCIT.2018040103

Das, A., Gottlieb, S., & Ivanov, D. (2019). Managing Disruptions and the Ripple Effect in Digital Supply Chains: Empirical Case Studies. In D. Ivanov, A. Dolgui, & B. Sokolov (Eds.), *Handbook of Ripple Effects in the Supply Chain* (pp. 261–285). Springer International Publishing. https://doi.org/10.1007/978-3-030-14302-2_13

Dev, N. K., Shankar, R., Zacharia, Z. G., & Swami, S. (2021). Supply chain resilience for managing the ripple effect in Industry 4.0 for green product diffusion. *International Journal of Physical Distribution & Logistics Management*, *51*(8), 897–930. https://doi.org/10.1108/IJPDLM-04-2020-0120

El Idrissi, A. H., Larfaoui, F., Dhingra, M., Johnson, A., Pinto, J., & Sumption, K. (2021). Digital technologies and implications for Veterinary Services. *Revue Scientifique Et Technique (International Office of Epizootics)*, *40*(2), 455–468. https://doi.org/10.20506/rst.40.2.3237

Fleury, A., & Fleury, M. T. (2020). A reconfiguração das Cadeias Globais de Valor (global value chains) pós-pandemia. *Estudos Avançados*, *34*(100), 203–222. https://doi.org/10.1590/s0103-4014.2020.34100.013

Hald, K. S., & Coslugeanu, P. (2021). The preliminary supply chain lessons of the COVID-19 disruption—What is the role of digital technologies? *Operations Management Research*. https://doi.org/10.1007/s12063-021-00207-x

Haleem, A., Javaid, M., Singh, R. P., Rab, S., & Suman, R. (2022). Perspectives of cybersecurity for ameliorative Industry 4.0 era: A review-based framework. *Industrial Robot: The International Journal of Robotics Research and Application*, *49*(3), 582–597. https://doi.org/10.1108/IR-10-2021-0243

Hopkins, J. L. (2021). An investigation into emerging industry 4.0 technologies as drivers of supply chain innovation in Australia. *Computers in Industry*, *125*, 103323. https://doi.org/10.1016/j.compind.2020.103323

How Ever Given has shaped the future for shipping canals—Ship Technology Global | Issue 77 | July 2021. (n.d.). Retrieved July 8, 2022, from https://ship.nridigital.com/ship_jul21/ever_given_canal_future

Hrustek, L., Kutnjak, A., & Tomašek, L. (2022). Contribution of Digital Marketing Activities to the Sustainability of Smart Products and Services. *Proceedings of 11th International Scientific Symposium "Region, Entrepreneurship, Development,"* 281.

Hu, L. (2022). The PPE industry in Italy during COVID-19: Supply chain disruption and the adoption of digital and social media in B2B firms. *Journal of Business & Industrial Marketing*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/JBIM-01-2021-0005

Irfan, I., Sumbal, M. S. U. K., Khurshid, F., & Chan, F. T. S. (2022). Toward a resilient supply chain model: Critical role of knowledge management and dynamic capabilities. *Industrial Management & Data Systems*, *122*(5), 1153–1182. https://doi.org/10.1108/IMDS-06-2021-0356

Ivanov, D. (2021). Digital Supply Chain Management and Technology to Enhance Resilience by Building and Using End-to-End Visibility During the COVID-19 Pandemic. *IEEE Transactions on Engineering Management*, 1–11. https://doi.org/10.1109/TEM.2021.3095193

Ivanov, D., & Dolgui, A. (2019). Low-Certainty-Need (LCN) supply chains: A new perspective in managing disruption risks and resilience. *International Journal of Production Research*, *57*(15–16), 5119–5136. https://doi.org/10.1080/00207543.2018.1521025

Ivanov, D., Dolgui, A., Das, A., & Sokolov, B. (2019). Digital Supply Chain Twins: Managing the Ripple Effect, Resilience, and Disruption Risks by Data-Driven Optimization, Simulation, and Visibility. In *International Series in Operations Research and Management Science* (pp. 309–332). https://doi.org/10.1007/978-3-030-14302-2_15

Ivanov, D., Dolgui, A., & Sokolov, B. (2018). The impact of digital technology and Industry 4.0 on the ripple effect and supply chain risk analytics. *International Journal of Production Research*, *57*, 1–18. https://doi.org/10.1080/00207543.2018.1488086

Kapoor, K., Bigdeli, A. Z., Dwivedi, Y. K., & Raman, R. (2021). How is COVID-19 altering the manufacturing landscape? A literature review of imminent challenges and management interventions. *Annals of Operations Research*. https://doi.org/10.1007/s10479-021-04397-2

Kazancoglu, Y., Ozbiltekin-Pala, M., Sezer, M. D., Kumar, A., & Luthra, S. (2022). Circular dairy supply chain management through Internet of Things-enabled technologies. *Environmental Science and Pollution Research*. https://doi.org/10.1007/s11356-021-17697-8

Lee, J. M., & Wong, E. Y. (2021). *Suez Canal blockage: An analysis of legal impact, risks and liabilities to the global supply chain*. https://doi.org/10.1051/matecconf/202133901019

Lohmer, J., Bugert, N., & Lasch, R. (2020). Analysis of resilience strategies and ripple effect in blockchain-coordinated supply chains: An agent-based simulation study. *International Journal of Production Economics*, *228*, 107882. https://doi.org/10.1016/j.ijpe.2020.107882

Martinelli, E. M., Farioli, M. C., & Tunisini, A. (2021). New companies' DNA: The heritage of the past industrial revolutions in digital transformation. *Journal of Management and Governance*, *25*(4), 1079–1106. https://doi.org/10.1007/s10997-020-09539-5

Mavlutova, I., Volkova, T., Natrins, A., Spilbergs, A., Arefjevs, I., & Miahkykh, I. (2021). Financial Sector Transformation In The Era Of Digitalization. *Studies of Applied Economics*, *38*. https://doi.org/10.25115/eea.v38i4.4055

McLoughlin, B. (2021). *What the blocked Suez Canal can teach us about supply chain technology | HERE*. https://www.here.com/company/blog/suez-canal-supply-chain

Menon, S., & Shah, S. (2020). Growth of Digital Supply Chains for SME Transformation. *2020 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD)*, 1–6. https://doi.org/10.1109/ICTMOD49425.2020.9380603

Moretto, A., & Caniato, F. (2021). Can Supply Chain Finance help mitigate the financial disruption brought by Covid-19? *Journal of Purchasing and Supply Management*, *27*(4), 100713. https://doi.org/10.1016/j.pursup.2021.100713

Moulds, K. (2021, November 25). *The Suez Canal—Could a virtual incident cause the same problems?* Intercede. https://www.intercede.com/the-suez-canal/

Nikkanen, M. (2021). *The Ever Given and Suez Canal: Avoiding future crisis with weather observation data*. Vaisala. https://www.vaisala.com/en/blog/2021-06/ever-

given-and-suez-canal-avoiding-future-crisis-weather-observation-data

Obade, V. de P., & Gaya, C. (2021). Digital technology dilemma: On unlocking the soil quality index conundrum. *Bioresources and Bioprocessing*, *8*(1), 6. https://doi.org/10.1186/s40643-020-00359-x

Paul, S. K., & Chowdhury, P. (2020). A production recovery plan in manufacturing supply chains for a high-demand item during COVID-19. *International Journal of Physical Distribution & Logistics Management*, *51*(2), 104–125. https://doi.org/10.1108/IJPDLM-04-2020-0127

Pyun, J., & Rha, J. (2021). Review of Research on Digital Supply Chain Management Using Network Text Analysis. *Sustainability*, *13*, 9929. https://doi.org/10.3390/su13179929

Quayson, M., Bai, C., & Osei, V. (2020). Digital Inclusion for Resilient Post-COVID-19 Supply Chains: Smallholder Farmer Perspectives. *IEEE Engineering Management Review*, *48*(3), 104–110. https://doi.org/10.1109/EMR.2020.3006259

Rajesh, R. (2016). Forecasting supply chain resilience performance using grey prediction. *Electronic Commerce Research and Applications*, *20*, 42–58. https://doi.org/10.1016/j.elerap.2016.09.006

Rapaccini, M., Saccani, N., Kowalkowski, C., Paiola, M., & Adrodegari, F. (2020). *Navigating disruptive crises through service-led growth*. https://doi.org/10.1016/j.indmarman.2020.05.017

Rauniyar, K., Wu, X., Gupta, S., Modgil, S., & Lopes de Sousa Jabbour, A. B. (2022). Risk management of supply chains in the digital transformation era: Contribution and challenges of blockchain technology. *Industrial Management & Data Systems*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/IMDS-04-2021-0235

Reyes, J., Mula, J., & Díaz-Madroñero, M. (2021). Development of a conceptual model for lean supply chain planning in industry 4.0: Multidimensional analysis for operations management. *Production Planning & Control*, *0*(0), 1–16. https://doi.org/10.1080/09537287.2021.1993373

Sajjad, A. (2021). The COVID-19 pandemic, social sustainability and global supply chain resilience: A review. *Corporate Governance: The International Journal of Business in Society*, *21*(6), 1142–1154. https://doi.org/10.1108/CG-12-2020-0554

Saleh, A., & Awny, M. M. (2020). Digital transformation strategy framework. *Proceedings of the 29th International Conference of the International Association for Management of Technology, IAMOT*. https://www.proceedings.com/56210.html

Santos, H., Oliveira, A., Soares, L., Satis, A., & Santos, A. (2021). Information Security Assessment and Certification within Supply Chains. *ACM International Conference Proceeding Series*, 1–6. https://doi.org/10.1145/3465481.3470078

Sermuksnyte-Alesiuniene, K. (2021). *Increasing the effectiveness of food supply chain logistics through digital transformation*. Instituto Federal de Educação, Ciência e Tecnologia de São Paulo.

Serrano, J. C., Mula, J., & Poler, R. (2021). Smart Master Production Schedule for the Supply Chain: A Conceptual Framework. *Computers*, *10*, 156. https://doi.org/10.3390/computers10120156

Shah, S. R., See, H. H., & Samadi, S. (2021). Applicability of Digital Tracking System on Third Party Logistics (TPL) Services. *2021 IEEE International Conference on Technology Management, Operations and Decisions (ICTMOD)*, 1–6. https://doi.org/10.1109/ICTMOD52902.2021.9750416

Shaw, B., & Chisholm, O. (2020). Creeping Through the Backdoor: Disruption in Medicine and Health. *Frontiers in Pharmacology*, *11*, 818. https://doi.org/10.3389/fphar.2020.00818

Smet, R. D. (2021, April 20). *Suez Canal blockage: Evidence that Data Transformation initiatives are needed to reduce risk and improve operational visibility*. Schneider Electric Blog. https://blog.se.com/machine-and-process-management/2021/04/20/suez-canal-blockage-evidence-that-data-transformation-initiatives-are-needed-to-reduce-risk-and-improve-operational-visibility/

Sun, L. (Sarina), & Zhao, L. (2018). Technology disruptions: Exploring the changing roles of designers, makers, and users in the fashion industry. *International Journal of Fashion Design, Technology and Education*, *11*, 1–13. https://doi.org/10.1080/17543266.2018.1448462

Syed, N. F., Shah, S. W., Trujillo-Rasua, R., & Doss, R. (2022). Traceability in supply chains: A Cyber security analysis. *Computers & Security*, *112*, 102536. https://doi.org/10.1016/j.cose.2021.102536

Taghizadeh, E., & Taghizadeh, E. (2021). *The Impact of Digital Technology and Industry 4.0 on Enhancing Supply Chain Resilience*.

Tortorella, G., Fogliatto, F. S., Gao, S., & Chan, T.-K. (2021). Contributions of Industry 4.0 to supply chain resilience. *The International Journal of*

Logistics Management, *33*(2), 547–566. https://doi.org/10.1108/IJLM-12-2020-0494

Varriale, V., Cammarano, A., Michelino, F., & Caputo, M. (2021). Sustainable Supply Chains with Blockchain, IoT and RFID: A Simulation on Order Management. *Sustainability*, *13*(11), 6372. https://doi.org/10.3390/su13116372

Weber-Snyman, A. (2021). Responding to supply chain disruptions caused by the COVID-19 pandemic: A Black Swan event for omnichannel retailers. *Journal of Transport and Supply Chain Management*, *15*, 2310–8789. https://doi.org/10.4102/jtscm.v15i0.628

Zhang, H. (2021). Blockchain facilitates a resilient supply chain in steel manufacturing under covid-19: 22nd European Conference on Knowledge Management, ECKM 2021. *Proceedings of the 22nd European Conference on Knowledge Management, ECKM 2021*, 964–972. https://doi.org/10.34190/EKM.21.058

Possible Changes in Academic Teaching as a Consequence of the COVID-19 Pandemic: International Findings and a Case Study from Croatia

Goran Bubaš, Antonela Čižmešija
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{gbubas, acizmesi}@foi.unizg.hr

Abstract. *After the first lockdown at universities in March and April 2020 due to the onset of the COVID-19 pandemic numerous studies have been conducted to investigate the effects of the pandemic circumstances on teaching and learning at higher education institutions (HEIs), as well as on their students and teachers. This paper briefly outlines the results of international studies that were performed during the pandemic, as well as some findings from surveys performed at the national level in Croatia, and at the level of one HEI in Croatia. The main focus of this paper is on students' satisfaction with online and hybrid instruction, their evaluation of various replacements for on-site (face to face) teaching, their preference for on-site, online and hybrid (partly on-site, partly online) teaching, the obstacles/barriers to online teaching, as well as on its advantages during the COVID-19 pandemic. An empirical pilot study was performed with students enrolled in two university courses (N=70) that investigated in more detail the effects of hybrid instruction on students' satisfaction, preferences for online content delivery, flexibility in course attendance (on-site vs. online), and possible reasons for students' inclination toward online teaching and learning. The main finding of the pilot study was that students' preference for online delivery of instruction during the COVID-19 pandemic may have been predominantly grounded on practical and pragmatic reasons.*

Keywords. distance education, e-learning, COVID-19 pandemic, hybrid learning, flexible learning, pilot study, survey

1 Introduction

It was in the period from mid-March to April 1st, 2020 that most countries in the world imposed a national lockdown due to the onset of the COVID-19 pandemic of the SARS-CoV-2 virus, including school closures, as can be seen from the interactive chart map for April 1st, 2020 that is presented in *Figure 1* (source: Our World in Data, 2022; for more popular reminders of the first closure/lockdown period, see BBC, 2020; DW, 2020). According to the United Nations report published in August 2020 (UN, 2020), the COVID-19 pandemic "has caused the largest disruption of education in history", primarily because by mid-April 2020 as many as 94 per cent of learners in 200 countries worldwide were affected, including approximately 1.58 billion children and youth from pre-primary to higher education.

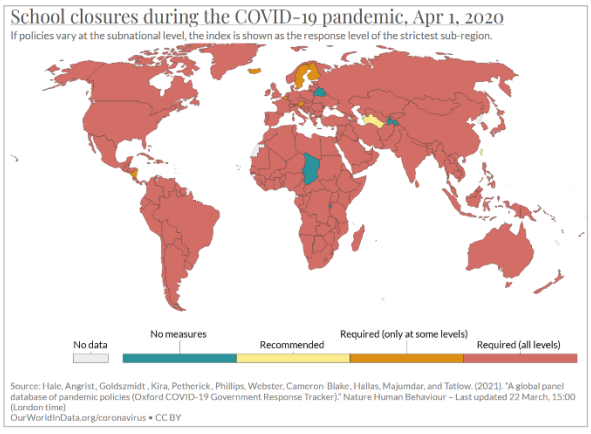


Figure 1. A global map of school closures during the COVID-19 pandemic on April 1st 2020

Shortly after the onset of the COVID-19 pandemic the expression "emergency remote teaching" became widespread in the education community to describe the sudden shift to fully online teaching and learning. In one of the most cited scholarly articles on e-learning (published on March 27th, 2020 in the *EDUCAUSE Review* and mentioned in more than 4000 publications by mid-June 2022 (according to *Google Scholar*), Hodges et al. (2020) state that "emergency remote teaching (ERT) is a temporary shift of instructional delivery to an alternate delivery mode due to crisis circumstances".

Bond et al. (2021) conducted a review of 282 primary empirical studies published between January 2020 and October 2020 on how teachers, students and academic institutions adapted to delivery of education in fully remote and online formats during the summer semester of the 2019/2020 academic year. They found that most of the research papers in that period were focused on (a) the experiences of students during the pandemic, (b) quality of online teaching and learning, (c) support to students by use of digital technology and (d) assessment issues.

Another review study was performed by Stewart (2021) that included 38 studies specifically about *emergency remote teaching* (ERT) that were published in the same period from January 2020 to October 2020. This review study was more focused on ERT and indicated that the major related empirical research themes in that period were (a) the diverse ERT experiences, (b) digital divide and vast educational / socioeconomic inequalities, (c) commonly experienced ERT problems, and (d) frequently performed adjustments in response to ERT.

1.1 Large-scale surveys at the beginning and during the COVID-19 pandemic

Several large-scale surveys were performed in Europe, USA and globally in the period after the onset of the COVID-19 pandemic and during the summer semester of the 2019/2020 academic year. A study conducted in 41 European country (Doolan et al., 2021) found that 93% of the surveyed higher education students (N=9,180) had their on-site (face-to-face) classes cancelled before or during the survey period. At that time, most of the substitutions for on-site teaching were online real time video lecturing (74.61%), delivering of lecturers' presentations to students (44.51%), asynchronous online video recordings of the lectures (32.10%) and asynchronous online audio recordings of the lecturer (20.58%). In USA, a survey was conducted among students who attended nine public research universities (Chirikov et al., 2020). One specific question in this survey asked the students how well they had adapted to remote instruction. As many as 49.7% (total N=22,974) of undergraduate students responded with *"Well"* and *"Very well"* to this question, and the same responses were given by 66.6% (total N=12,214) of graduate and professional students. Another report at an earlier phase of this survey performed at five universities in the USA (Soria et al., 2020) revealed that the most important obstacles to transitioning to online learning for graduate / professional (N=7,690) and undergraduate (N=22,519) students were (a) lack of motivation for online learning, (b) lack of interaction with other students, (c) inability to learn effectively in an online format, (d) distracting home environments or lack of access to appropriate study spaces, as well as (e) course content that was not appropriate for online delivery.

A very large global survey (Aristovnik et al., 2020) performed in the period from May 5th until June 15th, 2020 among higher education students from 6 continents and 62 countries (final sample N=30,383) revealed that 86.7% of students experienced cancellation of onsite classes before or during the survey because of the COVID-19 pandemic. The findings of this survey were that onsite teaching was primarily substituted with online lecturing in the form of (a) real-time video conferencing, (b) sending of lecturers' presentations to students and (c) asynchronous video recordings of teaching.

A more recent review study on the responses of education institutions at the beginning and in the later phases of the COVID-19 pandemic (Anthony & Noel, 2021) indicated that educational institutions have benefited from their adaptation to the emergency remote teaching (ERT) situation and adoption of virtual learning (VL). The findings of this study suggested that the institutions were adopting platforms like Zoom, Google Classroom and Microsoft Teams, as well as specialized e-learning platforms like Moodle and Blackboard. It can be concluded that the crisis related to the COVID-19 pandemic has accelerated a more widespread acceptance of such platforms.

From the beginning of the COVID-19 pandemic the American Council on Education (ACE) has repeatedly performed "Puls Point Surveys" with a brief set of standard questions to collect data from college and university leaders about the actions, insights and opinions regarding decisions and challenges that were faced by their higher education institutions (HEIs) due to the COVID-19 pandemic. In fact, eight survey reports were published by ACE in the period from April 2020 to October 2021. In the latest Fall 2021 survey (Melidona et al., 2021), which was performed about 19 months after the first lockdown, the respondents were 230 presidents of private, public and for-profit HEIs. In this survey, interesting responses were given regarding the mode of instruction that best described what their institution would be offering at the beginning of the 2021 fall semester. It was planned that (a) "predominantly in-person, with some online instruction" would be delivered by 59% of HEIs, (b) "exclusively in-person" instruction would be offered by 21% of HEIs, (c) "predominantly online, with some in-person instruction" by 17% of HEIs, and (d) "exclusively online instruction" was planned by 4% of HEIs. When compared to previous ACE surveys for Fall 2020 and Fall 2019 (Turk et al., 2020), a gradual shift back to more dominant in-person instruction and less dominant online instruction can be observed (see *Figure 2*).

Interestingly, even though 80% of the surveyed presidents in the Fall 2021 survey reported that their HEIs planned "predominantly in-person, with some online instruction" or "exclusively in-person" instruction, as many as 90% of them also stated that their institution had a contingency plan that would enable a transition to online teaching and learning if the pandemic conditions worsened.

The *Connected Student Report* (Anft, 2021) surveyed 1,128 students and 1,076 staff from 10 countries (USA and Australia, as well as 8 countries from Europe). This study was conducted in the period from February 18th to March 25th, 2021. Among the most interesting findings of this survey was that (a) campus officials plan to increase institutional efficiency with the use of digital technology and that (b) a large percentage of students (40% of them)

consider having more *flexible* learning options as an important policy their institutions could adopt to increase the level of students' wellbeing. This study also found that 43% of global respondents preferred hybrid courses, in comparison to 21% of them who preferred taking all of their classes online. The study also found that 50% of respondents expected their courses to be online at the end of the pandemic.

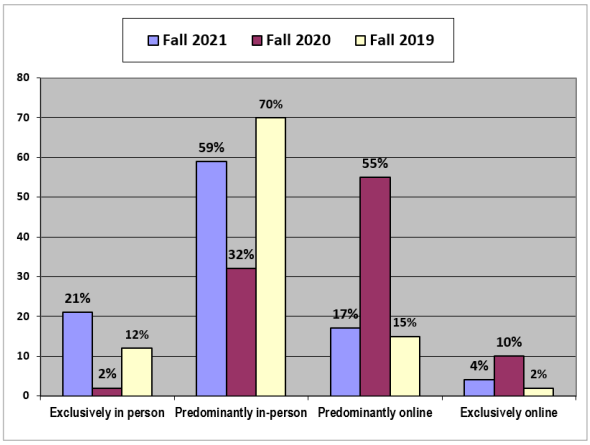


Figure 2. Planned modes of instruction for 1999, 2000 and 2021 fall terms according to ACE surveys (source: Melidona et al. 2021; Turk et al., 2020)

1.2 Motivation for our study

The first aim of the empirical study that is presented in our paper is to *collect and present representative research results of other authors* on the impact of the COVID-19 pandemic on various aspects of teaching and learning at HEIs. Those aspects were predominantly related to student satisfaction regarding online learning during the pandemic, as well as to obstacles/barriers to and advantages of that mode of instruction, and their preference for diverse modes of teaching and learning (fully on-site, hybrid, or fully online) during the pandemic. The second aim was to perform an empirical *pilot study* among students at one higher education institution in Croatia to investigate their interest in (a) on-site, (b) partly on-site and partly online, (c) fully online, and (d) flexible learning options, after the students have experienced more than two years of variable formats of delivery of teaching (fully online, partly online and partly on-site, fully on-site) since the beginning of the COVID-19 pandemic and first lockdown of Croatian HEIs in March 2020. Furthermore, perceived problems/obstacles and advantages of online learning were also in the focus of interest of our pilot study. It is hoped that the findings of our study will contribute to adaptive changes at HEIs to make those institutions more resilient to challenges during the forthcoming waves of the COVID-19 pandemic, as well as more flexible in meeting the actual and future expectations of students after they have gained experience with diverse teaching modes and acquired greater competence for online learning. Finally, additional importance of our study is associated with the demand for teleworking at HEIs, which is more widespread among education professionals than before the pandemic, as well as with the fact that digital "ecosystems" are being accepted and further advanced at primary, secondary and tertiary level of education.

Before presenting the results of the empirical part of our study, it must be noted that there are some limitations to our pilot study. First, the sample of respondents (students) that were included in our survey was rather small (N=70). Second, only students enrolled in two courses at the same higher education institution were included in our convenience sample which makes our *pilot study* close to the term *case study*. Finally, the results of our survey are at least partly reflective of a certain period during the COVID-19 pandemic, i.e. of the attributes of the higher education environment in Croatia during the summer semester of the 2021/2022 academic year when most of the teaching was performed fully online or in hybrid form (as a combination of online and on-site instruction).

2 Theoretical background

Most of the large-scale international and national survey studies related to the effects of the COVID-19 pandemic on higher education that are briefly reviewed in this paper (e.g.: Doolan et al., 2021; Aristovnik et al., 2020; Anft, 2021; Bezjak et al., 2020; Đorđević et al., 2021) were not designed according to a specific theoretical or pedagogical model. However, in these studies it was possible to apply concepts like *blended learning* and *hybrid learning* that are associated with distance education in general and, more specifically, e-learning.

It must be noted that during the COVID-19 pandemic higher education institutions were first closed and then reopened for *on-site* teaching or some *combination of on-site and online* instruction was utilized according to varying epidemiological measures. In the early phases of the COVID-19 pandemic, the previously discussed (in the *Introduction*) concept of *Emergency Remote Teaching* (Hodges, 2020) gained much popularity amongst researchers who were investigating the effects of the pandemic on teaching at HEIs. The terms *flexible learning* and *HyFlex* (for a practical example see Kelly, 2020) were much less used in academic literature in relation to the response of HEIs and their teachers to the COVID-19 pandemic, but they are still addressed in the continuation of our paper.

2.1 Blended learning

Even though the use of the *blended learning* concept in research on educational technology has received certain critique (most recently by Cronje, 2020), it was extensively utilized at the time of the COVID-19

pandemic, mostly in a plethora of papers reporting empirical research findings related to education in various fields like physics (Rizaldi et al., 2021), medicine (Bolatov et al., 2022) and sports science (Finlay et al., 2022), as well as in analytical work to provide recommendation during the pandemic circumstances (see: Kumar et al., 2021; Singh et al., 2021).

It must be emphasized that numerous studies have been performed to investigate the possible advantages of *blended learning* in comparison to *traditional teaching* that have yielded mixed results. For instance, the findings of a recent meta-analysis performed by Müller & Mildenberger (2021) were inconclusive on this issue. As a contrast, it must also be mentioned that a recent *systematic review of systematic reviews* on blended learning (Ashraf et al., 2021) revealed that most blended learning studies were based on students as subjects and discussed blended learning as a generic construct. Furthermore, the results of this study revealed that students' self-regulation toward learning, satisfaction and engagement, as well as academic performance were the most frequently mentioned positive effects of introducing blended learning. A review regarding online and blended learning in schools (from primary to higher) after the onset of the COVID-19 pandemic was recently conducted by Topping et al. (2022). The findings were that, in general, use of digital technology was better in comparison to regular instruction in 85% of studies, but that blended learning was also better than fully online learning, with computer-assisted instruction (CAI) as the most effective approach.

Various definitions of *blended learning* can be found in literature, but the most general view is that blended learning refers to teaching courses both on-site in the classroom (face to face) and online (remotely) with a possibility to combine various pedagogical strategies. Some scholarly papers that address the *blended learning pedagogical approach* imply the notion of blended learning as the use of *face to face* teaching and learning, on the one side, and *online* teaching and learning with the use of specific technology, on the other side, being *optimally combined* or *integrated* to maximize their conjoint benefits. Still, it must be noted that in the most recent COVID-19-related literature the term *blended learning* is predominantly used exclusively to denote teaching that combines on-site and online instructional activities within a single course. One example of a specific use of blended learning in the discipline of *Computer-Aided Language Learning* (CALL) at one HEI in Croatia was well described in a paper by Seljan et al. (2006).

2.2 Hybrid learning

The term *hybrid learning* has traditionally been used as a synonym for *blended learning*. This terminological equivalence is still relevant for most scholarly authors (see Eyal & Gill, 2022) even though taxonomies have been used to discern between the two (see Margulieux at al., 2016). However, more recently some other attributes have been added to the concept of *hybrid learning* in scholarly papers, especially after the onset of the COVID-19 pandemic: (1) at least half of the attendance to instruction has to be online, (2) the teaching in the classroom is broadcast in parallel online and students can choose between attending the instruction *on-site* in the classroom or *remotely* from home via synchronous videoconferencing. When discussing *hybrid learning* the notion of an "optimal blend" or a "combination of best of both worlds" (face to face and online) is less frequently mentioned than when the term *blended learning* is concerned.

3 Research problems and goals

The COVID-19 pandemic has forced higher education institutions (HEIs), students and their instructors to adapt to distant teaching and learning modes, technologies and models. The impacts of these changes on higher education (HE) have been extensively tracked and interpreted. In this paper the selected effects of the pandemic on HE are being reviewed having in mind the *research of other authors* and they are also investigated in an empirical *pilot study* performed at one HEI in Croatia. Therefore, the goals of our study were to (1) identify the effects of the COVID-19 pandemic on students' preferences for particular teaching modes, as well as to (2) determine their satisfaction with teaching that was delivered online at the beginning and later into the pandemic. Furthermore, the goals were to (3) identify the most frequent potential barriers/obstacles to online instruction, and also (4) the perceived advantages of online teaching and learning. Finally, the goals were also to (5) assess students' preferences for flexible learning, in other words, for greater choice of modality of attendance to instruction (on-site vs. online), as well as to (6) examine the potential benefits of online instruction that were not frequently mentioned in literature but which could further support the introduction of flexible learning.

According to the previously mentioned goals, the following research questions were defined for the empirical *pilot study*:

RQ1: What are the preferences for specific modes of instruction and their combinations (on-site, online, hybrid, synchronous, asynchronous) of the students included in our pilot study?

RQ2: How satisfied were the students in our pilot study with hybrid and fully online instruction, respectively?

RQ3: What are the most frequent problems (obstacles/barriers) to online teaching that can be identified in our pilot study, as well as those that were observed by other researchers during the pandemic?

RQ4: What are the greatest potential advantages of online teaching that can be identified in our pilot study,

as well as those that were observed by other researchers during the pandemic?

RQ5: What is the preference for flexible learning in possible continuation of the pandemic (or the post-pandemic period) among the students in our pilot study?

RQ6: Are there some other potential advantages of online instruction which are rarely mentioned in literature but can provide additional explanation of the students' preference for online instruction in our pilot study?

4 Methodology

In our study a combination of a brief literature review and an *empirical pilot study* is used to investigate and illustrate some of the effects of the COVID-19 pandemic on teaching at the higher education level. Also, in the continuation of this paper some of the results of our *pilot study* are discussed in comparison to similar findings of other researchers.

4.1 Data collection instrument

For our pilot study a paper-and-pencil survey was developed with the following thematic sections:

- general evaluation of teaching in a specific course;
- evaluation of online instruction in the course;
- comparison of on-site with online teaching;
- problems and advantages of online instruction;
- evaluation of synchronous vs. asynchronous instruction;
- preferences for various teaching modes and ratios of on-site and online instruction;
- preferences regarding the characteristics of the video image of the teacher during synchronous videoconferencing.

4.2 Respondents and data collection

In our pilot study a convenience sample was used that consisted of 70 students enrolled in two different university courses in the summer semester of the 2021/2022 academic year. There were 23 male and 47 female subjects in this convenience sample and they were between 19 and 24 years of age. Among them, 28 were enrolled in an *undergraduate* course related to computer-mediated communication and were in their 2nd study year, and the other 42 students were enrolled in a *graduate* course related to leadership that is delivered in the 1st study year. All of the students in our pilot study had experienced at least 2 years of teaching and learning in their tertiary education during the COVID-19 pandemic before the survey was conducted. Also, all of them had satisfactory ICT equipment and necessary skills to participate in online instruction.

The data were collected with a paper-and-pencil survey at the end of the summer semester of the 2021/2022 academic year. Students' participation in the survey was anonymous and on a voluntary basis. The survey was approved by the *Ethics Committee* of their higher education institution.

4.3 The educational context of our pilot study

It must be noted that most of the students in higher education in Croatia have experienced a variety of situations from the beginning of the COVID-19 pandemic until the end of the summer semester of the 2021/2022 academic year. As was presented at the beginning of our paper (in *Introduction*), at the beginning of the pandemic there was a lockdown of classes at academic institutions in Croatia for a period of 2-3 months. This was followed with periods of partly on-site and partly online (hybrid) teaching, as well as fully on-site teaching according to the epidemiological conditions at a certain point in time during the pandemic, which was mostly associated with the pandemic "waves". Therefore, the experiences of students who participated in our pilot study during the pandemic led us to define research questions from RQ1 to RQ4. However, RQ5 addresses *flexible teaching and learning* and a specific modification was therefore introduced for our pilot study in the delivery formats of lectures in two university courses that the respondents in the survey were enrolled in. This was done in form of a *parallel delivery* of all lectures (1) on-site in the classroom at college, as well as (2) online as synchronous videoconferencing via Zoom and, post festum, (3) as pre-recordings of the delivered lectures that were made available to the students online two weeks before mid-term exams. In fact, the students were able to freely choose on a weekly basis between attending a certain lecture on-site in the classroom, or online as a Zoom videoconference. Furthermore, for both courses two lectures during the semester were delivered exclusively as *asynchronous* pre-recorded voice over PowerPoint video webcasts. By contrast, less than 1/2 of the practical instruction in form of seminars or laboratory exercises were performed in parallel on-site and online, or as an asynchronous activity. In other words, more than 50% of the aforementioned practical types of instructional activities were delivered on-site in a college classroom/laboratory.

Finally, for both courses a Moodle learning management system was used with course material and links to Zoom web conferencing synchronous broadcast of on-site lectures, as well as asynchronous pre-recorded video webcasts of the lectures on equal topics. When the seminars and laboratory exercises were performed parallelly on-site in the classroom and online the link to BigBlueButton videoconferencing system was provided in Moodle.

4.4 Other unpublished survey results that were included in our study

The Agency for Higher Education in Croatia presented online the results of their two national surveys conducted during the COVID-19 pandemic that have not been published in scholarly papers (Bezjak et al., 2020; Đorđević et al., 2021). Some of the results of these surveys are presented and interpreted in parts of this paper. Also, several of their survey questions were repeated in our pilot study with slightly modified wording.

The other source of previously unpublished data in our study was the survey implemented by the higher education institution (University of Zagreb, 2022) at which our pilot study was also performed that has been repeatedly conducted in the second half of the winter semesters of 2020/2021 and 2021/2022 academic year.

5 Results and discussion

5.1 Students' preferences for various teaching modes

To investigate which mode of teaching delivery was most preferred by the students (N=70) in our pilot study, the following question was asked in the survey *"Which mode of teaching in this course do you prefer (choose only one of the answers)?"* Students' responses are presented in *Figure 3*. According to the data presented in *Figure 3*, most of the students (39%) stated that they preferred *hybrid teaching with synchronous online lecturing and seminars or laboratory exercises performed on-site in the classroom*. Interestingly, the second most frequent choice was *fully online teaching and predominantly synchronous lecturing* (20%). Only about 1/6 of the students (17%) preferred all the teaching *being performed in the traditional on-site manner*. The least preferred modes of teaching were *fully online asynchronous teaching* (14%) and *hybrid (online lecturing and the rest on-site) with predominantly asynchronous lecturing* (10%).

When the data collected by the survey in our pilot study at the end of the summer semester of the 2021/2022 academic year is compared with the data collected earlier in the pandemic (freely accessible online survey data provided by: University of Zagreb, 2022), in the period from late November to the beginning of December 2021 for all the studies, years of study and courses at the same HEI in Croatia (with a total of 1,595 responses to an equal survey question as in our pilot study), there is a considerable difference since the main students' preference at the level of our HEI at that earlier time of data collection was for *fully online teaching and predominantly synchronous teaching* (56%). In comparison, preference for *hybrid teaching with synchronous online lecturing and other activities performed on-site* was at that time reported by only 16% of surveyed students. However, the results of our survey conducted at the end of the summer semester of the 2021/2022 academic year are in greater correspondence to the results of a national survey in Croatia (Đorđević et al., 2021) performed at the beginning of September 2021 (N=4,273), which indicated that 29% of Croatian students would prefer *fully online teaching* in the forthcoming 2021/2022 academic year while 36% of them would prefer *a combination of online and on-site teaching*.

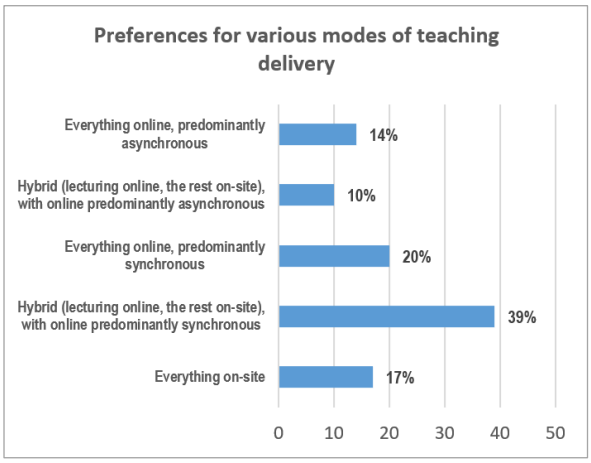


Figure 3. Students' preferences for various modes of teaching delivery in our pilot study (N=70)

5.2 Student satisfaction with online teaching

When the students in our pilot study (N=70) were asked *"To what degree are you satisfied with the performed hybrid mode of teaching in this course?"* as many as 64% responded with *"Totally satisfied"* and 29% with *"Mostly satisfied"* (see *Figure 4*).

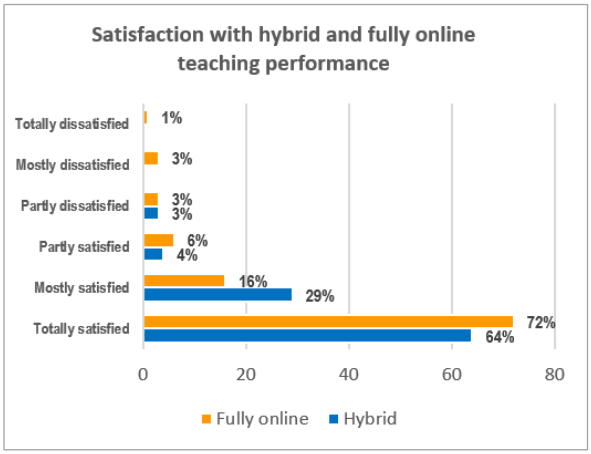


Figure 4. Students' satisfaction with hybrid and fully online teaching performance according to the survey in out pilot study (N=70)

Furthermore, as can also be seen in *Figure 4*, their responses to the question *"To what degree are you satisfied with the performed online part of teaching in this course?"* were also favourable, with 72% of the

subjects who chose the response *"Totally satisfied"* and 16% of them who chose the response *"Mostly satisfied"*. It must be noted that these findings of our pilot study at the level of two courses at the end of summer semester of the 2021/2022 academic year were considerably better for *hybrid teaching* and the *online part of teaching* when compared with responses to similar questions of the survey performed at the level of all courses of our HEI (N=427) in the winter semester (i.e. the previous semester) of the 2021/2022 academic year (freely accessible online survey data provided by: University of Zagreb, 2022).

The large-scale international survey performed for the European Students' Union (ESU) in 41 European countries (N=9,180) in the period from April 21st to May 3rd , 2020 (Doolan et al., 2020), as well as a more global survey performed on students from 62 countries (N=30,383) and conducted from May 5th until June 15th, 2020 (Aristovnik et al., 2020) both revealed a rather positive evaluation of organization of online lectures and specific types of online teaching like real-time video conferences and asynchronous video recordings. However, it must be emphasized that numerous other research papers were identified which revealed mixed results of surveys conducted in the corresponding period at a more specific level and with lesser number of participants. Actually, results of other studies were in the range from *dissatisfaction* of students with their online learning experiences (e.g. Maqableh & Alia, 2021), through *moderate satisfaction* with online teaching (e.g. Reed at al., 2022) and rather *high satisfaction* with online course delivery during the pandemic (e.g. Alturki et al., 2022).

The change of students' perception and evaluation of online teaching after the onset of the COVID-19 pandemic was well detected by the *two surveys* performed by the Agency for Science and Higher Education in Croatia. The *first* national survey (Bezjak et al., 2021), conducted immediately after the lockdown at Croatian HEIs in the period from June to July 2020 (N=1,114), revealed that (at the time of that survey) as many as 50% of the surveyed students evaluated online lectures as *"worse than before the quarantine"* or *"much worse than before the quarantine"*. However, the *second* national survey in Croatia (Đorđević et al., 2021) performed by the Agency for Science and Higher Education at the beginning of September 2021 (N=4,273) indicated that the satisfaction with online teaching had improved since at this later phase of the pandemic as many as 57% of the respondents stated that they were *"satisfied"* or *"very satisfied"* with the quality of *online lecturing* that they received at their HEI. Also, 43% of the respondents declared that they were *"satisfied"* or *"very satisfied"* with the quality of *practical and field work education*.

5.3 Perception of barriers/obstacles to online learning during the pandemic

Barriers to online learning have been investigated for more than two decades (for an example, see Fabry & Higgs, 1997). Two most comprehensive reports on research on barriers to online learning perceived by students were co-authored by Muilenburg & Berge (2004; 2005). These authors identified as many as 45 potential student barriers that were grouped into the following 8 categories as a result of factor analysis: (1) administrative/instructor issues, (2) social interactions, (3) academic skills, (4) technical skills, (5) learner motivation, (6) time and support for study, (7) cost and access to the Internet, and (8) technical problems. However, despite numerous studies that investigated barriers to online learning during the pandemic performed by researchers worldwide, no recent review studies were found so far that were performed across diverse disciplines and addressed the pandemic period.

Let us mention again the large global study (Aristovnik et al., 2020; Aristovnik et al., 2021) that was performed immediately after the pandemic outbreak, with respondents from HEIs in 62 countries (N=30,383), which identified some possible challenges to students at the beginning of the COVID-19 pandemic. When the students in this study were asked to compare their workload before the on-site classes were cancelled due to the first lockdown during the COVID-19 pandemic, 42.6% of them reported that their workload had become larger or significantly larger with online instruction after the first lockdown. The share of students who reported the same workload was 26.6% and, interestingly, 30.8% of students reported that their study workload had become smaller or significantly smaller. According to this large-scale survey, adequate access to course study material (e.g. compulsory and recommended literature) was available to only 50% of respondents, a quiet place to study at home to 55% of them, and a good internet connection to 59%. Some students reported that they had at least some more difficulty staying focused during the online teaching in comparison to onsite teaching, and that their study performance had weakened since the onsite classes had been cancelled. Finally, the most prevailing negative emotions among students were boredom, frustration and anxiety. The authors of this study also warned that since students may face a lack of self-discipline while studying in isolation at home during the lockdown, the lecturers should be careful not to overload students with study materials and assignments.

A similar set of problems was reported in a study by Doolan et al. (2021) regarding students in European countries. This study revealed that common problems related to online learning after the onset of the pandemic were related to inadequate infrastructure for studying from home (not having a quiet place to study, poor internet connection, or poor access to study materials), not being able to have practical classes, lack of face-to-face interaction with teachers and colleagues, monotony of everyday life at home, lack of motivation and increased procrastination, difficulty to concentrate at home and distractions from family members, challenges of studying on their own without the possibility to immediately clarify open questions that emerge during instruction with teachers, and

difficulty with learning complex subject matter in an online environment.

A more specific focus on obstacles to online learning at the beginning of the COVID-19 pandemic was made in a survey performed in the USA by Soria et al. (2020). As mentioned earlier, according to their findings, the most frequent obstacles reported by undergraduate (N=22,519) and graduate and professional students (N=7,690), ordered by their frequency, were: lack of motivation for online learning, lack of interaction with other students, inability to learn effectively in an online format, distracting home environments or lack of access to appropriate study spaces, course content that is not appropriate for online learning, lack of clear expectations for online learning from instructor(s). A separate report from their survey (Chirikov et al., 2020) revealed an increase in the prevalence of major depressive disorder and generalized anxiety disorder after the beginning of the pandemic in 2020 (in comparison to the pre-pandemic 2019). It must be emphasized that the prevalence of those mental health problems was higher among students who did not adapt well to remote instruction.

The Agency for Science and Higher Education of Croatia (Bezjak et al., 2020) performed a survey among Croatian higher education students at the national level in June and July 2020 (N=1114). The findings of this survey revealed that 56% of respondents perceived a higher or much higher work load than before the quarantine and 51% stated that their motivation for performing student's obligations was *lesser* or *much lesser* than before the quarantine.

In our *pilot study* (N=70) conducted at the end of the summer semester of the 2021/2022 academic year the students in two courses were asked *what they considered as greatest problems in online teaching of those courses.* In *Figure 5* their responses are presented in percentages in comparison to the results of the institutional survey at the same HEI (N=640) that was performed in the winter semester of the 2021/2022 academic year (University of Zagreb, 2022), in which an almost identical question was included regarding *problems in online teaching at that institution* in general. The students had to select up to 3 perceived problems from a predefined list of responses. As can be observed from the data presented in *Figure 5*, after five semesters of experience with fully online and hybrid teaching during the pandemic, the most frequently perceived problems in online teaching were associated with the *aggravated communication with the teacher* and *need for equipment for teaching and learning*. Other problems were related to online teaching *being boring* or *not adapted for online delivery*, as well as with *no instruction previously provided to students regarding how to learn in the online environment*. From the data displayed in *Figure 5* it can be concluded that specific problems in online teaching in two subsequent semesters were identified with similar percentages of responses from the students in our survey (N=70) as from the participants in the institutional survey (N=640) performed at the same HEI in the previous semester.

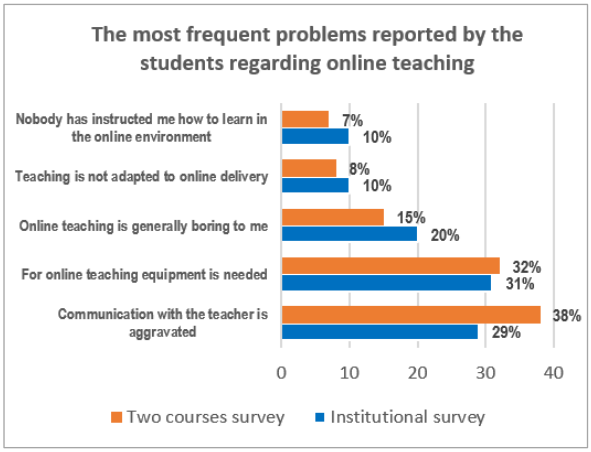


Figure 5. The most frequent problems in online teaching reported by the students at the level of two courses in our pilot study (N=70) and at the institutional level for all courses and all study years (N=640)

5.4 Advantages of online teaching and learning during the pandemic

A systematic review of literature by Saikat et al. (2021) revealed several potential benefits of mobile learning during the COVID-19 pandemic: it (1) enables learning despite social distancing measures, (2) helps in preventing the spread of COVID-19, (3) allows learners and educators to continue the education process from any location, (4) saves time because students do not have to commute to campus for lectures, (5) makes the lessons and class materials available online and offline via a variety of platforms, (6) enables a large number students to attend a course at the same time while attendance in offline courses is limited by the number of seats in the lecture room, (7) reduces cost of utilities and transportation costs, etc.

A survey study performed by Muthuprasad et al. (2021) identified some other benefits of online learning during the pandemic: (1) flexible schedule, (2) opportunity to study at one's own pace and time of convenience, (3) more comfortable environment, (4) enhancement of technical skills of learners, and (5) greater ability to concentrate. Similar findings regarding IT students were made in a survey by Akuratiya & Meddage (2021) in which respondents agreed with the following characteristics of online learning: the ability to learn at one's own pace (63.8%), comfortable surrounding (53%), and online learning being fun / enjoyable (50.7%). An extensive list of benefits of online classes was identified in a study by Li (2022): creating a digital community where students can post questions and inquires, and share materials; enhancing students' familiarity with digital learning technologies and allowing access to further learning materials; making students feel connected with others during the pandemic and creating a sense of community; offering an escape from daily pressures and reduced exposure to the massive amount of

negative news; flexibility, convenience, and more control over the learning process; strengthening self-learning skills and self-discipline; bridging the education gap and enabling continuity during the pandemic; enabling watching playbacks and repeated access to recorded materials.

Several benefits of online learning during the COVID-19 pandemic were also identified in a national survey in Croatia (the sample consisted of 4,273 higher education students) conducted in September 2021 by the Agency for Science and Higher Education (Bezjak et al., 2020). For instance, as many as 60% of the respondents in this survey stated that because of online teaching they had saved on their travel expenses and 56% of respondents stated that during the 2020/2021 academic year they had an opportunity to better organize their time (percentages represent joint responses *"Totally agree"* and *"Mostly agree"* for each of the two questions).

The results of the survey in our pilot study regarding the advantages of online teaching of specific courses (N=70) are presented in *Figure 6*. The greatest advantages were (a) that the students did not have to come to the college but were able to do everything from home, (b) the students were able to manage the process of learning by themselves, and (c) the students were able to choose the time to learn.

Again, similar responses were collected from the students in the institutional survey at the same HEI (N=640) that was performed in the period from November 29th to December 8th, 2021 (University of Zagreb, 2021) with the same three greatest advantages identified.

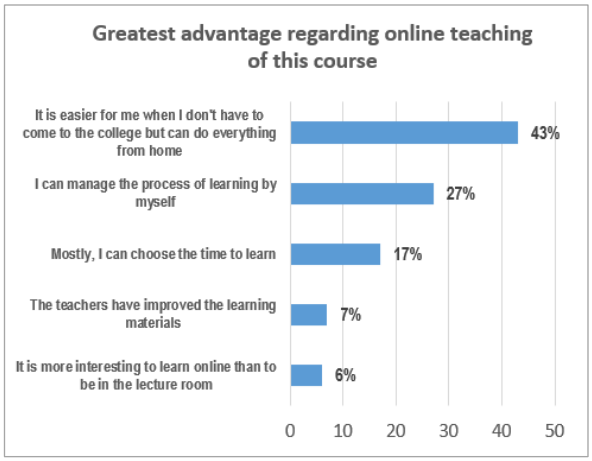


Figure 6. Greatest advantage of online teaching that was perceived by the students in our survey (N=70)

5.5 Preference for flexible learning in possible continuation of the pandemic (or the post-pandemic period)

In an exploratory study aimed at predicting educational approaches and practices in the post-pandemic world, Singh et al. (2021) recognized *flexibility of learning* as one of the important topics emphasized by teachers, students, university administrators and instructional designers. The assumptions were that online learning environments should be *flexible and adaptable*, with greater students' opportunity to manage their own time, as well as with the faculty offering *hybrid* and *blended* options that better fit students' learning styles. In fact, due to the COVID-19 pandemic, there has been an *increasing interest in the expectations of students regarding flexible learning* and how to manage the factors which facilitate the implementation of flexible learning at the level of individual courses and education institutions (see Reginaldo and Ching, 2021).

Even before the COVID-19 pandemic, researchers identified *flexibility* and consideration of unique life situations of students, with the possibility of learning at one's own pace, as an important component of a university learning environment (for an example, see Valtonen et al., 2021). It must be mentioned that the interesting concept of *"HyFlex"*, introduced by Beatty (2007), has been gaining popularity after the onset of the COVID-19 pandemic (for an example see Verrecchia & McGlinchey, 2021). According to Beatty, *HyFlex* is (1) *hybrid* because it combines online and face-to-face teaching and learning activities, and (2) *flexible* because students can choose to attend or not to attend face-to-face course sessions without negative effects on their learning. In other words, at any given week of the course the students can participate either in *face-to-face* sessions or in *online* activities, depending on what their needs and desires are in that week. Pressley (2022) elaborated the difference between *hybrid* learning and *HyFlex* by emphasizing that in the latter format the students can choose the way in which to participate in the course on a session-by-session basis. They can engage with the course material in a format that works best for them at any given time: (1) synchronously in-person, (2) synchronously online or (3) asynchronously online. Simply stated, in a single weekly session three students could each participate in the course in a different format. For a thorough example of an implementation of *HyFlex* during the pandemic regarding teaching of seminars see Detyna et al. (2022).

While investigating the students' preferences for returning to college during the COVID-19 pandemic, Steimle et al. (2022) revealed that the surveyed students expressed different levels of concern regarding the risk of infection due to participation in on-site (face to face) classes. There was also a difference among students regarding their preference for on-campus versus online classes in case of a possible COVID-19 outbreak during the current (or forthcoming) semester. This implies that the *flexible* approach could be the best instructors (or institutional) response regarding the diversity of related students' concerns and interests. Furthermore, the results of a national survey in Croatia (Đorđević et al., 2021) that were mentioned earlier in this paper, like the results of the survey from our pilot study that are presented in *Figure 3*, revealed the variety of students' interest regarding the mixture of instruction delivery formats, from (a) fully online and (b) different ratios of hybrid

(online versus on-site) teaching, to (c) fully on-site course attendance. Another survey (Bagarić et al. 2021) performed among students enrolled in 20 higher education institutions at the beginning of the COVID-19 pandemic in Croatia (N=249) revealed that "flexibility of work" was one of their most frequently mentioned positive personal experiences regarding online learning in that period.

To further investigate the students' interest for *flexibility in course delivery*, two related questions were included in the survey in our pilot study (N=70). In *Figure 7* the responses to the first survey question (*"If you had been given a choice at the beginning of the summer semester of the 2021/2022 academic year, how would you have mostly preferred /provide an approximate estimate/ that the teaching in this course be delivered to you?"*) are presented.

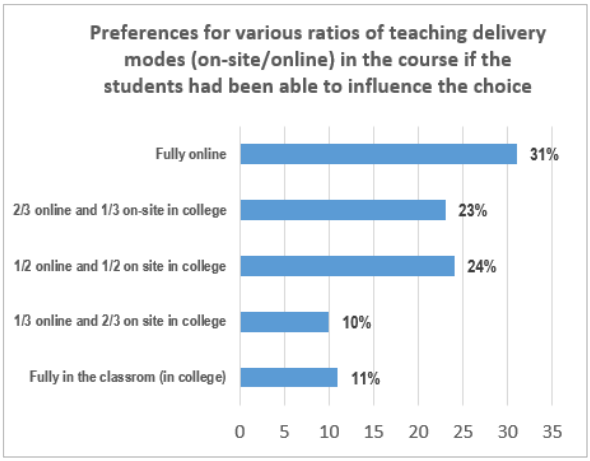


Figure 7. Most preferred ratio of on-site (in the classroom) and online teaching in the course in case the students had been given a choice at the beginning of the summer semester of the 2021/2022 academic year (N=70)

To examine the preferences of participants in our pilot study for more *flexibility* in their choice of mode of teaching, the following survey question was used: *"Do you believe that the teacher of this course should enable that in the future the students can independently choose to attend online teaching of this course instead of on-site teaching in college for any lecture or seminar / laboratory exercise in which they wish to do so?"* Even though the results of data analysis presented in *Figure 8* were derived from the responses of a relatively small convenience sample of students, it is interesting that as many as 76% of students (who responded with *"Certainly yes"* and *"Probably yes"* to the previous question) indicated that they would *prefer to have a choice between attending teaching on-site or online on a weekly basis*. Only 4% (who responded with *"Certainly not"* and *"Probably not"*) stated the opposite. These results should be observed as a possible trend in students' expectations in relation to the ongoing COVID-19 pandemic (and, perhaps, also the forthcoming *post-pandemic* period).

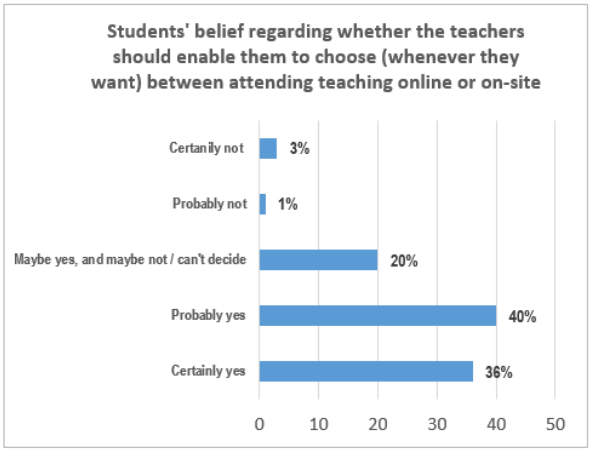


Figure 8. Percentage of students who believe that the teachers should give them the opportunity to choose to attend on-site or on-line teaching on a weekly basis (N=70)

However, it must be emphasized that, especially when a large number of students is enrolled in a specific course, enabling such flexibility to the students would require more time for the preparation of teaching materials and administrative work (i.e. grading of on-site and online assignments) from the instructor(s). In a popularly written article related to the use of *HyFlex* during the COVID-19 pandemic, Kelly (2020) lists the following possible negative aspects of *HyFlex*: fewer students participating in on-site lecturing; greater need for planning by teachers; more administrative work performed for tracking students who participate in different teaching modalities; greater difficulty with engaging students in asynchronous learning; more responsibility required from students for successful learning.

It is important to investigate if the *flexible approach* to teaching and learning results in *inferior learning outcomes* in comparison to *traditional on-site teaching*, as was done in one study during the COVID-19 pandemic by Verrecchia & McGlinchey (2021), and also in other studies conducted immediately before the pandemic (e.g. Feldhammer-Kahret al., 2021), or those performed even earlier (for a list of earlier studies see the review essay by Beatty, 2019).

5.6 Some additional advantages of online versus on-site teaching

It is unmistakable that in our pilot study most of the students preferred *fully online* delivery or *hybrid* delivery of instruction. As can be observed in the data presented in *Figure 7*, if the students had been given the opportunity to determine the preferred ratio of online and on-site teaching (in the classroom) at the beginning of the course they attended in the summer semester of the 2021/2022 academic year, the greatest percentage of students would have chosen fully online delivery (34% of students), followed by 1/2 online and 1/2 on-site delivery (24% of students), and 2/3 online

and 1/3 on-site delivery (23% of students). Also, according to the data presented in *Figure 4*, 88% of the students in our pilot study were *"mostly satisfied"* or *"fully satisfied"* with *fully online* teaching performance, and 93% of them were equally satisfied with *hybrid* teaching performance. Finally, the data displayed in *Figure 3* revealed that 59% of students in our pilot study stated that they preferred either (a) *hybrid* teaching with synchronous online lecturing and seminars or laboratory exercises performed on-site in the classroom, or (b) *fully online* teaching and predominantly synchronous lecturing.

A question arises as to what are the reasons for such preference for online or hybrid teaching, other than those presented in *Figure 6*: (a) the students do not have to come to the college but can do everything from home, (b) the students can manage the process of learning by themselves, and (c) the students can choose the time to learn. To investigate this issue in our pilot study several items that are not commonly found in related literature were also included in the survey. The first of such items was the evaluation of *"Strain and fatigue because of sitting at the desk in the classroom in college versus learning at home".* The students had to choose one of the following responses: *"There is no advantage with online instruction"*, *"There is a minor advantage with online instruction"*, and *"There is great advantage with online instruction"*. The students' responses (N=70) are presented in *Figure 9*.

As can be concluded from the data displayed in *Figure 9*, most of the students (54%) stated that there was a *great advantage* of online instruction regarding their experience of strain and fatigue when sitting at the desk in the classroom during on-site lectures in college.

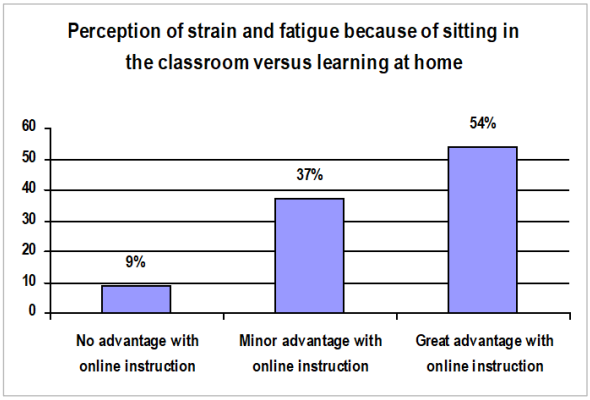


Figure 9. Evaluation of strain and fatigue because of sitting at the desk in the classroom in college versus learning at home (N=70)

The second item in our pilot study was associated with the evaluation of the *"possibility to avoid a feeling of hunger or thirst during lectures, seminars of laboratory exercises in college versus synchronous videoconferencing that is watched and listened to at home."* The students' responses (N=70) are presented in *Figure 10*. Clearly, online instruction, with students at home, is a means for avoiding students being hungry or thirsty during their participation in the process of teaching and learning.

Finally, one more potential advantage of online instruction was investigated in our pilot study (N=70) with the use of a survey item related to *"visibility / sharpness of presentation on the projector screen in college versus the computer screen at home"*. In this case as many as 57% of the students responded with *"There is great advantage with online instruction"*. This means that the visibility of the teaching content is greater if the students use their home computers to attend videoconferencing lectures in comparison to them viewing the projection of teaching materials in lecture halls, classrooms and laboratories on-site in college.

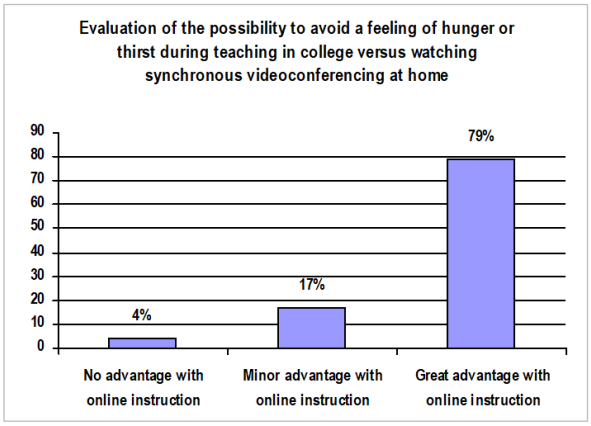


Figure 10. Students' evaluation of the possibility to avoid a feeling of hunger or thirst during teaching in college versus watching synchronous videoconferencing at home (N=70)

When the data presented in *Figure 9* and *Figure 10* are supplemented with those in *Figure 5*, it can be noticed that, according to our pilot study, *the students' perceptions of the advantages of online instruction can be to a considerable degree of practical and not of pedagogical nature*: they feel less strain related to sitting in the classroom, they can have a snack at home instead of feeling hungry in college, they do not have to travel to the college but can do everything from home, they can manage the process of learning by themselves, and can also choose the time to learn.

To briefly conclude the comparison of *on-site* (in the classroom) instruction with *online* instruction, we also emphasize that the analysis of students' responses to several other survey items in our pilot study indicated the following:

- Most of the students stated that it was *more interesting to them to attend online instruction* (71% of them *partly*, *mostly* or *fully agreed* with the corresponding statement).
- Most of the students *did not agree* that it was *easier for them to follow face to face teaching* of the course on-site in college (live) than online teaching (56% of them *partly*, *mostly* or *fully disagreed* with such a notion).

- Most of the students *did not agree* that with the *same amount of used time* during face to face teaching of the course on-site in college (live) they *learn more* than with the use of online synchronous teaching (i.e. with the use of videoconferencing lectures).

The previously listed findings of our pilot study contradict stereotyped assumptions that online learning (a) is less interesting, (b) more difficult to follow, and (c) uses more of students' time in comparison to on-site learning in the classroom. However, this is probably more applicable to *synchronous* teaching via videoconference than a*synchronous* learning activities.

Since *synchronous videoconferencing* was one of the most preferred forms of online delivery of instruction in our pilot study (see *Figure 3*), several additional survey items were devoted to the *use of the streaming of the video image of the instructor*. The students' responses to such survey items indicated that the synchronous projection of *voice over PowerPoint* slides with a *small* or *moderately large* video image of the instructor is on the average best evaluated. This is in concordance with the findings by Korving et al. (2016) that the *visibility of the lecturer* increased attention in web lectures and that the *size of the image* of the lecturer also had a positive effect on attention. Wermeskerken and van Gog (2017) also found that the instructor's face in video lecturing affected the attention in demonstration videos.

6 Conclusion

The COVID-19 pandemic has created the first and only opportunity so far to investigate a global replacement of traditional teaching with fully online and/or hybrid instruction at higher education institutions. As was documented in the Fall 2021 survey of the American Council on Education (Melidona et al., 2021), and is visible in the data displayed in *Figure 2*, by October 2021 the situation at HEIs in the USA appeared to be close to the pre-pandemic fall term of year 2019 in relation to the planned proportion of on-site versus online instruction. That means that the *optimal time frame* for in-depth investigations of massive transitions from traditional to fully online or hybrid instruction due to the pandemic may have elapsed. However, what remains in practical terms is the more advanced skills of teaching staff at HEIs to deliver online education, as well as greater competence and motivation of students to participate in and make use of online instruction delivery. For instance, a recent study by Kovacs et al. (2022) documented that the Austrian students' desired use of e-exams (e-assessment), audiovisual tools (audio, video and online tutorials) and interactive tools (Zoom, MS Teams) after the COVID-19 pandemic was greater in comparison to the self-reported intensity of their use of those media before the first lockdown in the COVID-19 pandemic. Furthermore, HEIs have upgraded their ICT infrastructure and procedures for more effective distance education and have increased their readiness to utilize various communication and collaboration tools in case of potential deterioration of the epidemiological situation.

In this paper we have outlined some of the important findings of other researchers regarding teaching at HEIs from the beginning of the COVID-19 pandemic to the period more than 2 years after it started to affect academic education. In the brief literature review and in our pilot study we have supported the popular claim that the COVID-19 pandemic has changed the average student's preferences toward hybrid and fully online teaching modes. In fact, according to most studies, the students' *satisfaction with teaching* that was delivered online was of acceptable level at the beginning of the COVID-19 pandemic and it even increased in later phases of the pandemic. Numerous studies, including our pilot study, have identified diverse *barriers (obstacles)* to online instruction as well as potential *advantages* of online teaching and learning. In our pilot study we have also determined considerable interest of students in *flexible learning* (i.e. in their greater choice of modality of attendance in instruction – on-site vs. online) and uncovered some potential benefits of online instruction that had not been frequently mentioned in literature.

In terms of a more specific response to the research questions (RQ1-RQ6), the following conclusions could be drawn from the previously presented results of our pilot study (N=70):

RQ1: As many as 78% of the students in our *pilot study* jointly stated that if they had a choice they would prefer one of the following instruction delivery modes: (a) fully online, (b) 2/3 online and 1/3 on-site or (c) 1/2 online and 1/2 on-site instruction. More specifically, the form of predominantly *synchronous* online instruction, that is (a) hybrid or (b) fully online, was preferred by 59% of the students in our pilot study.

RQ2: As many as 64% of the students in our *pilot study* were *"Totally satisfied"* or *"Mostly satisfied"* with *hybrid* part of instruction and 72% were equally satisfied with the *online* part of teaching.

RQ3: The most frequent problems (obstacles/barriers) to online teaching that were identified in our *pilot study* were *aggravated communication with the teacher, need for equipment for teaching and learning, online teaching being boring and teaching not being adapted for online delivery*. A more detailed list of barriers/obstacles can also be derived from large-scale international studies: lack of motivation for online learning, lack of interaction with teachers and other students, inability to learn effectively in an online format, difficulty to concentrate at home, distracting home environments or lack of access to appropriate study spaces, and course content that is not appropriate for online learning.

RQ4: The greatest potential advantages of online teaching that were identified in our *pilot study* were mostly of *practical nature* and included: no need to commute to college and possibility to do everything from home, possibility to manage the process of learning by the students themselves and more potential to choose the time to learn. Similar findings were also

most frequently mentioned in research performed by other authors. However, in our *pilot study* we also noticed that a substantial proportion of students identified several other potential *conveniences of online learning*: less strain or fatigue can be associated with online learning in comparison to sitting at a desk in the classroom; online learning from home implies that students can avoid feeling hunger or thirst while attending lectures; the visibility of the teaching content can be greater if the students use their home computers to attend videoconferencing lectures in comparison to viewing the image on the projector screen on-site in classrooms.

RQ5: The results of our *pilot study* indicate that in possible continuation of the pandemic (or the post-pandemic period) most of the students would probably prefer *flexible* teaching and learning that enables them to choose between (a) attending instruction on-site or (b) attending synchronous videoconferencing lecturing that is in parallel broadcasted online.

RQ6: The "common sense" additional findings obtained by the students in our *pilot study* (and previously discussed in relation to RQ4) include their perception of more strain or fatigue when sitting at a table in the classroom than in online learning from home, as well as of greater possibility of avoiding feeling hunger or thirst while attending lectures from home. Together with other identified advantages of online teaching and learning, such findings highlight the predominantly *practical* (or *convenience*) nature of most of the advantages of studying *fully online* or in a predominantly *hybrid* format that were identified in literature, as well as in our *pilot study*. However, these simple *utilitarian benefits* may in fact have a significant positive impact on the students' prevalent perception of and preference for *fully online* and *hybrid* instruction. In fact, saving *time* that would be otherwise spent on commuting between home and campus and also on waiting between traditional classes attended on-site at college is perhaps the greatest advantage of *fully online* or *hybrid* modality of delivering education. As a support to the notion that *time* is an issue in online versus traditional on-site instruction let us mention that Barbera et al. (2014) have tried to draw attention to the "paradox" of time in the implementation of technologically mediated pedagogy.

The term "paradox" has been used by researchers for various aspects of e-learning and one such "contradiction" can be located in occasional findings that the average students' satisfaction with online instruction may be in discord with the potential demand for more study time and increased effort in this type of instruction (in comparison to traditional teaching and learning). It is the opinion of the authors of this paper that this problem should be particularly dealt with in case of *asynchronous* teaching and learning. In the continuation of our research after this *pilot study* it would be opportune to address such issues in more depth with an advanced survey instrument, more sophisticated methodology and a greater number of subjects selected across diverse courses and higher education institutions.

To conclude (as was mention earlier in this paper), the findings of our study could contribute to greater *resilience* of HEIs in the continuation and after the COVID-19 pandemic, as well as to their *flexibility* in meeting expectations of students. In addition, our findings could motivate HEIs to responding to the growing interest for *telework* among education professionals, especially as *digital learning "ecosystems"* are advanced worldwide at the tertiary level of education.

7 Limitations of our study

The limitations of our *pilot study* are predominantly related to the number of respondents (N=70) and limited number of courses at only one HEI that were included in the survey. However, comparative results of statistical analyses from other surveys were presented in this paper wherever possible. Furthermore, some items in our survey were carefully articulated to better correspond to items from similar earlier studies performed by other researchers.

References

Akuratiya, D.A., & Meddage, D.N.R. (2020). Students' perception of online learning during COVID-19 pandemic: A survey study of IT Students. *International Journal of Research and Innovation in Social Science (IJRISS)*, 4(9), 755-758. Retrieved June 15th, 2022, https://www.rsisinternational.org/journals/ijriss/Digital-Library/volume-4-issue-9/755-758.pdf

Alturki, S., Alija, S., & Stuckenschmidt, H. (2022). Online delivery in higher education during pandemics: Students' perspective. *TEM Journal*, 11(2), 882-892, https://www.temjournal.com/content/112/TEMJournalMay2022_882_892.pdf

Anft, M. (2021). *Connected Student Report, 2nd edition: Insights into Global Higher Education Trends from over 2,000 Students and Staff*. Salesforce, Inc., San Francisco, CA, United States. Retrieved June 15th, 2022, https://www.salesforce.org/highered/download-connected-student-report-second-edition/

Anthony Jnr, B., & Noel, S. (2021). Examining the adoption of emergency remote teaching and virtual learning during and after COVID-19 pandemic. *International Journal of Educational Management*, 35(6), 1136-1150. https://doi.org/10.1108/IJEM-08-2020-0370

Aristovnik, A., Keržič, D., Ravšelj, D., Tomaževič, N., & Umek, L. (2020). Impacts of the COVID-19 pandemic on life of higher education students: A

global perspective. *Sustainability*, 12, 8438. https://doi.org/10.3390/su12208438

Aristovnik, A., Keržič, D., Ravšelj, D., Tomaževič, N., & Umek, L. (2021). *Impacts of the Covid-19 Pandemic on Life of Higher Education Students: Global Survey Dataset from the First Wave. Data in Brief 2021*, 107659. https://doi.org/10.1016/j.dib.2021.107659

Ashraf, M.A., Yang, M., Zhang, Y., Denden, M., Tlili, A., Liu, J., Huang, R., & Burgos, D. (2021). A systematic review of systematic reviews on blended learning: Trends, gaps and future directions. *Psychology Research and Behavior Management*, 14, 1525-1541. https://doi.org/10.2147/PRBM.S331741

Bagarić, Ž., Plantak, D., & Škof, B. (2021). Croatian experiences of distance learning at the beginning of the public health crisis caused by COVID-19. *Proceedings of the 15th International Technology, Education and Development Conference* - INTED2021 (pp. 2859-2866), Valencia, Spain. https://dx.doi.org/10.21125/inted.2021.0610

Barbera, E., Gros, B., & Kirschner, P. (2014). Paradox of time in research on educational technology. *Time & Society*, 24(1), 96-108. https://doi.org/10.1177/0961463X14522178

Beatty, B. (2007). Transitioning to an online world: Using HyFlex courses to bridge the gap. In C. Montgomerie & J. Seale (Eds.), *Proceedings of ED-MEDIA 2007--World Conference on Educational Multimedia, Hypermedia & Telecommunications* (pp. 2701-2706). Vancouver, Canada: Association for the Advancement of Computing in Education (AACE). Retrieved June 15th, 2022, https://www.learntechlib.org/primary/p/25752/

Beatty, B. (2019). Evaluating the impact of hybrid-flexible courses and programs: Highlight from selected studies. In Brian J. Beatty (Ed.), *Hybrid-Flexible Course Design: Implementing Student-Centered Hybrid Classes* [Online], EdTech Books. Retrieved June 15th, 2022, https://edtechbooks.org/hyflex/impact

Bezjak, S., Đorđević, M., & Plužarić, Ž. (2020). *Izazovi u visokom obrazovanju za vrijeme pandemije bolesti COVID-19 i socijalne izolacije: iskustva i potrebe studenata i djelatnika visokih učilišta* [*Challenges in higher education during the COVID-19 pandemic and social isolation: Experiences and needs of students and employees of higher education institutions* / in Croatian language], Agency for Science and Higher Education, Croatia. Retrieved June 15th, 2022, https://www.azvo.hr/images/stories/publikacije/Rezultati_istra%C5%BEivanja_Izazovi_u_visokom_obrazovanju_za_vrijeme_pandemije_bolesti_COVID19_i_socijalne_izolacije.pdf

BBC (2020). Coronavirus: The world in lockdown in maps and charts. BBC News Services, British Broadcasting Corporation (BBC), London, UK. Retrieved June 15th, 2022, https://www.bbc.com/news/world-52103747

Bolatov, A.K., Gabbasova, A.M., Baikanova, R.K., Igenbayeva, B.B., & Pavalkis, D. (2022). Online or blended learning: The COVID-19 pandemic and first-year medical students' academic motivation. *Medical Science Educator*, **32,** 221–228. https://doi.org/10.1007/s40670-021-01464-y

Bond, M., Bedenlier, S., Marín, V.I., & Händel, M. (2021). Emergency remote teaching in higher education: mapping the first global online semester. *International Journal of Educational Technology in Higher Education*, 18:50. https://doi.org/10.1186/s41239-021-00282-x

Chirikov, I., Soria, K. M, Horgos, B., & Jones-White, D. (2020). *Undergraduate and graduate students' mental health during the COVID-19 pandemic*. UC Berkeley: Center for Studies in Higher Education. Retrieved June 15th, 2022, https://escholarship.org/uc/item/80k5d5hw

Cronje, J. C. 2020. Towards a new definition of blended learning. *The Electronic Journal of e-Learning*, 18(2), 114-121. https://doi.org/10.34190/EJEL.20.18.2.001

Detyna, M., Sanchez-Pizani, R., Giampietro, V., Dommett, E.J., & Dyer, K. (2022). Hybrid flexible (HyFlex) teaching and learning: Climbing the mountain of implementation challenges for synchronous online and face-to-face seminars during a pandemic. *Learning Environments Research*. https://doi.org/10.1007/s10984-022-09408-y

Doolan, K., Barada, V., Burić, I., Krolo, K., & Tonković, Ž. (2021). *Student life during the COVID-19 pandemic lockdown: Europe-wide insights.* European Students' Union (ESU). Retrieved June 15th, 2022, https://esu-online.org/?publication=student-life-during-the-covid-19-pandemic-lockdown-europe-wide-insights

DW (2020). *Coronavirus: What are the lockdown measures across Europe?* Deutsche Welle (DW), Bonn/Berlin, Germany. Retrieved on June 15th, 2022, https://www.dw.com/en/coronavirus-what-are-the-lockdown-measures-across-europe/a-52905137

Đorđević, M., Plužarić, Ž., & Bišćan, I. (2021). *Studenti i pandemija: Kako smo (pre)živjeli?* [*Students and the pandemic: How we survived* / in Croatian] Agency for Science and Higher Education, Croatia, 2021. Retrieved June 15th, 2022, https://www.azvo.hr/images/stories/novosti/Rezultati_istra%C5%BEivanja _Studenti_i_pandemija_-_Kako_smo_pre%C5%BEivjeli_lektorirano.pdf

Eyal, L., & Gil, E. (2022). Hybrid learning spaces — a three-fold evolving perspective. In: Gil, E., Mor, Y., Dimitriadis, Y., Köppe, C. (eds), *Hybrid Learning Spaces. Understanding Teaching-Learning Practice.* Springer, Cham. https://doi.org/10.1007/978-3-030-88520-5_2

Fabry, D. L., & Higgs, J. R. (1997). Barriers to the effective use of technology in education: Current status. *Journal of Educational Computing Research*, 17(4), 385. Retrieved June 25, 2022 from https://www.learntechlib.org/p/86209/.

Feldhammer-Kahr, M., Dreisiebner, S., Arendasy, M., & Paechter, M. (2021). One month before the pandemic: Students' preferences for flexible learning and what we can learn. In: C. Pranca & M. Wang (Eds.), *Psychological Applications and Trends 2021*, pp. 187-191. https://doi.org/10.36315/2021inpact039

Finlay, M. J., Tinnion, D. J., & Simpson, T. (2022). A virtual versus blended learning approach to higher education during the COVID-19 pandemic: The experiences of a sport and exercise science student cohort. *Journal of Hospitality, Leisure, Sport & Tourism Education*, 30, 100363. https://doi.org/10.1016/j.jhlste.2021.100363.

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The Difference between Emergency Remote Teaching and online learning. *EDUCAUSE Review*, March 2020. Retrieved June 15th, 2022, https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Kelly, K. (2020). *COVID-19 Planning for Fall 2020: A Closer Look at Hybrid-Flexible Course Design*. PhilOnEdTech blog. Retrieved June 15th, 2020, https://philonedtech.com/covid-19-planning-for-fall-2020-a-closer-look-at-hybrid-flexible-course-design/

Korving, H., Hernández, M., & De Groot, E. (2016). Look at me and pay attention! A study on the relation between visibility and attention in weblectures. *Computers & Education*, 94, 151-161.

Kovacs, C., Jadin, T., & Ortner, C. (2022). Austrian college students' experiences with digital media learning during the first COVID-19 lockdown. *Fronteers in Psychology*, 7 February 2022. https://doi.org/10.3389/fpsyg.2022.734138

Kumar, A., Krishnamurthi, R., Bhatia, S., Kaushik, K., Jyothi Ahuja, N., Nayyar, A., & Masud, M. (2021). Blended learning tools and practices: A comprehensive analysis. *IEEE Access*, 9, 85151-85197. https://doi.org/10.1109/ACCESS.2021.3085844

Li, D. (2022). The shift to online classes during the Covid-19 pandemic: Benefits, challenges, and required improvements from the students' perspective. *The Electronic Journal of e-Learning*, 20(1), 1-18. https://doi.org/10.34190/ejel.20.1.2106

Margulieux, L. E., McCracken, W. M., & Catrambone, R. (2016). A taxonomy to define courses that mix face-to-face and online learning. *Educational Research Review*, 19, 104-118. https://doi.org/10.1016/j.edurev.2016.07.001.

Maqableh, M., & Alia, M. (2021). Evaluation online learning of undergraduate students under lockdown amidst COVID-19 Pandemic: The online learning experience and students' satisfaction. *Children and Youth Services Review*, Elsevier, 128(C). https://doi.org/10.1016/j.childyouth.2021.106160

Melidona, D., Taylor, M., & McNamee, T. C. (2021). *2021 Fall Term Pulse Point Survey of College and University Presidents*. American Council on Education, Washington, DC, USA. Retrieved June 15th, 2021, https://www.acenet.edu/Documents/Pulse-Point-Survey-Fall2021.pdf

Muilenburg, L., & Berge, Z. (2004). A factor analytic study of barriers perceived by students to online learning. In J. Nall & R. Robson (Eds.), *Proceedings of E-Learn 2004--World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 2050-2053). Washington, DC, USA: Association for the Advancement of Computing in Education (AACE). Retrieved June 15th, 2022, https://www.learntechlib.org/primary/p/11623/.

Muilenburg, L., & Berge, Z. (2005). Student barriers to online learning: A factor analytic study, *Distance Education*, 26(1), 29-48. https://doi.org/10.1080/01587910500081269

Muthuprasad, T., Aiswarya, S., Aditya, K. S., & Jha, G. K. (2021). Students' perception and preference for online education in India during COVID -19 pandemic. *Social Sciences & Humanities Open*, 3(1), 100101. https://doi.org/10.1016/j.ssaho.2020.100101.

Müller, C., & Mildenberger, T. (2021). Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education. *Educational Research Review*, 34, 100394, https://doi.org/10.1016/j.edurev.2021.100394

Our World in Data (2022). School closures during the COVID-19 pandemic [interactive chart map for April 1st, 2020]. *Oxford COVID-19 Government Response Tracker*, Blavatnik School of Government, University of Oxford. Retrieved June 15th, 2022, https://ourworldindata.org/grapher/school-closures-covid

Pressley, J.P. (2022, March). Explaining the difference between HyFlex and hybrid teaching models. *EdTech: Focus on Higher Education*. Retrieved June 15th,

https://edtechmagazine.com/higher/article/2022/03/hyflex-hybrid-teaching-models-whats-the-difference-eperfcon

Reed, P., Adjei-Mensah, P., & Truzoli, R. (2022). Satisfaction with online-teaching is affected by COVID-status for university students. *The Open Psychology Journal*, 15. https://doi.org/10.2174/18743501-v15-e2203020

Reginaldo, A. L., & Ching, D. A., (2021). Online learning expectations among engineering students: Analyzing pre-determined factors in the implementation of flexible learning. *International Journal of Educational Management and Development Studies*, 2(4), 24-43. https://doi.org/10.53378/352076

Rizaldi, D. R., Doyan, A., Makhrus, M., Fatimah, Z., & Nurhayati, E. (2021). Adaptation to new normal conditions: Students physics learning outcomes using the blended learning model. *International Journal of Asian Education*, *2*(3), 369–376. https://doi.org/10.46966/ijae.v2i3.171

Saikat, S., Dhillon, J. S., Wan Ahmad, W. F., & Jamaluddin, R.A. (2021). A systematic review of the benefits and challenges of mobile learning during the COVID-19 pandemic. *Educational Sciences*, 11(9), 459. https://doi.org/10.3390/educsci11090459

Seljan, S., Banek Zorica, M., Špiranec, S., & Lasić-Lazić, J. (2006). CALL (Computer-Assisted Language Learning) and Distance Learning. *Proceedings of the 29th International convention MIPRO 2006* (pp. 145-151), Opatija, Croatia.

Singh, J., Evans, E., Reed, A., Karch, L., Qualey, K., Singh, L., & Wiersma, H. (2022). Online, hybrid, and face-to-face learning through the eyes of faculty, students, administrators, and instructional designers: Lessons learned and directions for the post-vaccine and post-pandemic/COVID-19 world. *Journal of Educational Technology Systems*. 50(3), 301-326. doi:10.1177/00472395211063754

Singh, J., Steele, K., & Singh, L. (2021). Combining the best of online and face-to-face learning: Hybrid and blended learning approach for COVID-19, post vaccine, & post-pandemic world. *Journal of Educational Technology Systems*, 50(2), 140-171. https://doi.org/10.1177/00472395211047865

Soria, K. M., Chirikov, I., & Jones-White, D. (2020). *The Obstacles to Remote Learning for Undergraduate, Graduate, and Professional Students*. UC Berkeley: Center for Studies in Higher Education, 2020. Retrieved June 15th, 2022, https://escholarship.org/uc/item/5624p4d7

Steimle, L.N., Sun, Y., Johnson, L., Besedeš, T., Mokhtarian, P., & Nazzal, D. (2022). Students' preferences for returning to colleges and universities during the COVID-19 pandemic: A discrete choice experiment. *Socio-Economic Planning Sciences*, 82, 101266. https://doi.org/10.1016/j.seps.2022.101266.

Stewart, W. H. (2021). A global crash-course in teaching and learning online: A thematic review of empirical Emergency Remote Teaching (ERT) studies in higher education during year 1 of COVID-19. *Open Praxis*, 13(1), 89–102. https://search.informit.org/doi/10.3316/informit.758902304536019

Topping, K. J., Douglas, W., Robertson, D., & Ferguson, N. (2022). Effectiveness of online and blended learning from schools: A systematic review. *Review of Education*, 10, e3353. https://doi.org/10.1002/rev3.3353

Turk, J. M., Soler, M. C., & Ramos, A. M. (2020). *College and University Presidents Respond to COVID-19: 2020 Fall Term Survey*. American Council on Education, Washington, DC, USA. Retrieved June 15th, 2021, https://www.acenet.edu/Documents/Presidents-Respond-COVID19-Fall2020.pdf

Valtonen, T., Leppänen, U., Hyypiä, M., Kokko, A., Manninen, J., Vartiainen, H., Sointu, E., & Hirsto, L. (2021). Learning environments preferred by university students: A shift toward informal and flexible learning environments. *Learning Environments Research*, 24, 371–388. https://doi.org/10.1007/s10984-020-09339-6

Verrecchia, P. J., & McGlinchey, M. J. (2021). Teaching during covid: The effectiveness of the HyFlex classroom in a 300 Level Statistics class. *Journal of Education and Training Studies*, 9(3), 23-27. https://doi.org/10.11114/jets.v9i3.5146

UN (2020). *United Nations "Policy Brief: Education during COVID-19 and beyond"*, 2020. Retrieved June 15th, 2022, https://www.un.org/sites/un2.un.org/files/sg_policy_brief_covid-19_and_education_august_2020.pdf

University of Zagreb (2022). *FOI studentske ankete* [FOI student surveys / data from surveys performed at the Faculty of Organization and Informatics in Varaždin, University of Zagreb; available publicly in Croatian language]. Faculty of Organization and Informatics, University of Zagreb, Varaždin, Croatia. Retrieved June 15th, 2022, https://ankete.foi.hr

Wermeskerken, M., & van Gog, T. (2017). Seeing the instructor's face and gaze in demonstration video examples affects attention allocation but not learning. *Computers & Education*, 113, 98-107.

Digital Transformation and Corporate Culture

Albena Antonova, Yavor Dankov
Digital Platforms Building Blocks: Focusing on Agile Organizations and Product-Service Systems

Tea Mijač, Mario Jadrić, Maja Ćukušić
Information Quality of Digital Services in Higher Education: Higher Order Construct Analysis

Rudi Pleteš, Martina Tomičić Furjan, Željko Dobrović
Factors influencing project manager's ability to deliver a software engineering project

Charalampos Alexopoulos, Igor Pihir, Martina Tomičić Furjan
Automatic End-to-End Decomposition and Semantic Annotation of Laws Using High-Performance-Computing and Open Data as a Potential Driver for Digital Transformation

Olivia Doina Negoita, Mirona Ana-Maria Popescu, Albena Antonova, Dominik Bork
A comparison of business platforms used by SMEs to digitalize management activities

Clara Viganego, Anna Akhmedova, Marta Mas-Machuca
Digital skills gap: Industry demand vs. Higher Education Institutions offerings

Digital Platforms Building Blocks: Focusing on Agile Organization and Product-Service Systems

Albena Antonova and Yavor Dankov
Faculty of Mathematics and Informatics
Sofia University "St. Kliment Ohridski"
5 "James Baucher" Blvd, 1164, Sofia, Bulgaria
{a_antonova, yavor.dankov}@fmi.uni-sofia.bg

Abstract. *Many companies struggle from dynamic business environment, volatile markets, lack of resources, and multiple restrictions. At the same time, digital platforms emerge as industry disruptors, providing new business models, networking patterns and channels of cooperation.*
The present research aims to outline how the principles behind agile organization and product-service systems can support companies to explore digital platform opportunities. Stepping on theoretical research, the paper proposes a conceptual model enabling companies to consider how to fit to the logic of the digital platforms. The paper is part of the project DEMO, focusing on preparing the next professionals in digital platform enterprises.

Keywords. Digital Platforms, Agile Organization, Product-Service System, Digital Platform Enterprises

1 Introduction

The evolving market dynamics and increased global risks impose many new hardships for companies. To survive and remain competitive, organizations need to be flexible and constantly innovate, adopting new digital strategies, new working styles, and new product-service offerings. Many of businesses consider the transition from established organizational practices and governance to digital platforms, adopting more adaptive methodologies and capacity for real-time reactivity. In this regard, more and more organizations are considering new platform technologies to restructure their value offering and governance structure.

Digitalization further challenges traditional business principles, business models, and value-creation mechanisms. For example, digital technologies increasingly blur the distinction and the differences between "goods" and "services". In the digital world, the value offerings are organized in complex "product-service offerings" (Gaiardelli et al., 2014), often constructed within complex ecosystems, combining institutions, companies, and infrastructure (Vargo & Lush, 2016). The ongoing digital transformation further steps on personalized and adaptable smart services (Beverungen et al., 2017) provided by digitally-enhanced products within the aligned systems and platforms.

At the same time, the digital platform companies and their disruptive business models expanded dramatically in size and scale over the last decade (de Reuver et al., 2018). Platform thinking offers many benefits such as better models to connect third parties, facilitating transactions and efficiency in assets management and innovations. However, very few platform businesses originate from the EU and Central and Eastern Europe (CEE). Furthermore, not every industry or business can benefit from adopting new platforms. Even more, it is still unclear how existing companies and mainly small and medium enterprises (SMEs) can benefit from platform principles. Considering that digital platforms play a crucial role in digital transformation, for example in the field of Industry 4.0, and smart manufacturing, it is necessary to study and conceptualize more precisely which industries and business processes can successfully transit toward platform businesses. Moreover, digital platforms continue to raise public debates on various social, technological, and tax issues and need further attention, regarding policy regulations, data protection and much more.

The present paper aims to investigate agile organizations and product-service systems as building blocks toward considering digital platform opportunities. On one hand, agile organizations' principles support companies to become more adaptive, agile, fast, and reactive to changes. In parallel, servitization approaches enable them to transfer value offerings from pure products to more complex and digitally supported product-service systems. However, adopting a digital platform enterprise thinking is not just about online operations and leveraging the networking effects.

This paper is based on the literature review of the ongoing EU Erasmus+ project DEMO. DEMO focuses to design and deliver quality training solutions,

learning materials, and practice-oriented case studies for CEE graduate students in business and technical specialties, in order to prepare next professionals in the field of digital platform enterprises.

The first part of the paper provides a general summary of the key concepts and characteristics of agile organizations and PSS. Next are outlined the building blocks of digital platforms, described in a general conceptual model. Finally, the discussion provides several considerations on the platform's constraints, benefits and factors for success.

2 Background

The first section focuses on main theoretical principles behind agile organization and product-service-systems, serving as starting point for next theoretical and practical investigations toward digital platform transition.

2.1 Agile organizations

Agility is defined as "the ability to both create and respond to change to profit in a turbulent business environment." (Highsmith, 2002; Highsmith, 2010). In this context, the term *Agile* or *Business Agile* refers to "the approach of providing greater flexibility and faster decision-making in the modern business world, a world where organizations that aren't Business Agile will take longer to succeed and be less flexible in this modern, demanding business world, and fail at a faster rate potentially" (Taylor, 2021). The *Organizational Agility*, described in (Ozbayrac, 2020) represents "building cross-functional teams that can deliver an end-to-end business or process in the organization. *Agile organizations* aim to combine both features and create a system where the bricks are built on the platform. The *Agile organization* is sustainable as a large corporation and is fast and flexible as a start-up at the same time."

Concepts of Agile Methodologies and Project Management

Agile methodologies are closely related to software development projects, management of customer requirements, and general project management practices. Although the biggest popularity of Agile methodologies is within project-based organizations (PBO), these principles are increasingly suitable for all types of businesses.

The introduction of innovative and flexible methodologies and practices often result from different challenges related to the management of linear business processes. In 2001, the "Manifesto for Agile Software Development" (Beedle et al., 2001) was created, presenting the following four central values:

- Individuals and interactions over processes and tools;
- Working software over comprehensive documentation;
- Customer collaboration over contract negotiation;
- Responding to change over following a plan.

This gives the launch of agile methodologies and practices and their permanent place in software development and the adoption of agile practices in organizational management and every business area.

Examples of such agile methodologies are Extreme Programming, SCRUM, Dynamic Systems Development Method (DSDM), Crystal, Feature-Driven Development, and Pragmatic Programming (Taylor, 2021). Other examples are Lean, Agile modelling, Agile Unified Process (UAP), Kanban, and Scrumban.

Aligning with the "Manifesto for Agile Software Development" which describes twelve main principles, Perkin & Abraham (2021) propose **twelve agile business principles**. Three of them are especially relevant for digital platform enterprise building blocks (Perkin & Abraham, 2021):

1. "The primary orientation is towards customer needs, delivered through constant improvement of customer experience";
2. "Strategies and tactics are highly adaptive and responsive, and change is welcomed";
3. "The best results emerge from small teams with a high degree of autonomy";

Considering these twelve agile business principles, the management board of an organization should align and synchronize in order to meet all the requirements of the business environment and stakeholders and reach a higher degree of flexibility and adaptability.

Furthermore, management often has to face the challenges of doing agile and being agile at the same time in order to improve performance and achieve business goals and performance, and also to be adaptive leaders who react adequately to the fast-changing ecosystems and requirements. Figure 1 represents the perception of the enterprise agility vision, adapted from (Highsmith, 2014). The diagram illustrates the overall vision concerning the management and its positioning toward doing agile or being agile leaders.

Figure 1 presents the two main directions of the *Enterprise Agility Vision* as a relationship between doing things in an agile way and being an agile leader. Part of the things directly related to doing your job in an agile way and involve indicators of capability, quality, do less, speed-to-value, and so on. All this leads to being an agile leader, which includes indicators and capabilities such as engaging, adapting, exploring, and so on, which undoubtedly lead to the adaptation and perception of agile methods and principles.

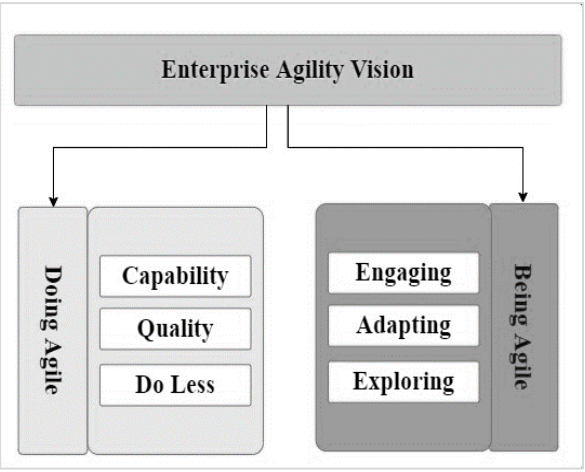


Figure 1. Enterprise Agility Vision adapted from (Highsmith, 2014)

2.2 Product-Service Systems

The main idea behind the emergence of the Product-Service System (PSS) is the changed perspective from "products" (material artifacts) to "services" (intangible offerings), considering the whole system that facilitates value creation and value exchange. The term PSS is first mentioned in a Dutch policy paper on the environment and economy (Goedkoop et al., 1999). In this definition, the authors state that PSS is "a system of products, services, infrastructure and network support that continually strives to be competitive, satisfy customer needs and result in less environmental impact than traditional business models" (Goedkoop et al., 1999).

A more general summary of other PSS definitions highlights some of the main background concepts:

- PSS is a complex value offering composed of a mix of tangible and non-tangible elements that can be labelled as products, services, and infrastructures (Goedkoop et al., 1999; Barravecchia et al., 2021);
- PSS is a systematic package in which intangible services are attached to tangible products to finish various industrial activities in the whole product life cycle (Zhang et al., 2012);
- PSS and customer-centricity – PSS elements interact to fulfil customer needs (Manzini and Vezzoli, 2003); Interactions between providers and customers are extended to different phases of the PSS lifecycle (Cavalieri & Pezzotta, 2012);
- PSS is a new business model: A system of products, services, supporting networks, and an infrastructure designed to be: competitive, satisfy customer needs, and have a lower environmental impact than traditional business models (Mont, 2002);
- PSS and sustainability - PSS-based strategies have positive environmental effects compared to traditional business models (Roy, 2000); PSS integrates a product and service into a system that provides the same necessary functionally but reduces the environmental impact (Baines et al., 2007).

Some PSS definitions further emphasize "selling the use" instead of "selling the product", focusing on value-in-use and not on the value of "ownership". As discovered in (Tukker, 2015), the focus is not on the products and services per se but on the customer needs and how organizations define their responsibilities and relationships in a more complex value-chain system.

The field of PSS research is multidisciplinary, addressing various problem areas and research directions as discovered in the reviews, focusing on PSS and economy "servitization" (Baines et al., 2007), PSS and business model innovation (Reim et al., 2015); PSS and sustainability (Tukker, 2015), and others. From the very beginning, the PSS covers a large area of sustainability, prospects for sustainable economic growth, and a circular economy (Fernandez et al., 2020). It is interesting to outline how the literature review by (Barravecchia et al., 2021) ranges the accumulated literature within the main PSS topics – PSS design, sustainability, environmental, social impact, and servitization (Table 1). Figuring out that some of the papers can address few topics, the sum is exceeding 100%.

Table 1. Main PSS topics, Barravecchia et al., (2021)

PSS design (methods/tools)	21%
PSS environmental and social impact	18%
PSS and servitization process	13%
Sustainable PSS	11%
PSS Business models	10%
PSS performance analysis	10%
PSS requirements analysis	9%
Industrial PSS	9%

The emerging concepts of Service Science (Maglio & Spohrer, (2008) and Service-Dominant logic (Lush & Vargo, 2008, Vargo & Lush, 2016) contribute to defining a general model of service systems research, mainly focusing on service innovations and service ecosystem view (Vargo et al., 2017). Thus, PSS can be explored from socio-technical systems theory view (Baxter & Sommerville, 2011). There, the socio-technical system consists of two components (subsystems):

- **the technical subsystem**, covering assets such as machines and equipment, as well as processes and tasks that are responsible for the conversion of input resources into outputs;
- **the social subsystem**, made up of people (such as customers and employees), structured in groups, roles, and functions.

Most of the papers and current research is focused on applying the PSS systems in the context of Industry 4.0 and manufacturing digitalization. There many PSS

involve intersectoral interactions and configurations of vertically or horizontally integrated agents (Sony & Naik, 2020). This way, the rising digitalization poses new challenges for designing and delivering complex value offerings, combining services and tangible products within interconnected socio-technical systems.

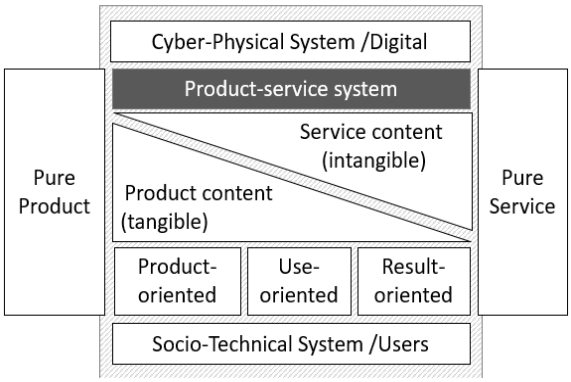


Figure 2. Product-Service System model, adapted from (Tukker, 2004)

Further, Cyber-physical systems represent the convergence of PSS with digital technologies. Figure 2 explore the general model of PSS, CPS and company value creation. The CPS definition is "an intelligent system, connecting the physical and the digital/cyber world through influence and control using sensors and actuators" (Martin et al., 2019). The authors further reflect its capabilities to recognize and act upon signals from the environment, connecting through machine-to-machine communication. The definition of CPS by (Gölzer et al., 2015) states that CPS are "able to communicate with each other, to detect their environment, to interpret available data and to act on the physical world". CPS combines a cyber part that provides computational capabilities, sensors for collecting data, and actuators. It is more often characterized as a technical system than a socio-technical one. However, considering the emergence of increasingly intelligent CPS machines that can further support autonomous actions, multi-criteria decision making, and complex interaction between humans and technology, and reflecting that these terms are often used within different disciplines, CPS should be further reflected when defining smart service, smart service system, and smart products capabilities (Martin et al., 2019).

2.3 Digital Platform Enterprises

The fast-accumulating literature in the field of "digital platforms" finds many overlapping definitions, mixing interchangeably concepts of "digital platforms" with "platform ecosystems", "multi-sided platforms", or "digital platform enterprises". Stepping on the building blocks metaphor, the authors perceive the definition of a digital platform as "a building block that provides an essential function to a technological system and serves as a foundation upon which complementary products, technologies, or services can be developed", formulated in (Spagnoletti et al., 2015).

Digital platforms specifics

Considering some of the main classifications of platforms, such as (Evans & Gawer, 2016), the digital platform enterprises can be identified as:

- **Transaction platform**: focusing to excel at demand and supply matching of products and/or services;
- **Innovation platform:** providing backbone technology or facility for other firms to design their technologies, products, or services;
- **Integrated/Hybrid platform:** facilitating transactions and forming an ecosystem of stakeholders;
- **Investment platforms,** supporting a portfolio and investment strategies.

The main functions of the digital platforms are to facilitate (1) effective matching and easy interactions (and transactions), (2) efficient innovation backbone, speed and scale of innovation, ROI, and profitability; (3) ecosystem building, connecting a pool of investors, partners, buyers and relevant stakeholders (Figure 3).

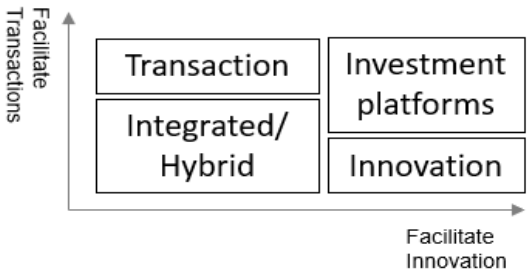


Figure 3. Digital platforms differentiation, adapted from the classification of (Evans & Gawer, 2016)

3 Conceptual Model Building Blocks

Considering the concepts behind digital platforms, our aim is to define both organizational and technical elements, addressing the main challenges behind company management. On one hand, agile methodologies come from software industry and explore the easiness to build-on and excel prototypes until reaching the final version and output. On the other hand, the product-service-system thinking come from manufacturing industry and focus on delivering physical artefacts. Therefore, the authors propose to combine the both elements (fig.3) in order to develop a conceptual model for better understanding the digital platform enterprise way of thinking, management

processes, value creation mechanisms and platform model of functioning.

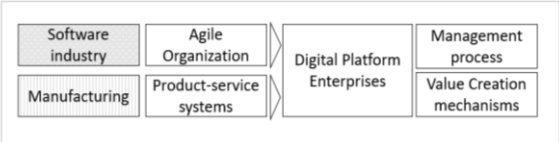


Figure 4. Building blocks of the Digital Platforms Enterprise – the authors' conceptualization

This way, the building blocks of a digital platform (fig.4) cover the social sub-system - the business and management level and the technological sub-system - components, products and artefacts, both corresponding to relevant agile and PSS concepts.

3.1 Agile levels

There are three different levels of agility that an organization can perceive and achieve, considering the benefits and impact on organizations, business strategies and goals (Highsmith, 2014):

- **Agility at the operational level** represents the company's transition to an agile organization. There is *operational agility* that does not significantly affect the expansion of the organization and responsiveness or other changes within the organization, despite that some of the staff have become more agile. Still, project, product, and managers are the same and do the same job they did before introducing the agile practices in the organization.
- **Agility at the strategic level** – is focused on achieving responsiveness in the organization. This level is achieved mainly by the organization's managerial and leadership staff, which strategically contributes to the company's overall agility.
- **Agility at the portfolio level** – focuses on the agile practices introduced in the organization and the strategic decisions that the management has to make when applying them for high-value and vital projects. Therefore, applying these practices takes place strategically and is targeted only for essential projects. Thus, all other projects may not apply agile practices. The management of the agility at the portfolio level represents the bridge between enterprise agility and the successful execution of agile projects.

There are existing requirements for agile organizations, such as agile office, governance, facilities, and technology, that the organization must consider to achieve the desired level of agility in the organization.

3.2 Types of PSS

The Product service system (PSS) is traditionally classified into three different types (Baines et al., 2007) as *Product-oriented PSS (POS), Use-oriented services (UOS), and Result-oriented services (ROS).* A brief description of the terms is provided next, based on (Baines et al., 2007).

Product-oriented PSS. This involves the sale of a product in a traditional way, in which the customer retains the ownership of the product, although including some associated services. For instance: after-sales service to ensure functionality and durability of the product (maintenance, repair, reuse, recycling, training, and consulting). In this case, the company is motivated to introduce a PSS to minimize costs of use throughout the product lifecycle;

Use-oriented services. The producer owns the product, and sells the use or functions of a product (e.g., leasing, sharing, etc.). In this case, the company is motivated to create a PSS to maximize the use of the product, to meet the demand, to extend product life, and reuse materials;

Result-oriented services. This involves the sale of a result or competence rather than a product. Companies offer a mix of personal services, and when the result is delivered in a product, the PSS provider maintains its ownership, and the customer only pays for the results.

4 Conceptual Model of a Digital Platform Enterprise

The platform business models differentiate from pipeline business mainly by network effects, increasing value by adding more users, self-reinforcing cycle of growth; digital technologies, capturing, transmitting, and monetizing data over the Internet; economy of scale, and lock-in effects (Gawer, 2014). Although platform businesses cannot be applied to all companies and all sectors, many companies can improve their model of functioning by applying both agile and service-oriented methodologies.

In this perspective, before considering how and if the firm should implement a Digital Platform Enterprise business logic, it can address the following elements on strategic, development, and portfolio level, as described in Table 3. The conceptual model, proposed in Table 3, builds on the service science priorities, extending and improving the model of (Ostrom et al. 2010).

First, on strategic level, the company has to reconsider the focus on customer and customers' needs. Design thinking models offer appropriate management tools and persona to outline customer needs, considering ethnographic research methods and focusing on user expectations and real situation. In this model, servitization thinking and PSS models can

improve value-creation mechanisms transferred to multiple-platform experiences.

Second, on development level, agile methodologies improve company business processes, considering innovation, service design, and organizational channels. Third, on portfolio level, companies can expand the platform model of PSS for further personalization and adaptation, adequately transforming customer data into customer-centred offerings.

Table 3. Digital Platform Enterprise building block model, adapted by (Ostrom et al., 2010)

Strategic Level- PSS	Development level - Agile	Portfolio Level - Platform
Improve the organization's ability to offer PSS successfully.	Stimulate Innovation	Effectively brand and sell PSS through personalized user-centered offerings;
Improve well-being through transformative PSS	Enhance Service Design	Enhance PSS user experience by enhancing contextualizatio n and personalization approaches
Create and maintain a relevant organizational culture and mindset	Optimize service networks and value chains around customer experiences;	Measure and optimize PSS value offerings by adopting complex KPI;

Furthermore, companies adopting digital platform business thinking have to:

- Use socio-technical system paradigms to build the PSS ecosystem and drive PSS innovations;
- Adopt appropriate business models, supported by new PSS and CPS technology offerings (e.g., combining smart services, cloud computing, machine learning, artificial intelligence, and others);
- Accelerate the adoption and acceptance of new service-oriented PSS technologies (connect to and invest in ecosystem and innovation providers);
- Capture and deliver PSS service-oriented information for real-time decision making;
- Enable and accelerate mobile access for consumers and employees;
- Enhance privacy and security to protect service consumers, employees, companies, and society;
- Enable agility and integration through PSS architecture and appropriate technology platforms.

5 Discussion and future work

The identified two fundamental building blocks of the Digital platform enterprises – agile organizations and product-service systems, enable companies to decompose organizational and technical challenges and to reflect on practical steps to address platform thinking in their digital transformation strategies.

On a strategic level, agile principles allow companies to easily extend their capacity for structural changes on a horizontal and vertical level. For example, many companies add new departments to respond to the current business dynamics, such as Data Analysis, Digital and Innovation departments, and Customer Experience departments, leading to vertical expansion and additional management staff. However, if not appropriately managed, this expansion can lead to substantial difficulties most commonly related to (Ozbayrac, 2020):

- Lower rate of adjustment to market dynamics;
- Inefficient decision-making;
- Lack of cross-functional collaboration;
- Less innovation or emerging new revenue streams;
- Talent attrition and low employee engagement, and so on.

The positive effects of the business transformation to an agile organization include (but are not limited to): employee engagement, resource optimization, time to market (Ozbayrac, 2020), data analytics use for extracting valuable knowledge from data (Dankov and Birov, 2018), and so on. In the same time, a summary of the main benefits of adopting agile organizations include: (1) resource optimization; (2) improved cross-functional collaboration; (3) higher employee satisfaction; (4) faster realization of benefits. Thus, neither adopting agile methodologies or PSS model will automatically lead to better business performance.

Table 4. Main advantages of adopting PSS for customers and companies, based on (Baines et al., 2007)

For Customers
• More customized offers and higher quality; • New functionalities and PSS combinations; • End-of-life tasks transferred to the producer.
For Companies
• New market opportunities and competitive advantages; • Alternative to standardization and mass production; • Improvement in the total value delivered for the customer through increasing service elements; • Access to data about product performance during the use phase.

On the development level, adopting PSS allows companies to consider new technologies as supporting infrastructure, improving value offerings, business model structuring, innovation potential, boosting ecosystem and networking effects and circle economy, green and sustainability principles. Table 4 lists the main advantages of PSS based on the target groups (Baines et al., 2007).

Future work

The study represents only the first step toward for researching implementation and deployment of digital platform enterprise. It represents a theoretical model, conceptualizing some relevant topics of platform management and platform value creation. Further work will contribute for evaluating existing platform companies and exploring more examples and case studies, contributing to fine-tune the model and its elements.

Furthermore, the current complexity and dynamics in the domain of Digital Platform Enterprise deployment require further attention on other issues such as digital platforms business models, digital marketing and customer-centricity, digital platform architecture management, platform interoperability, all of them to be further addressed in the future DEMO innovative learning solutions.

The final aim is to facilitate companies and experts from the CEE to better understand and approach digital platforms design and development in a consistent and relevant to the theory and practice way, further advancing their companies' digitalization and servitization.

6 Conclusion

The present research outlined the theoretical concepts behind agile organizations and product-service systems as building blocks of digital platform. In the near future it can be expected further advancement of PSS and CPS, evolving to increasingly "smart" (product) service systems, focusing on further personalization, adaptation, and customization of product-service offerings, fuelling innovations and new digital platform business models.

Thus, this paper provides a conceptual model for adopting digital platform thinking, improving understanding of the main platform value-creation mechanisms. This can help both the young companies and start-ups, defining new value-creation offerings, and the established businesses, benefiting from new platform-based user-oriented and experience-focusing offerings, further preparing for digital economy transformation.

Acknowledgments

The current research is partially supported by project DEMO - Digital Platform Enterprise, funded by EU, ERASMUS+ programme, project No. 2021-1-RO01-KA220-HED-000027576.

References

Baines, T. S., Lightfoot, H. W., Evans, S., Neely, A., Greenough, R., Peppard, J., ... & Wilson, H. (2007). State-of-the-art in product-service systems. Proceedings of the Institution of Mechanical Engineers, Part B: journal of engineering manufacture, 221(10), 1543-1552.

Barravecchia, F., Franceschini, F., Mastrogiacomo, L., & Zaki, M. (2021). Research on product-service systems: topic landscape and future trends. Journal of Manufacturing Technology Management.

Baxter, G., & Sommerville, I. (2011). Socio-technical systems: From design methods to systems engineering. Interacting with computers, 23(1), 4-17.

Beedle, M., Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin R., Schwaber, K., Sutherland, J. and Thomas, D. (2001). Manifesto for Agile Software Development. Online: https://agilemanifesto.org/history.html Last Accessed: 19.04.2022

Berkovich, M., Leimeister, J. M., Hoffmann, A., & Krcmar, H. (2014). A requirements data model for product service systems. Requirements Engineering, 19(2), 161-186.

Beverungen, D., Müller, O., Matzner, M., Mendling, J., & Vom Brocke, J. (2019). Conceptualizing smart service systems. Electronic Markets, 29(1), 7-18.

Cavalieri, S., & Pezzotta, G. (2012). Product–Service Systems Engineering: State of the art and research challenges. Computers in industry, 63(4), 278-288.

Dankov, Y., Birov, D. (2018). General Architectural Framework for Business Visual Analytics. In: Shishkov, B. (eds) Business Modeling and Software Design. BMSD 2018. Lecture Notes in Business Information Processing, vol 319. Springer, Cham. https://doi.org/10.1007/978-3-319-94214-8_19

De Reuver, M., Sørensen, C., & Basole, R. C. (2018). The digital platform: a research agenda. Journal of Information Technology, 33(2), 124-135.

da Costa Fernandes, S., Pigosso, D. C., McAloone, T. C., & Rozenfeld, H. (2020). Towards product-service system oriented to circular economy: A systematic review of value proposition design approaches. Journal of Cleaner Production, 257, 120507.

Gaiardelli, P., Resta, B., Martinez, V., Pinto, R., & Albores, P. (2014). A classification model for product-service offerings. Journal of cleaner production, 66, 507-519.

Gawer, A. (2014). Bridging differing perspectives on technological platforms: Toward an integrative framework. Research policy, 43(7), 1239-1249.

Goedkoop, M. J., Van Halen, C. J., Te Riele, H. R., & Rommens, P. J. (1999). Product service systems, ecological and economic basics. Report for Dutch Ministries of environment (VROM) and economic affairs (EZ), 36(1), 1-122.

Gölzer, P., Cato, P., & Amberg, M. (2015, May). Data Processing Requirements of Industry 4.0-Use Cases for Big Data Applications. In ECIS.

Gummesson, E. (2008). Extending the service-dominant logic: from customer centricity to balanced centricity. Journal of the Academy of Marketing Science, 36(1), 15-17.

Highsmith, J. (2014). Adaptive Leadership: Accelerating Enterprise Agility. Addison-Wesley, ISBN 978-0-13-359844-5

Highsmith, J. (2010). Agile Project Management: Creative Innovative Products. Upper Saddle River, NJ: Addison-Wesley.

Highsmith, J. (2002). Agile Software Development Ecosystems. Boston: Addison-Wesley.

Lusch, R. F., & Vargo, S. L. (2008). The service-dominant mindset. Service science, management and engineering education for the 21st century (pp. 89-96). Springer, Boston, MA.

Maglio, P. P., & Spohrer, J. (2008). Fundamentals of service science. Journal of the academy of marketing science, 36(1), 18-20.

Manzini, E., & Vezzoli, C. A. (2002). Product-service systems and sustainability: Opportunities for sustainable solutions. UNEP-United Nations Environment Programme.

Martin, D., Hirt, R., & Kühl, N. (2019). Service Systems, Smart Service Systems, and Cyber-Physical Systems—What's the difference? Towards a Unified Terminology. 14th International Conference on Wirtschaftsinformatik, February 24-27, 2019, Siegen, Germany

Manzini, E., & Vezzoli, C. (2003). A strategic design approach to develop sustainable product service systems: examples taken from the 'environmentally friendly innovation'Italian prize. Journal of cleaner production, 11(8), 851-857.

Mont, "Clarifying the concept of product –service system," Journal of Cleaner Production., vol. 10, pp. 237–245, 2002.

Ostrom, A. L., Bitner, M. J., Brown, S. W., Burkhard, K. A., Goul, M., Smith-Daniels, V., Demirkan, H. & Rabinovich, E. (2010). Moving forward and making a difference: research priorities for the science of service. Journal of service research, 13(1), 4-36.

Ozbayrac, G. (2020). Enterprise Agility: A Practical Guide to Agile Business Management. CRC Press, Taylor & Francis Group, ISBN: 1032139099, London, UK.

Perkin, N. and Abraham, P. (2021). Building the Agile Business through Digital Transformation. Second Edition, Kogan Page Limited, ISBN:9781789666540, NY, USA

Roy, R. (2000). Sustainable product-service systems. Futures, 32(3-4), 289-299.

Reim, W., Parida, V., & Örtqvist, D. (2015). Product–Service Systems (PSS) business models and tactics–a systematic literature review. Journal of Cleaner Production, 97, 61-75.Zhang, F., Jiang, P., Zhu, Q., & Cao, W. (2012). Modeling and analyzing of an enterprise collaboration network supported by service-oriented manufacturing. Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 226(9), 1579-1593.

Information Quality of Digital Services in Higher Education: Higher Order Construct Analysis

Tea Mijač, Mario Jadrić, Maja Ćukušić
Faculty of Economics, Business and Tourism
University of Split
Cvite Fiskovića 5, Split, Croatia
{tea.mijac, mario.jadric, maja.cukusic}@efst.hr

Abstract. *Digital transformation in higher education leads to new digital services for stakeholders, particularly students. The importance of service quality and quality assurance is particularly emphasised in the shift to the "digital-first" or "digital-only" mode. In that context, the paper focuses on the multidimensionality of digital services in higher education and the subdimension of information quality based on the feedback from over 561 students in Croatia.*

Keywords. Information quality, higher education, digital transformation, higher-order-construct.

1 Introduction

Digital transformation (DT) represents the most significant paradigm shift in Higher Education (HE) (Biedermann, Kalbfell, Schneider, & Drachsles, 2019). It manifests in novel concepts and solutions, not only as a technology-driven approach but also resulting in new digital business models, digital user engagement, and so on (Moreira, Ferreira, & Seruca, 2017). Consequently, an increase in the number of digital services in HE has been notable (Julião & Gaspar, 2021; Pasini, Estevez, & Pesado, 2019). DT in HE enables modular growth, development, and access to resources, ensuring that institutions become more efficient in teaching and non-teaching activities (Fisher, 2006; Xiao, 2019).

DT is not only about *moving* to digital, but due to the *marketisation* of HE, those digital services must also be aligned with end-users needs (Díaz-Méndez, Paredes, & Saren, 2019; Elsharnouby, 2015; Ng & Forbes, 2009). Higher Education Institutions (HEIs) strive to foster relationships with stakeholders by delivering timely, accurate information and, most importantly, aligning with the stakeholders' interests (Pinho, Franco, & Mendes, 2022). Even though there are numerous stakeholders in HE (Kettunen, 2015), due to increased competition among HEIs, students' needs are particularly highlighted (Nguyen, Shirahada, & Kosaka, 2012).

In the DT projects, HEIs should place a particular emphasis on evaluating and measuring issues linked to service quality (Annamdevula & Bellamkonda, 2016). It is well argued that *service quality* (primarily focusing on information) is a crucial aspect of *student satisfaction* (Annamdevula & Bellamkonda, 2016; Sultan & Wong, 2019) which has a mediator role between service quality and student loyalty (Annamdevula & Bellamkonda, 2016).

Reliable, accurate, and up-to-date information is cogitated as *information quality*. In the public sector, the information quality of digital services proved significant for citizen satisfaction (Rasool & Warraich, 2018). Similarly, information quality in HE has been emphasised (Sultan & Wong, 2019). Because of DT, consuming information online has become commonplace in HE (Chen & Chengalur-Smith, 2015), and as the information is now consumed online, its quality needs to be assured both online and offline.

Although HEIs began their digital transformation (Grajek, 2020), *there is still a lack of empirical studies* in this area (Khalid, Ram, & Khalee, 2018). Even though information quality has a significant role in HE and is considered a complex multidimensional phenomenon, it has not been fully understood (Ge, Helfert, & Jannach, 2011). Also, for evaluating digital services' characteristics, studies have generally modelled information quality as first-order constructs (Forsgren, Clay, Wang, & Durcikova, 2016). It has been implied that future studies should inspect it as the second-order construct and explore it in a different context (Rasool & Warraich, 2018).

In line with this, the study presented herein aimed to *model the construct of information quality of digital services in HE and determine preferred characteristics (subdimensions) of those digital services from students' perspective, i.e. from the end-users side.*

The paper is structured as follows. Section 2 presents the role of information quality in the DT of HE. Section 3 describes the research methodology, while section 4 presents the study's empirical results. Highlights and the main contributions of the study are

presented in section 5, together with future research directions.

2 Theoretical Background

Technology development is changing the operations in the public sector and contributing significantly to DT in HEIs (Sandkuhl & Lehmann, 2017; Santos, Batista, & Marques, 2019). DT presupposes the transition from the traditional value creation and delivery to the mass application of different technologies (Sandkuhl & Lehmann, 2017; Verina, Titko, & Shina, 2019). In that regard, it is expected that HEIs will focus even more on developing and implementing ICTs that represent appropriate channels for communication with their stakeholders (Santos et al., 2019). HEIs are using these technologies both to improve their performance and adapt to changes in a society that is increasingly driven by technology (Rodrigues, 2017) since the structures and processes in HE are just as susceptible to far-reaching changes as the ones in the business environment (Gilch et al., 2019). A recent survey revealed that 13% of HEIs are engaged in DT projects, 32% are developing DT strategies, and 38% of HEIs are still discovering DT as a concept (Grajek, 2020). At the same time, 17% of institutions do not invest time in DT.

Another transformation impetus is a competitive market where HE institutions compete for students and staff nationally and internationally (Díaz-Méndez et al., 2019; Sutjarittham, Gharakheili, Kanhere, & Sivaraman, 2019). Embracing *marketisation* in HE leads to adopting concepts and practices such as advertising, branding, monitoring and managing customer satisfaction in a customer-oriented approach (Díaz-Méndez et al., 2019; Elsharnouby, 2015; Ng & Forbes, 2009). Customer-oriented, i.e. the student-oriented approach, has become particularly important (Elsharnouby, 2015; Fader, 2012). It aims to look at the product or service from students' perspectives to understand their needs (Levaniemi, 2015). Ensuring the student-oriented approach by using the technology is becoming a norm (Balyer & Oz, 2018; Bond, Marín, Dolch, Bedenlier, & Zawacki-Richter, 2018; Seres, Pavlicevic, & Tumbas, 2018) and a key factor in creating a successful HE environment (Arkhipova, Belova, Gavrikova, Pleskanyuk, & Arkhipov, 2020). To maintain their role in the development of society, HEIs must respond in a timely and adequate manner to these challenges.

Providing sufficient and reliable *information* to students can impact their perception of HEIs' service quality (Sultan & Wong, 2019). The findings show that information is more statistically significant than experience as the antecedents of service quality.

There are four types of digital services in HE, and for each, information quality is of utmost importance (Pasini et al., 2019):

1. Informational digital services – delivering useful information to the HE community; for example, notifying students of important dates.
2. Digital certification services – verification and confirmation of different student statuses; for example, proof that a student has obtained a diploma.
3. Digital records services – these are digital services that generally control the presence/attendance. For example, checking the presence of students, teaching or administrative staff working hours.
4. Digital authorisation services – refers to authorisation and digital signatures.

Despite the inevitable importance, numerous business initiatives have been postponed or cancelled in various industries, citing poor information quality as the main reason (Ge et al., 2011). Subdimensions often used to measure information quality in HE are (1) understandability (Aldholay, Abdullah, Isaac, & Mutahar, 2019; Efiloğlu Kurt, 2019; Xu & Du, 2018), (2) reliability (Cidral, Oliveira, Di Felice, & Aparicio, 2018; Daghan & Akkoyunlu, 2016), (3) accuracy (Aldholay et al., 2019; Çelik & Ayaz, 2021; Fadelelmoula, 2018), (4) up-to-dateness (Çelik & Ayaz, 2021; Efiloğlu Kurt, 2019), (5) timeliness (Daghan & Akkoyunlu, 2016; Lwoga, 2013), (6) relevancy (Çelik & Ayaz, 2021; Efiloğlu Kurt, 2019; Hasan et al., 2018; Xu & Du, 2018) and (7) completeness (Al-Azawei, 2019; Aldholay et al., 2019; Efiloğlu Kurt, 2019; Hagos, Garfield, & Anteneh, 2016).

3 Methodology

Based on the literature review, an instrument was prepared. An online survey tool (Lime Survey) was used for data collection with the participants, students from Croatia.

Higher-Order Constructs (HOC) analysis using Covariance-Based Structural Equation Modelling (CB-SEM) has been applied. It contributes to each dimension, as opposed to connecting all manifest variables with a single first-order construct when it can be said that the explanation of the construct is incomplete (Mohammadi & Mahmoodi, 2019). A minimal sample size is achieved (Hair, Black, Babin, & Anderson, 2010), and the HOCs have been evaluated the same way first-order constructs are (ibid). HOC analysis has been done using the IBM Amos software.

The sample was stratified and made of students from ten HEIs in Croatia in the field of social sciences (economics) who use any digital services for non-teaching-related activities. The survey began in May 2021 and lasted until June 2021. Nine responses (outliers) are removed from the total number of completed survey questionnaires (N=551).

4 Results

Information quality has been modelled as a second-order construct. The initial list of proposed subdimensions has been changed. *Completeness* has been removed, as discriminant validity was not achieved. Therefore, a second model analysis was performed with six subdimensions. Confirmatory Factor Analysis (CFA) was used to check validity and reliability, and the results are presented in Figure 1. In Table 1, Factor Loadings (FL), Average Variance Extracted (AVE), and Composite Reliability (CR) are presented (Hair, Babin, & Krey, 2017). Calculated FLs range from 0.723 to 0.917 (each above the referent value of 0.60). Calculated AVE indicators exceed the value of 0.50, as suggested. Finally, CRs for each first-order construct are over 0.70. Therefore, convergent validity is confirmed. Discriminant validity is also confirmed as all calculated correlations are lower than 0.85 (Brown, 2006). Additionally, the square root of AVE surpasses calculated correlations (Afthanorhan, Ghazali, & Rashid, 2021).

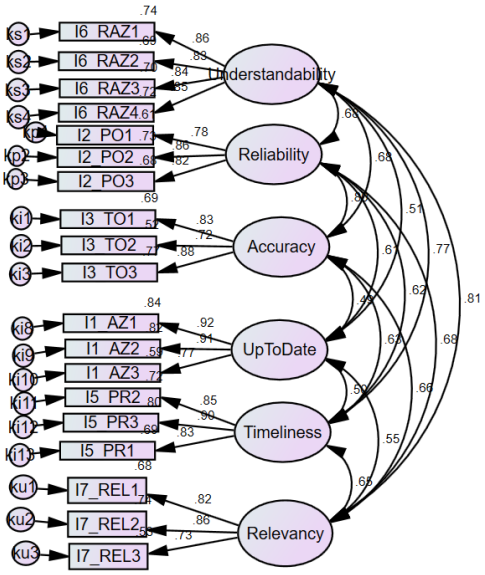


Figure 1. CFA model

Table 1. FL, AVE and CR

Item	First-order construct	FL	AVE	CR
I3_SQ001	Accuracy	0.83	0.662	0.854
I3_SQ002	Accuracy	0.723		
I3_SQ003	Accuracy	0.88		
I7_SQ001	Relevancy	0.824	0.648	0.846
I7_SQ002	Relevancy	0.859		
I7_SQ003	Relevancy	0.727		
I2_SQ003	Reliability	0.824	0.672	0.860
I2_SQ002	Reliability	0.856		
I2_SQ001	Reliability	0.778		
I5_SQ002	Timeliness	0.849	0.739	0.894
I5_SQ003	Timeliness	0.895		
I5_SQ001	Timeliness	0.833		
I6_SQ001	Understandability	0.862	0.713	0.909
I6_SQ002	Understandability	0.829		
I6_SQ003	Understandability	0.838		
I6_SQ004	Understandability	0.848		
I1_SQ001	Up-to-date	0.917	0.753	0.901
I1_SQ002	Up-to-date	0.908		
I1_SQ003	Up-to-date	0.77		

Results of calculated discriminant validly are in Table 2.

Table 2. Discriminant validity

	Understan.	Reliability	Accuracy	UpTo Date	Timeliness	Relevancy
Understandability	**0,84**					
Reliability	0,68	**0,82**				
Accuracy	0,68	0,80	**0,81**			
UpToDate	0,51	0,61	0,49	**0,87**		
Timeliness	0,77	0,62	0,63	0,50	**0,86**	
Relevancy	0,81	0,68	0,66	0,55	0,65	**0,81**

After validity and reliability are confirmed, path coefficients are calculated. Figure 2 shows standardised path weights, and all weights are significant.

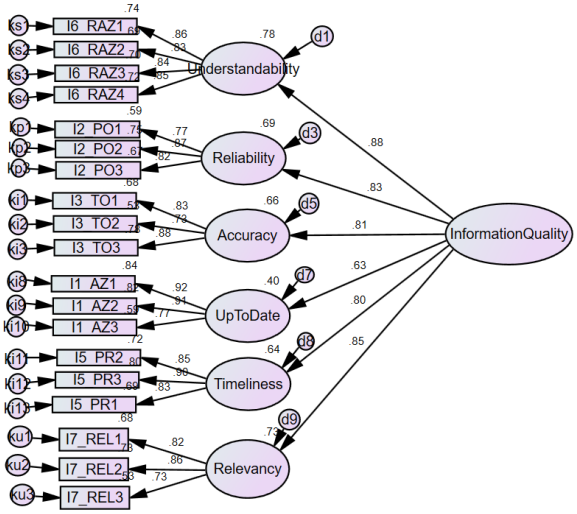


Figure 2. Path model with standardised estimations

Several relevant fit indices have been calculated to determine if the model fits the data (Hair et al., 2010; Marsh & Hocevar, 1985). Chi-square and freedom degrees ratio (CMIN/DF) is 3.67, Comparative Fit Index (CFI) is 0.953, Goodness-of-fit (GFI) is 0.903, and Root mean square error of approximation (RMSEA) 0.070. All calculated indices are within proposed referent values.

5 Discussion and conclusion

First-order constructs, i.e., subdimensions of information quality for digital services in HE, have been identified and confirmed as foreseen based on the theoretical foundations of the study.

By modelling information quality as a second-order construct, the multidimensional nature of the observed construct is confirmed. Hence, it would be more appropriate for future research to avoid using the *overall information quality* and instead measure it as a second-order construct. For that matter, this study provides a better understanding of the characteristics of digital services regarding information quality in the context of HE. Namely, the research study conducted with end-users confirmed that *accuracy*, *relevancy*, *up-to-dateness*, *timeliness*, *reliability*, and *understandability* are essential information quality subdimensions. Initially, *completeness* has been tested as a subdimension of information quality. However, although convergent validity was confirmed and the model fit indicators met the threshold, *discriminant validity was not confirmed.* The proposed model with six subdimensions, however, meets all the thresholds.

Further to confirming the six subdimensions, the findings have other implications relevant to the user-oriented design of digital services in HE. As some authors point out – the information for end-users in the information systems of HEIs is often not aligned with their needs (Husain, Syafar, Sabara, & Syafar, 2019). Therefore the results of the study could be used as guidelines for HEIs regarding specific aspects of information system (re)design in the process of DT.

However, the study has notable limitations. Participants are students from Croatia, so it would be interesting to conduct the same research in other countries and compare the results. Also, since the study focuses on the end-users perspective, it would be helpful to inspect and compare the results from the perspective of service providers and identify possible differences. This study focused on digital services unrelated to learning and teaching; however, since DT represents a much bigger perspective, future research should take a more comprehensive approach.

Acknowledgements

The research is supported by the Croatian Science Foundation [grant number IRP-2017-05-7625].

References

Afthanorhan, A., Ghazali, P. L., & Rashid, N. (2021). Discriminant Validity: A Comparison of CBSEM and Consistent PLS using Fornell & Larcker and HTMT Approaches. *Journal of Physics: Conference Series*, *1874*(1), 012085. https://doi.org/10.1088/1742-6596/1874/1/012085

Al-Azawei, A. (2019). What drives successful social media in education and e-learning? A comparative study on Facebook and moodle. *Journal of Information Technology Education: Research*, *18*, 253–274. https://doi.org/10.28945/4360

Aldholay, A., Abdullah, Z., Isaac, O., & Mutahar, A. M. (2019). Perspective of Yemeni students on use of online learning: Extending the information systems success model with transformational leadership and compatibility. *Information Technology and People*, (May). https://doi.org/10.1108/ITP-02-2018-0095

Annamdevula, S., & Bellamkonda, R. S. (2016). The effects of service quality on student loyalty: the mediating role of student satisfaction. *Journal of Modelling in Management*, *11*(2), 446–462. https://doi.org/10.1108/JM2-04-2014-0031

Arkhipova, M. V., Belova, E., Gavrikova, Y. A., Pleskanyuk, T. N., & Arkhipov, A. N. A. (2020). *Reaching Generation Z. Attitude Toward Technology Among the Newest Generation of School Students*.

Balyer, A., & Oz, O. (2018). Academicians' views on digital transformation in education. *International Online Journal of Education and Teaching (IOJET)*, *5*(4), 809–830. Retrieved from http://iojet.org/index.php/IOJET/article/view/441/295

Biedermann, D., Kalbfell, L., Schneider, J., & Drachsles, H. (2019). *Stakeholders attitudes towards digitalisation in higher education institutions*. 187–192. https://doi.org/10.18420/delfi2019

Bond, M., Marín, V. I., Dolch, C., Bedenlier, S., & Zawacki-Richter, O. (2018). Digital transformation in German higher education: student and teacher perceptions and usage of digital media. *International Journal of Educational Technology in Higher Education*, *15*(1), 1–20. https://doi.org/10.1186/s41239-018-0130-1

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: The Guilford Press.

Çelik, K., & Ayaz, A. (2021). Validation of the Delone and McLean information systems success model: a study on student information system. *Education and Information Technologies*, (0123456789). https://doi.org/10.1007/s10639-021-10798-4

Chen, Y. H., & Chengalur-Smith, I. (2015). Factors influencing students' use of a library Web portal: Applying course-integrated information literacy instruction as an intervention. *Internet and Higher Education*, *26*, 42–55.

https://doi.org/10.1016/j.iheduc.2015.04.005

Cidral, W. A., Oliveira, T., Di Felice, M., & Aparicio, M. (2018). E-learning success determinants: Brazilian empirical study. *Computers and Education*, *122*, 273–290. https://doi.org/10.1016/j.compedu.2017.12.001

Daghan, G., & Akkoyunlu, B. (2016). Modeling the continuance usage intention of online learning environments. *Computers in Human Behavior*, *60*, 198–211. https://doi.org/10.1016/j.chb.2016.02.066

DeLone, W. H., & McLean, E. R. (2016). Information Systems Success Measurement. *Foundations and Trends® in Information Systems*, *2*(1), 1–116. https://doi.org/10.1561/2900000005

Díaz-Méndez, M., Paredes, M. R., & Saren, M. (2019). Improving society by improving education through service-dominant logic: Reframing the role of students in higher education. *Sustainability (Switzerland)*, *11*(19). https://doi.org/10.3390/su11195292

Efiloğlu Kurt, Ö. (2019). Examining an e-learning system through the lens of the information systems success model: Empirical evidence from Italy. *Education and Information Technologies*, *24*(2), 1173–1184. https://doi.org/10.1007/s10639-018-9821-4

Elsharnouby, T. H. (2015). Student co-creation behavior in higher education: the role of satisfaction with the university experience. *Journal of Marketing for Higher Education*, *25*(2), 238–262. https://doi.org/10.1080/08841241.2015.1059919

Fadelelmoula, A. A. (2018). The impacts of the quality dimensions of the ERP system on the realisation of the fundamental business objectives and perceived usefulness. *International Journal of Enterprise Information Systems*, *14*(4), 89–107. https://doi.org/10.4018/IJEIS.2018100107

Fader, P. (2012). *Customer centricity* (2nd ed.). Wharton.

Forsgren, N., Clay, P. F., Wang, X., & Durcikova, A. (2016). The integrated user satisfaction model: Assessing information quality and system quality as second-order constructs in system administration. *Communications of the Association for Information Systems*, *38*(1), 803–839. https://doi.org/10.17705/1CAIS.03839

Ge, M., Helfert, M., & Jannach, D. (2011). Information quality assessment: Validating measurement dimensions and processes. *19th European Conference on Information Systems, ECIS 2011*.

Gilch, H., Beise, A. S., Krempkow, R., Muller, M., Stratmann, F., & Wannemacher, K. (2019). *Digitalisierung der Hochschulen: Ergebnisse einer Schwerpunktstudie für die Expertenkommission Forschung und Innovation*.

Grajek, S. (2020). *How Colleges and Universities Are Driving to Digital Transformation Today*. Retrieved from https://er.educause.edu/-/media/files/articles/2020/1/er20sr214.pdf

Hagos, Y., Garfield, M., & Anteneh, S. (2016). Measurement factors model for e-learning systems success. *Proceedings - International Conference on Research Challenges in Information Science*, *2016-Augus*. https://doi.org/10.1109/RCIS.2016.7549361

Hair, J.F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis*. Pearson Prentice Hall.

Hair, Joseph F., Babin, B. J., & Krey, N. (2017). Covariance-Based Structural Equation Modeling in the Journal of Advertising: Review and Recommendations. *Journal of Advertising*, *46*(1), 163–177. https://doi.org/10.1080/00913367.2017.1281777

Hasan, M., Maarop, N., Samy, G. N., Mohammad, R., Azmi, N. F., Hassan, N. H., & Ghaffar, N. A. (2018). Measurement tool for assessing research information management system success. *Journal of Telecommunication, Electronic and Computer Engineering*, *10*(3–2), 53–57.

Husain, H., Syafar, F., Sabara, E., & Syafar, S. (2019). Data and information quality framework development: Proposed for Indonesia higher education. *Proceedings of the 33rd International Business Information Management Association Conference, IBIMA 2019: Education Excellence and Innovation Management through Vision 2020*, (April), 7648–7653.

Julião, J., & Gaspar, M. C. (2021). Lean thinking in service digital transformation. *International Journal of Lean Six Sigma*, *12*(4), 784–799. https://doi.org/10.1108/IJLSS-11-2020-0192

Kettunen, J. (2015). Stakeholder relationships in higher education. *Tertiary Education and Management*, *21*(1), 56–65. https://doi.org/10.1080/13583883.2014.997277

Khalid, J., Ram, B., & Khalee, M. (2018). *Promising digital university : a pivotal need for higher education transformation Promising digital university : a pivotal need for higher education transformation Jamshed Khalid *, Braham Rahul Ram , Mohamed Soliman , Anees Janee Ali , Muhammad Khaleel a*. *12*(January), 13. Retrieved from https://www.researchgate.net/publication/324844227%0APromising

Levaniemi, V. (2015). Knowing the Difference

Between Customer Focused and Customer Centric is a Serious Competitive Advantage. Retrieved from https://www.linkedin.com/pulse/knowing-difference-between-customer-focused-centric-ville-levaniemi/

Lwoga, E. T. (2013). Measuring the success of library 2.0 technologies in the African context: The suitability of the DeLone and McLean's model. *Campus-Wide Information Systems*, *30*(4), 288–307. https://doi.org/10.1108/CWIS-02-2013-0011

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin*, *97*(3), 562–582. https://doi.org/10.1037/0033-2909.97.3.562

Mohammadi, F., & Mahmoodi, F. (2019). Factors Affecting Acceptance and Use of Educational Wikis: Using Technology Acceptance Model (3). *Interdisciplinary Journal of Virtual Learning in Medical Sciences*, *In Press*(In Press). https://doi.org/10.5812/ijvlms.87484

Moreira, F., Ferreira, M. J., & Seruca, I. (2017). DIGITAL TRANSFORMATION - A NEW CHALLENGE FOR HIGHER EDUCATION: A CURRICULUM PROPOSAL. *10th Annual International Conference of Education, Research and Innovation*, 8124–8131.

Ng, I. C. L., & Forbes, J. (2009). Education as service: The understanding of university experience through the service logic. *Journal of Marketing for Higher Education*, *19*(1), 38–64. https://doi.org/10.1080/08841240902904703

Nguyen, T. D., Shirahada, K., & Kosaka, M. (2012). A consideration on university branding based on SDL (Service Dominant Logic): The lens of stakeholders' value co-creation. *2012 9th International Conference on Service Systems and Service Management - Proceedings of ICSSSM'12*, 779–784. https://doi.org/10.1109/ICSSSM.2012.6252346

Pasini, A., Estevez, E., & Pesado, P. (2019). Assessment Model for Digital Services provided by Higher Education Institutions. *ACM International Conference Proceeding Series*, 468–477. https://doi.org/10.1145/3325112.3325268

Pinho, C., Franco, M., & Mendes, L. (2022). Influence of web portals on management support and university performance: an information quality perspective. *Quality and Quantity*. https://doi.org/10.1007/s11135-021-01305-5

Rasool, T., & Warraich, N. F. (2018). Does quality matter: A systematic review of information quality of E-government websites. *ACM International Conference Proceeding Series*, 433–442. https://doi.org/10.1145/3209415.3209473

Rodrigues, L. S. (2017). Challenges of digital transformation in higher education institutions: A brief discussion. *Proceedings of the 30th International Business Information Management Association Conference, IBIMA 2017 - Vision 2020: Sustainable Economic Development, Innovation Management, and Global Growth*, *2017-Janua*(November 2017), 4490–4493.

Sandkuhl, K., & Lehmann, H. (2017). Digital Transformation in higher education-The role of enterprise architectures and portals. *Lecture Notes in Informatics*, 12.

Santos, H., Batista, J., & Marques, R. P. (2019). Digital transformation in higher education: The use of communication technologies by students. *Procedia Computer Science*, *164*, 123–130. https://doi.org/10.1016/j.procs.2019.12.163

Seres, L., Pavlicevic, V., & Tumbas, P. (2018). Digital Transformation of Higher Education: Competing on Analytics. *INTED2018 Proceedings*, *1*(March), 9491–9497. https://doi.org/10.21125/inted.2018.2348

Sultan, P., & Wong, H. Y. (2019). How service quality affects university brand performance, university brand image and behavioural intention: the mediating effects of satisfaction and trust and moderating roles of gender and study mode. *Journal of Brand Management*, *26*(3), 332–347. https://doi.org/10.1057/s41262-018-0131-3

Sutjarittham, T., Gharakheili, H. H., Kanhere, S. S., & Sivaraman, V. (2019). Realising a Smart University Campus: Vision, Architecture, and Implementation. *International Symposium on Advanced Networks and Telecommunication Systems, ANTS*, *2018-Decem*. https://doi.org/10.1109/ANTS.2018.8710084

Verina, N., Titko, J., & Shina, I. (2019). Digital Transformation in Higher Education: Academic and Administrative Staff Perception. *34th IBIMA Conference: 13-14 November 2019, Madrid, Spain*.

Xu, F., & Du, J. T. (2018). Factors influencing users' satisfaction and loyalty to digital libraries in Chinese universities. *Computers in Human Behavior*, *83*, 64–72. https://doi.org/10.1016/j.chb.2018.01.029

Factors influencing project manager's ability to deliver a software engineering project

Rudi Pleteš, Martina Tomičić Furjan, Željko Dobrović
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{rpletes, martina.tomicic, zeljko.dobrovic}@foi.hr

Abstract. *Mostly pushed by the maturity of information and communication technology (ICT) and readiness of customers to embrace the change, the software engineering organizations (the Organizations) have built their businesses, engineering teams and managers around culture, working processes and methodological frameworks scaled mainly for onsite work, accompanied with managerial (authoritative or delegator) influence style. With already present lack of available software engineers (SE's), new era for software engineering managers begun with the need for response to change called "COVID-19". Relying on their experience gathered in years long projects, authors are analyzing challenges that the Software Engineering Project Manager (PM) is faced while delivering software project with a group of SE's, working together in geographically distributed environment.*

Keywords. software engineering project management; team culture; distributed teams; agile software development, servant leadership.

1 Introduction

The goal of any project is to deliver added value (product or a service) defined by the Contract, constrained with three variables – timeframe, budget, and quality. According to Mir and Pinnington (Mir and Pinnington, 2014), traditional project management approaches which exclusively pursue the mentioned success criteria are considered ineffective, and despite advancement in project management processes, tools and systems, project success will not significantly improve. From the other side, project cannot be delivered without people (team) working on it. Bonebright (Bonebright, 2010) analyzed that there are more than 250 different models and theories that were being used in team development practice, whereby one of the most common theory is the Tuckman's five stages of group development – forming, storming, norming, performing, and adjourning. Each of these stages requires from PM to adapt the leadership and managerial style, approach skills for both – the SE team and every individual. The idea is to get the best from each individual and from the group.

The Research Problem. Today's requirements for SE's grew and current destabilization of workforce requirements within global economy under crisis, highly influenced by COVID-19 situation, has changed the way engineering organizations are running their businesses. That is why the whole project's organization is dependent on the PM to assemble all required resources (physical, financial, human, political etc.). With lack of available potential employees on the Market and related high cost, lack of opportunity to retain people, Customers more often rely on their subcontractors like software companies to support their business in both – technical and industry specific processes and knowledge. Being faced with the fact "I can work anywhere in the world" and "I can be paid a lot more that I earn now", SE is lured into everyday recruiting process. Since teamwork, team culture and task management is more complex in distributed working environment, PM should not rely only on technical knowledge of SE's – their skills and motivation are more and more important. There is much that is not understood in context of what PM is doing on the software engineering project. Different stakeholders see their picture and no one, even the PM itself, cannot perceive the whole picture with ease. The software engineering PM profession started to be drastically complex with two new changes that came in parallel – (a) lack of SE's on the market and (b) presence of remote working environment.

Expected outcomes and results of this paper are: (a) To contribute to a better understanding of how remote work and insufficient work power influence the complexity of a software project management and (b) Face the authors' collected experiences from software engineering projects with recent theoretical background and provide the baseline for better high-level understanding of SE project management and its complexity.

The significance of the study for audiences. This paper can help any stakeholder around software engineering project or development team to better understand the complexity of PM environment.

Researchers interested in software engineering project management can find this paper useful for further research studies related to software engineering project management, team culture, distributed teams, agile software development and servant leadership.

The Case. This Research covers analysis of factors influencing PM's ability to lead a distributed SE team, working in strong project-oriented matrix organization, delivering a complex custom-made software, based on time and material or fixed scope contractual obligations, for an external Customer (mid-sized, large-sized enterprise or a governmental institution).

Recent studies that have addressed the research problem are covered trough indictive design in **Section 3**. **Section 4** covers **a narrative review of factors** influencing Project Manager's ability to deliver a software engineering project, strongly relying on subjective experience of authors, gathered in years long project activity. Discussion is identifying the gap between theory covered and practice presented, while conclusion has been made as a wrap-up, giving indications on research limitations and future research possibilities.

2 Research methodology

Social constructivism is an interpretive framework whereby individuals seek to understand their world and develop their own subjective meanings that correspond to their experience (Creswell, 2009). By designing the research using the **Social Constructivist** worldview, authors rely as much possible on their own views of the situation being studied, bringing personal value into the research. To contextualize the problem, authors have highlighted their own understanding and knowledge related to factors influencing PM to deliver SW engineering project, collected through their extensive practical experience working as a PM or a team member.

To challenge authors' subjective interpretation of knowledge and related observations, **theoretical background** of latest scientific findings is given beforehand and compared in the discussion section.

Theoretical background has been made through the analysis of scientific findings related to the topic of this research, sourced in the bibliographic database Web of Science (WoS). Initial WoS search (criteria: *"software engineering project management (Topic) and 2020-2022 (Year Published) and Article (Document Type) or Computer Science, Software Engineering (Web of Science Categories)"* has been made on 29th of June 2022 and returned 62 results. Further analysis has been made through sorting of titles according to citation (highest first) and further subjective perception of authors. Further analysis of articles that are rated within upper 25% most quality papers (first category quartile; Q1), resulted in 13 articles cited at least one time. Additional reduction has been made, based on following criteria

1. Article needs to be published in a scientific journal,
2. the content needs to be relevant for this Research,
3. Article must be cited at least one time,
4. At least one author of article that meets above mentioned criteria's, must have issued more than 20 scientific publications in total AND Sum of times cited is larger than 100 AND the research area must be primarily related to Computer Science (Software Engineering) AND the author must have more than 5 years of scientific research experience.

Concluding, six articles were chosen for inclusion. Although the number is very limited, the purpose of the theoretical background is not to give a comprehensive analysis on the topic, but rather to focus onto most recent findings by relevant authors.

3 Theoretical background

Šmite et al (Šmite et al., 2021) are exploring barriers between national and organizational cultures and their implications to the agile software engineering. In their work, authors are recommending that the organization needs to pay special attention to the process of onboarding new team members into the cultural norms and the term "trust" as a prerequisite for shared leadership in agile team is mentioned, so it is impossible in everchanging environment, with insufficient number of people on the project (of required seniority, knowledge, skills etc.) to expect that the team can work without senior leadership and management**.** Self-organizing agile teams need to be navigated and secured from their senior colleagues with more widen experience, knowledge, and skills – experts that can coach SE's, encouraging them, managing team's boundaries, dealing with unexpected problems within a team, managing the expectations of stakeholders outside of a team. Concluding, to support a SE, the atom of a 'container' called software engineering project team in matrix organizational structure, the organizations need PM with new skills that were not needed through the history of software engineering work.

Shastri et al., (Shastri et al., 2021) presented and described the role of a Scrum Master (SM) in agile projects. They analyzed the role of the Scrum Master (SM) in agile projects and described following group of SM's work: (a) process adapting, (b) negotiating, (c) mentoring, (d) protecting, (e) coordinating, and (f) facilitating. Their study for the first time presents a multifaceted study of the multiple dimensions of the SM role (facilitation, mentoring, negotiation, process adaption, coordination, and protection).

Chapetta and Travassos (Chapetta and Travassos, 2020) proposed framework to support researchers to observe and control SE's productivity. They mentioned Conway's law ("A system reflects the organizational structure that built it") as a project concern in context of communicational barrier among SE's that might influence the productivity of SE team and concluded

that the number of SE's and their variation over time affects the software development productivity. Same authors are also referencing to a Curtis' law ("Good designs require deep application domain knowledge") and concluding that the rate at which software projects get inputs and produce outputs is concerned with the domain knowledge of their SE's. Long-term effect of increasing the project duration promotes slight losses in software development productivity, because software development is creative work and highly human-related, so the complexity of leading, managing, and motivating SE's can be found as complex set factors that can influence the productivity of individual SE in environment where change is defined as constant.

By analyzing internal dynamics of the different stakeholders, Russo (Russo, 2021) is identifying critical success factors and their relations to a large-scale agile project and its success, and the role of SE's as one of the most crucial aspects to lead to a SE project.

Mendes et al. (Mendes et al., 2021), analyzed the relationship between decision-making and personality, and concluded that SE's personality can influence the decision-making style.

4 Analysis of factors influencing ability to lead a software engineering project

Existing technological and communicational readiness, society, government, business environment and world economy trends faced the software engineering companies with lack of adequate manpower able to support customer from different segments – consumer, industrial and government. The more stakeholders are present in project environment, the more complex is to run the project, satisfy communicational flows and manage all the expected outcomes. The new working paradigm defined a SE working in a distributed environment as a key factor that is influencing the organizational design, where new standards and expectations are strongly influencing the organizational and project culture and values, their ability to motivate individuals and deliver successful software projects (shown in Fig. 1).

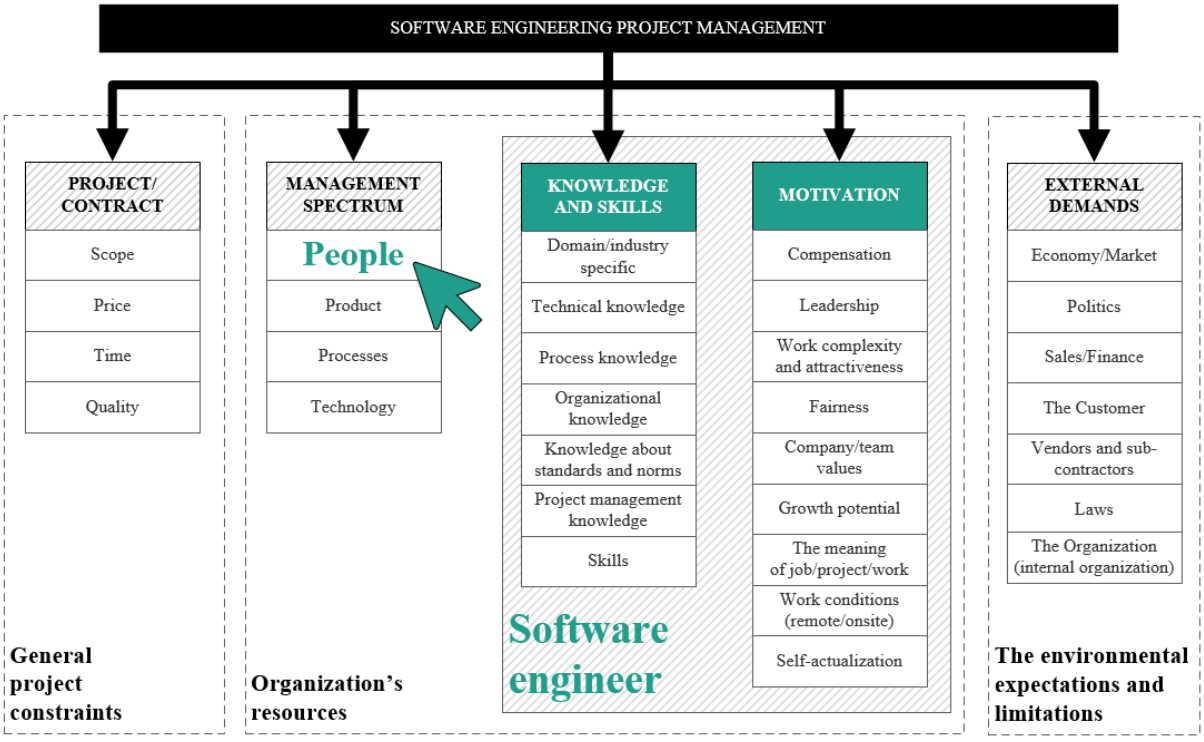


Figure 1: High-level decomposition of software engineering PM focus areas

Current economic trends brought to the software engineering projects a high level of uncertainty. Due to these high economic demands, Project can be at risk if PM is not skilled to motivate engineering team, positively affecting team spirit and values, create safe environment based on loyalty, respect, and trust. Considering that every individual has its unique motivational factors, PM needs to adapt its personality, communication style, motivational approaches, and decision-making style to every individual SE uniquely.

4.1 Perceiving the stakeholders' expectations of project outcome

If PM understands and influences interactions and expectations of a power structure (political awareness and sensitivity) around and within the project environment, he will more easily run the project.

Stakeholder is any individual or a group, internally within the Organization or externally, that have an interest from positive project outcome. All about the project is about meeting the objective and subjective expectations of stakeholders. Constrained with their knowledge, available information and attitude, and influence level, different stakeholders are shaping different perception, understanding and expectation of project outcome. If stakeholders would interpret successful project as a project that is delivered within contractual obligations (objective constraints), strictly defined by the contract – project could be treated as successful. But there are objective (e.g. budget and time related) and subjective (e.g. stakeholders' expectations, unpredictable risks related to economic and market flows) constraints that are shaping the perception of the term "successful project" and are influencing the project outcome. To satisfy the overall, stakeholders' objective and subjective perception on whether the project is delivered successfully, these expectations will need to be embraced, challenged, and managed. In case these additional expectations are accepted, formally or not, they are negatively influencing on PM's ability to control the Project outcomes. More importantly, these expectations require from PM, SE team and other stakeholders to spend more time for informal project communication, always questioning subjective thoughts and their value to the Project.

If the Project is initiated to fulfil goals that are outside of the contractual obligations (e.g., strategic goals), it will be exposed to additional political, business and commercial assessments, expectations, influences, and pressures from different stakeholders. Below are analyzed the most important stakeholders that have influence on SE project environment:

- **Sponsor** is the most powerful representative of a client, and most often this is the Chief Executive Officer (CEO). His sponsorship is defined by the fact that he pays the software, so he can use his power to influence all other stakeholders.
- **Product Owner (PO).** In traditional organizational setups PO represents a Key Customer's expert, but in case the Organization is using any of agile approaches (e.g., Scrum) to deliver software solutions, PM authorizes an individual within the SE team to act as a PO. PO is in this case an individual, SE who is focused on product (domain knowledge), who can prioritize value of a different requirements and is skilled to communicate with both – business and technically sound individuals in project environment. Usually, subjected individual is the SE functionally working as a senior business analyst. As such, Product Owner represents "the right hand" of a PM and by this, he must have strong project management skills.
- **User** is the individual who uses the software.
- **Supervisory Board** is group of individuals from Customer's (e.g., CEO, Procurement, etc.) and Supplier's (e.g., CEO, Sales, etc.) side, headed by the Project Sponsor, and being responsible and authorized to bring project decisions that are outside/above PM's power and influence.
- **Project Manager (PM).** PM is central communication, leadership, managerial, people oriented, financially and business aware, decision-making role on a project. According to the contractual obligation, PM is most responsible for an overall project delivery, oversees many functional areas through the software project delivery and development processes.
- **Functional Manager (FM).** FM's are boosting SE's technical and functional knowledge of SE's. They are taking care the Project's technical and functional environment is aligned with the organization's strategy, processes, methodological approaches etc.
- **Software engineering team (SE team).** SE team, also called the development team, are a group of individuals having dispersed technical, functional, domain, process, methodological, project and other knowledge and skillset. They are using knowledge and skills to accomplish the same goal – deliver the software product/service defined by the project.
- **Software engineer (SE).** This construct represents an individual, a team member, the core of s Project, working on a delivery of a software product/service in role of business analyst, developer, quality assurance specialist (tester), database administrator, consultant etc.
- **Third parties.** With economic demands that created the SE's as the most valuable asset, third parties (e.g., vendors, subcontractors, agencies renting SE's etc.) and their SE's that are engaged on the Project by the Organization, have a big influence on the Organization, the Project and other individuals working on the Project.

4.2 Perceiving and contextualizing influencing factors within project environment

Through this paper are analyzed factors that are influencing perceived core values, labor costs, subjective and objective factors shaping the SE's

perception and environment, attitude, and eagerness to contribute to an SE team. Organizations being able to manage career development path of SE's, motivate and stimulate people – will be able to retain knowledge, retain customers, projects and potentially gain comparative advantage on the market. Leading distributed SE team and running multiple projects at the same time, with all analyzed objective and subjective factors influencing the project environment requires from PM to be a leader, manager, and multidisciplinary expert with extensive skillset. PM's working as a project tracker will fail.

In other to meet the stakeholders' expectations, PM needs to determine and manage different factors that shape the complexity of project environment. Since PM is constantly switching from one context to another, below given factors are on purpose not grouped in meaningful containers, so the reader could more easily perceive different levels of thinking in project management, demanding and constantly changing environment, skills needed for influencing others without formal authority, dependencies and different interdisciplinary approaches in decision making, different knowledge areas and skillsets, communication, and coordination complexity:

a. **The commercial Contract with the Customer.** If the Contract is not favorable for the Organization due to any reason (e.g., bad decisions, assumptions, or estimations through the sales process), PM and project team will have difficulties from the early start of the Project. Potentially, the PM, SE team or key SE can feel unmotivated because they will need to work harder to compensate bad decision of others. If not leaded and motivated with high level of commitment and support, SE's could easily leave the Project and/or Organization and leave the project with additional risks (more unsatisfied people on the project, new onboarding processes consuming already overallocated SE's, etc.).
b. **Project's commercial value is determining the stakeholder's involvement.** Managers are driven by financial aspects, so larger expected revenue and potential income imply that the interest of stakeholders for project outcome will grow. If the Project is worth more, it means the PM will need to spend more time communicating and managing expectation of stakeholders on different layers.
c. **The Organization's ability to timely ensure people for the Project.** PM and the Organization should be able to spend time to the socialization, functional, domain, project, organizational and cultural training of a new SE on a Project. Due to the lack of SE's on the market, bad sales process and related sales prices, inadequate motivational and stimulation measures – the Organizations are not able to deliver projects within given constraints.
d. **Number of SE's working on a project.** Communication requirements of all SE's individually, and SE team, are unique, need to be managed and satisfied. The bigger the SE team is, the more complex is to manage the project.
e. **Ability of the Organization to support projects.** To be able to compete on the market, larger organizations strive to be compliant with different industrial standards. Also, the bigger the organization is, the processes are more complex, it takes longer to satisfy all political, personal, procedural, objective, and subjective, formal, and informal procedures to bring project decision. Today PM's need to bring fast decision and they need supportive organizations, their functional/technical teams, that can respond to project requirements and constant changes.
f. **Expectations of Agile mindset and its perceived values**. Agile teams should be self-organizing, have a shared leadership, with complementary knowledge and skills to run the project. Often a lot of decision cannot be taken within the engineering team without the intervention or support of senior management outside of the team. Since the PM is the managerial role with the most dispersed knowledge, the PM by nature supports the SE team.
g. **Number of externalized SE's on the Project.** Interest for remote workers is shaped with the fact that there is not sufficient manpower on global level – and the whole world is trying to attract as many SE professionals as they can get. This reshaped the SE financial expectations, so some organizations are not able to compete a market and retain their SE's, and that is why they are going into commercial contracts with third-party companies that are renting SE's. The more externalized team member's Project have, the PM will be exposed to additional activities – commercial and legal work, administrative work, additional communication and coordination, objective and subjective expectations of additional stakeholders with different influence levels, potentially hidden costs, coordination meetings and making things happen, coordinating different roles, teams, vendors, seniorities, potentially bad contracts, track and report about service level, satisfying product requirements, get to know new external processes that need to be satisfied and aligned with project's and organizational standards and requirements, manage invoicing, contractual obligations, purchase orders, timesheets, motivate people and align different cultural and other expectations. On the top of that, managing the team of individual SE's from different entities, in distributed environment is,

additional level of complexity for all included stakeholders. All these can consume a lot of PM's time and due to that, strongly influences to project complexity and potentially satisfaction of all stakeholders.

In case externalized SE is unsatisfied or demotivated, PM has limited authority over his motivational or stimulative factors. Also, the Organization is not able to fully control how external SE's values and culture are influencing the Organization's and Project culture and team values.

h. **Domain knowledge.** SE's working for an Organization and its projects are having specific domain knowledge about customers and their processes. With high level of fluctuation of SE's working on a Project, organizations and project teams are losing its productivity, expertise, and safe environment for other team members. In addition to these findings, by subcontracting or partially externalizing project activities or tasks, the Organization is further risking by consciously losing this unique and hard to get knowledge. This may result in (a) inability of the Organization to ensure high-level quality of a maintenance services for the Customer, (b) losing the Customer's trust and perception of value, (c) further losing their SE's, and consequently (d) losing business.

i. **Different commercial types of projects.** Depending on type of contract (time and material based versus fixed scope based), from PM is expected to be skilled, experienced, and able to plan and use different approaches that will ensure successful project delivery. Different project types require different knowledge and skillsets.

j. **The Project duration.** The longer projects are more complex since the motivation of every individual SE is decreasing over time.

k. **The world became a big workforce village.** SE's, PM's and other stakeholders became aware that the corona crisis extended the boundaries of what the whole world though is acceptable working environment. The new perception, where SE's can work from anywhere in the world, for an employer from any country in the world created new challenges for the organizations and the way they plan and deliver projects. Engineering organizations across the world are redesigning their working processes to support remote SE workers and this will strongly influence on the SE market and the way we plan and deliver software engineering projects.

l. **The maturity of SE team.** If SE's have a previous positive experience working together on a Projects, sharing core values, believing each other, they will more easily became productive and efficient as a Team (or a Group). Constant fluctuations of team members and dispersed values within a team can negatively influence outcome of a Project. In case of fluctuation of people, new onboarding processes can be a burden to a most valuable SE's within a team and if they are repeatedly mentoring/coaching new colleagues – they will feel (a) unsafe because managers cannot control fluctuation, (b) angry/sad because they will feel that managers do not see their burden with every change, etc.

m. **Alignment of Organization's seniority levels with global standards and expectations**. Assumption that senior SE is the individual only with great technical knowledge is high project risk. Individual labeled as „senior SE" needs to have good understanding of overall project environment, organizational and project processes, understanding of different methodological approaches, good interpersonal communicational skills, being able to write project/product documentation on its own , being able to communicate with different stakeholders, with different managerial, he needs to understand business acumen, always questions decisions being made etc.

n. **The answers of Organization to high global demand for SE workforce.** Due to the exponential use of IT in all segments of human lives, the global economy does not have enough SE's to satisfy all the needs. Companies are using different tactical approaches to retain and/or grow their businesses and projects, like:

- Retention policies ensuring financial stimulation like bonus policies, extra benefices (company vehicle 24/7, extra healthcare, financing sport activities, etc.). Also, FM's are having more 1:1 conversation with their SE's, searching for feedback and trying to use more supportive attitude, continuously improve working atmosphere.
- Attraction policies like opening additional offices to expand reach on new markets, more being focused on students and cooperation with universities, using aggressive organization's self-promotion trough dispersed set of marketing activities (e.g., helping community), hiring attractive and pushy recruiters (and sometimes even non-competent) to attract SE's on social and other virtual and real-world environments etc.
- Process improvement can be done through activities like strategic alliances with other organizations to optimize and improve efficiency and productivity.

If dominant, the fluctuation of SE's within the team can be a burden to a most valuable SE's who are expected to repeatedly act as a mentor and/or a coach to a new colleague or to work overtime (to compensate lost time), to override bad coding, correct bugs or redesign the specification. Due to the long recruiting process for a new SE if not supported correctly, existing SE's can suffer from burnout/ exhaustion, team members can feel unsafe, and the project may be endangered etc.

o. **Geographically distributed SE team.** SE teams working in geographically distributed (also: hybrid, virtual or remote) environment are more exposed to:
   - divergent thoughts about project goals and methods being used,
   - emotional exhaustion,
   - health issues,
   - less teamwork,
   - unavailability of team members,
   - bad decision making of senior management, insufficient empathy etc.,
   - need for supportive management,
   - language barriers,
   - additional administration,
   - bad onboarding processes,
   - wrong subjective perception of Project and Organization's expectations,
   - hard to control working hours,
   - the PM's and Organization's ability to manage and control the need of introverted and extroverted SE's,
   - feeling a need for psychologically safety and perception of trusted environment,
   - overexposed to virtual communication and lack of interpersonal, face to face communication and social gatherings,
   - cultural differences,
   - hard to establish team or project culture and to align with Organizational culture, etc.

p. **Geopolitical changes**, like the one started on 24th February 2022, by the Russian invasion on Ukraine, redefined already disrupted economic and SE workforce flows. These changes are additionally supporting and shaping people fluctuation flaws.

q. **Compensation for the lack of the matrix organizational structure.** If working in a matrix organization structure, and managers that have the lack of leadership skills – SE can feel unsafe and frightened that his FM will not be able to bring the right decision related to his professional career, financial stimulation etc. In this case, from PM is expected to support the SE, motivate him, and create safe environment around individual.

r. **Organization's Management and its understanding of software engineering PM environment.** If the Organization's CEO and other managers are having low level of understanding of the complexity and specifics of SE project management, this can negatively influence on PM, who will need spend a lot more energy more thoroughly argumenting objective and subjective influencers to a Project.

s. **Alignment of organizational and project goals.** The organization need to have and constantly analyze and improve metrics that will ensure the alignment of new opportunities and projects to its mission, vision, strategy, organizational processes, norms, standards, and core values. Organizations often generate a lot of sales opportunities and dive into projects just to gain revenue. The goal of organizations should challenge and questioning the goals and expected outputs of every potential project with organization's mission and vision. Often, the PM is the role that have the highest strategic overview over the Project and could be able to align Organizational and Project goals and culture. But, due to the demanding environment, often nobody takes care of these alignments.

t. **Self-actualization in society and working environment**. Self-actualization is perception of an individual rated in terms of how valued he is, his role, job, employer, financial compensation, and stimulation compared to other individuals he knows in private or business world.

u. **Competencies and abilities of a PM.** Due to the sum of all uncertainties related to software engineering project, most of the projects fail. That is the reason why it is important to choose the right PM, the one that is competent and capable leading and managing all aspects of project environment. PM needs to be able to think about different aspects of equally treatment for all team members, taking into consideration seniority level, gender, functional role, private needs and obligations (e.g., family, friends etc.).

5 Discussion

Theoretical background showed that relevant authors individually focused **their areas of interest on dispersed topics** – (a) factors influencing software development productivity (Chapetta and Travassos, 2020), (b) understanding cultural barriers in distributed SE projects and correlation with agile organizational approaches (Šmite et al., 2021), (c) demystifying the role of a Scrum Master in agile projects (Shastri et al., 2021), (d) study of a large-scale Agile transformation in a mission-critical environment, where stakeholders' commitment was a critical success factor (Russo, 2021) and (e) relationship between decision-making

style and personality within the context of software project development (Mendes et al., 2021).

A research gap, characteristic for the disruptive environment today's organizations are facing, is identified in the fact that neither all authors together neither any of them separately did not cover a dispersed high-level analysis of subjective and objective parameters influencing PM's ability to lead a geographically distributed SE team working on a complex software engineering project. the published research represents valuable knowledge and asset in specific, narrow areas of scientific research. For instance, the SM's role and related work that Shastri et al. (Shastri et al., 2021) analyzed, represents only one dimension of PM work – the one related to providing servant leadership (the one who serves) to a SE team. But, PM needs to be an advocate of all stakeholders (including the SE team), and be able to represent different opinions, motivating, guiding, and supporting stakeholders to act in accordance with project goals. In case the Organization has its other Agile processes in place but did not ensure the Scrum Master role due to any reason, PM will have additional obligations within SE team.

6 Conclusion

By covering different factors and aspects influencing the complexity of leading software project, and facing them with existing literature findings, authors of this paper provided the baseline for better high-level understanding of SE project management and its complexity. Also, this Research additionally contributes to the topic by addressing the influence of remote working environment and insufficient SE workpower to a complexity of a software project management.

Research limitations and future work. This paper is presenting a preliminary research study, based on limited number of focused, most relevant, recently published papers and a subjective opinion of the authors, enriched by objective experience findings. Authors plan to use the results presented as a baseline for future research that will be extended by (a) interviewing software engineering project managers and different stakeholder types identified in this paper and (b) bringing the interview results together with existing literature reach, authors' experience and conclusions being made through this paper. Also, authors will deeply analyze how governmental and law institutions are creating safe environment for remote workers and analyze how this could contribute to standardize the requirements, knowledge areas and skills that SE working in a distributive environment should have.

References

Bonebright, D.A., 2010. 40 years of storming: a historical review of Tuckman's model of small group development. Human Resource Development International 13, 111–120. https://doi.org/10.1080/13678861003589099

Chapetta, W.A., Travassos, G.H., 2020. Towards an evidence-based theoretical framework on factors influencing the software development productivity. Empir Software Eng 25, 3501–3543. https://doi.org/10.1007/s10664-020-09844-5

Creswell, J.W., 2009. Research design: qualitative, quantitative, and mixed methods approaches, 3rd ed. ed. Sage Publications, Thousand Oaks, Calif.

Mendes, F., Mendes, E., Salleh, N., Oivo, M., 2021. Insights on the relationship between decision-making style and personality in software engineering. Information and Software Technology 136, 106586. https://doi.org/10.1016/j.infsof.2021.106586

Mir, F.A., Pinnington, A.H., 2014. Exploring the value of project management: Linking Project Management Performance and Project Success. International Journal of Project Management 32, 202–217. https://doi.org/10.1016/j.ijproman.2013.05.012

Russo, D., 2021. The Agile Success Model: A Mixed-methods Study of a Large-scale Agile Transformation. ACM Trans. Softw. Eng. Methodol. 30, 1–46. https://doi.org/10.1145/3464938

Shastri, Y., Hoda, R., Amor, R., 2021. Spearheading agile: the role of the scrum master in agile projects. Empir Software Eng 26, 3. https://doi.org/10.1007/s10664-020-09899-4

Šmite, D., Moe, N.B., Gonzalez-Huerta, J., 2021. Overcoming cultural barriers to being agile in distributed teams. Information and Software Technology 138, 106612. https://doi.org/10.1016/j.infsof.2021.106612

Automatic End-to-End Decomposition and Semantic Annotation of Laws Using High-Performance-Computing and Open Data as a Potential Driver for Digital Transformation

Charalampos Alexopoulos
Department of Information and Communication Systems Engineering
University of the Aegean |
Karlovassi, 83200 Samos, Greece
alexop@aegean.gr

Igor Pihir, Martina Tomičić Furjan
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin
{ipihir, mtomicic}@foi.unizg.hr

Abstract. *This paper is dealing with automatic end-to-end law analysis conducted by decomposition and semantic annotation, by using the high-performance-computing and government open data. Legal data and law texts are a category of open data and thus they possess the potential to unlock digital innovation and transformation capacity in governments and businesses, regarding the development of new, better, and more cost-effective services for citizens. For that reason, they can be recognized as a potential digital transformation driver. This research presents a baseline for automation of decomposition and annotation with process and service elements developed for utilization on high-performance-computing infrastructure based on government laws open data and gives insights on how the results of it can initiate digital transformation.*

Keywords. Digital Transformation, Laws Decomposition and Annotation, Automation, Text mining, Hight-Performance-Computing, Open Data.

1 Introduction

The legislation of each state is very important to be open and available online in order for every part of society to have free and unhindered access (Fulton, 2011). Financial and legal services of academia and research institutions are constantly dealing with laws and their interpretation. All new and already existing procedures have been designed or will be designed based on the analysis of legal documents. In addition, significant advancements in the 'legal informatics' research field are observed since governments have started to promote the development of legal information systems (Casanovas et al., 2016). Disruptive technologies like text mining could be used for the development of novel legal information systems and services that able to automatically extract information from unstructured legal data.

This opens possibilities for innovation and digital transformation driven by these newly available legal data. Transformation can thereby be achieved through one or more new business concepts: new service, new product, business model, etc. (Tomičić Furjan, Tomičić-Pupek and Pihir, 2020) enabled by the results of performed laws analysis. "Digital transformation (DT) can be seen as an approach which assumes a radical change of doing business, driven highly by today's necessity to change and adapt to the digital age. This change is mostly made by creating new business models, which define the way how an organization provides value to the customers" (Hrustek, Tomičić Furjan & Pihir, 2019). Data retrieved or created from laws, analysed as proposed in this paper and provided as open data could initiate this change and become a potential DT driver. Open Data Process model for open data paradigm is described in paper (Hrustek, Tomičić Furjan & Pihir, 2020)

The main objective of this paper is to present the service towards the automatic decomposition (creation of legal data) and semantic annotation (based on specific and well-recognized standards) using text mining techniques on laws. It is meant to find a way to build and test a fully automated process towards the creation of big, linked, open legal data.

The Legal documents have specific characteristics which differentiate them from the "normal daily-used" documents due to their length and complexity (Nguyen et al., 2011). The complexity of the legislation is identified in the comprehending of the legal terminology, in other words, legal language, in combination with the existence of correlations as references among laws or other kinds of legal documents. There are different types of legislative documents such as laws, presidential decrees and ministerial decisions. Every type consists of a specific

form including specific structural elements (components), but in this paper we are dealing only with laws. Each reference may amend one or more different laws, and, in many cases, an in-depth search of previous laws is required to detect which component is in force. Such a service of creating and providing big, linked, open legal data (BLOLD) should be supported by a powerful high-performance-computing infrastructure in order to be able to analyse the large amount of data structures and the correlations among them. Based on the results of this service additional and added-value services could be developed. This way, we envision the facilitation of public access for all types of users to open legal data accessed and served via a Legal Web platform in a customizable, structured, intuitive and easy-to-handle way (Lachana et al., 2020). Furthermore, the data will be available for all countries complying with the developed standard and developed text techniques.

The added-value services could be described as follows: i) research through legal corpora, analyzing the alignment of national legislation with EU legislation, ii) comparing national laws which target the same life events, iii) analyzing the references to European legislation by national laws, iv) analyzing related laws within the same Member State, v) timeline analysis for all legal acts, vi) visualization of the progress and current status of a specific national or European piece of legislation and vii) sentiment analysis towards new legislation.

The paper starts with the introduction to the problem, followed by a methodology section and furthermore development of the idea across a literature review in form of a state-of-the-art chapter. The main part of the paper is oriented to the metadata scheme for decomposition and semantic annotation, HPC processing infrastructure, and how the decomposition and semantic annotation automated process is working, and which services are used/developed to support it. It ends with concluding remarks, acknowledgment, and bibliography.

2 Methodology

Research presented in this paper is part of larger project initiative streaming to automate the decomposition and semantic annotation of laws in such a way, so that the purpose and meaning of laws could be reused, compared, and automatically processed in process and semantic way. Research uses literature review, process and data modelling, metamodeling, datamining and high-performance computation to extract data and automation metadata-based annotation.

This research is oriented to problem description, literature review to position the problem, and architecture development for automation of decomposition and semantic annotation process. Research is conducted on results of previously ended hundreds of processes analysis and by finding and classifying the paths to automated processing of legal documents which in the end can trigger digital transformation.

3 The state of the art in legal text mining

One of the most challenging, but also most potential developments, comes with the web of data (Avila-Garzon, 2020) and the inherent mass of freely-available information, i.e., open data (Zeleti, Ojo, & Curry, 2016). Legal data and law texts is a category of open government data and thus they possess the potential to unlock innovation in governments and businesses, regarding the development of new, better, and more cost-effective services for citizens (Zuiderwijk & Janssen, 2014). In case of public administrations and governments, the distribution, availability and access towards legal information is crucial. Yet, there exist some severe issues at the moment regarding this access. One of them is found in form of available APIs, which are not always up and running on a 24/7 basis, paired with slow systems and often non-compliant data towards standard or even self-issued schemata. This in turn makes the use of automated crawling and analysis more than difficult (Charalabidis, et al., 2018).

Text mining, also known as text data mining, intelligent text analysis (Gupta & Lehal, 2009; Erhardt, Schneider, & Blaschke, 2006) or knowledge discovery from textual (structured) databases (Justicia De La Torre et al., 2018; Delen & Crossland, 2008), has been defined as “the discovery by computer of new, previously unknown, information by automatically extracting information from different written resources” (Hassani et al., 2020). Generally, text mining refers to the process of extracting interesting and non-trivial patterns or knowledge from unstructured text documents (Justicia De La Torre et al., 2018; Gupta & Lehal, 2009). Legal text mining analyses legal texts in order to extract useful legal information such as an overview of text’s content (Merkl & Schweighofer, 1997). Legal text documents are being unstructured stocked except in cases that online legal databases provide easy access to citizens, businesses etc. Furthermore, as pointed out by (Hassani et al., 2020) legal text documents are stored using natural language, so text mining can be suitably used for efficient analysis of such documents.

4 Metadata schema for law decomposition and semantic annotation based on Open Data

The Manylaws metadata schema is based on three standard ontologies, each one describing specific

elements of a document. The Data Catalog Vocabulary (DCAT)[1] is used to transform legal data in open legal data, the European Legislation Identifier (ELI)[2] to describe correlations between laws and EU Directives and Akoma Ntoso (AKN)[3] to describe parliamentary, legislative documents and the document lifecycle. Manylaws specific elements are added to cover the need of searching across borders for a specific term.

The scope of this complicated schema is to support the services of the Manylaws portal and to convert the legal documents into big, linked, open legal data. The basic elements that are connected with the Manylaws Services are included in Table 1. Specifically, the main scope of using DCAT-AP[4] is to convert the legal document to open data and the opportunity to being harvested by the European Data Portal (EDP) according to its guidelines, using all the mandatory fields that are required for a dataset in order to be harvested from the EDP[5] (e.g the title, the keywords, the theme of the legal document, the location, the description etc). ELI is used in order to imprint all the references that are included to a legal document because more and more European countries are using this ontology to describe legal documents. AKN is used in order to cover all the parliamentary supporting documents and steps that are required for the publication and adoption of a law. Furthermore, AKN is using a more flexible way to present the body of a legal document (e.g. lifecycle, workflow and body). Finally, the Manylaws Platform Extra (MLEXTRA) elements are being used to cover all the data that are not covered from the previous ontologies/ vocabularies. Particularly, the elements contain information for the similarity between two laws, between the articles of two laws, the translated body of the law, the nouns that are included in the body of the law and the ngrams that are a combination between adjectives and nouns, verbs and nouns, etc.

These three basic ontological approaches have been aligned and matched to each other in order to avoid duplication of metadata and data fields. This way, the final metadata schema includes the mapped fields and increases the flexibility of the whole approach. As the final step of the process, all data is stored into a virtuoso endpoint, so they are offered as linked data enabling also direct acquisition from the European Data Portal.

5 The process using text mining and HPC

Figure 1 describes the automated process of decomposition and semantic annotation. All the components and they are not triggered by user actions. They are related to processes that either retrieve information when available or produce new information from the retrieved data which will be stored in XML/RDF files when ready. These components operate periodically, at fixed intervals and their content is often renewed. These micro-services could be connected to any source available retrieving the available data, in most of the cases in .pdf format. For example, in the Greek case, a web crawler is used for retrieving the Greek legislation from the Greek National Printing Office (Greek NPO). The retrieved data (PDFs) are being stored in a pre-processing repository.

Whenever a new file is being stored, the RapidMiner is triggered and initiates Hadoop. The Hadoop Cluster is used for parallelized processing of massive data in a reasonable time. Specifically, for the Greek data Hadoop is used for extracting all the necessary metadata from the legal documents and for annotating these metadata with semantics in rdf/xml (based on ManyLaws metadata schema).

As soon as the semantic annotation of the metadata is finished, Hadoop transmits all the data to be translated to the eTranslation DSI which updates the Virtuoso triple store after the translation is finished with the updated rdf/xml.

[1] https://dcat.org/

[2] https://eur-lex.europa.eu/eli-register/about.html

[3] http://www.akomantoso.org/

[4] https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/dcat-application-profile-data-portals-europe

[5] https://data.europa.eu/en

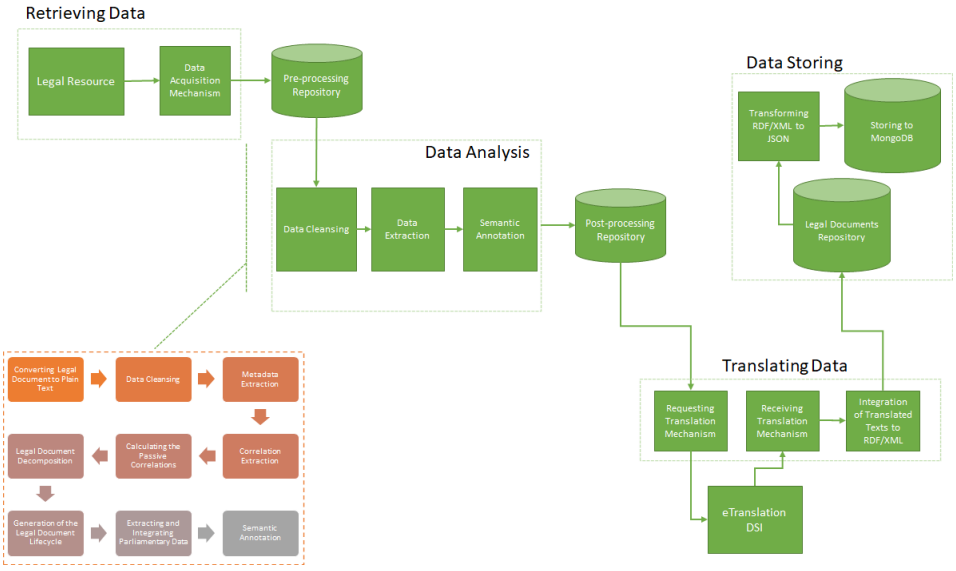


Figure 1: The decomposition and semantic annotation automated process

Given the vast amount of legal data published in the participating EU Member States, as well as their respective languages, analyzing it requires powers beyond that of a simple desktop computer, and thus High-Performance Computing is necessary.

6 Utilization of high-performance-computing infrastructure

This section presents the tools and algorithms of the designed process and the results of the high-performance computing (HPC) infrastructure usage in order to process and annotate big linked open legal data. The necessary tools and algorithms capable of handling large volumes and high velocity of data in the HPC environment have been developed and tested in the Greek and European case. The Greek National HPC Infrastructure "ARIS"[6] has been utilized in order to analyse a big number of legal documents.

The services and applications ran for 188 hours in the HPC system. The resources consumed are as follows: 8 nodes each one with 40 processing cores (totalling 320 processing cores); the total utilization of computing power amounted to 60,246 core hours; during the HPC operation 11,363,596 datasets were generated including the testing periods. It worth mentioning that the results achieved through the HPC operation in about 20 days, would require a high-end personal computer to work for almost 3,5 years.

Table 1 describes the micro-services that are necessary for the decomposition and semantic annotation. These micro-services are designed to work in parallel, so the utilization of an HPC infrastructure is possible. Thus, the development of the Message Passing Interface (MPI) was necessary.

MPI is an open library standard for distributed memory parallelization. The library API (Application Programme Interface) specification is available for C and Fortran. There exist unofficial language bindings for many other programming languages, e.g. Python a, b or Java1, 2, 3. The first standard document was released in 1994. MPI has become the de-facto standard to program HPC cluster systems and is often the only way available. There exist many implementations, Open Source and proprietary. The latest version of the standard is MPI 3.1 (released in 2015). The MPI interface is meant to provide essential virtual topology, synchronization and communication functionality between a set of processes (that have been mapped to nodes/servers/computer instances) in a language-independent way, with language-specific syntax (bindings), plus a few language-specific features. MPI programs always work with processes, but programmers commonly refer to the processes as processors. Typically, for maximum performance to be achieved, each CPU (or core in a multi-core machine) will be assigned a single process. This assignment happens at runtime through the agent that starts the MPI program, normally called mpirun or mpiexec.

The developed MPI program used to support our process divides the main process into forty (40) micro-services. For this purpose, we developed a script to divide the bulk data into forty parts. Each micro-service runs a specific part of the process on the bulk legal data that requires processing (data mining and semantic annotation) to be aligned with the developed metadata schema.

[6] https://hpc.grnet.gr/en/

Table 1. Micro-services Produced in the ARIS HPC

Service Name	Description
Title Extraction	This micro-service is used to extract the title of the legal document.
Number Extraction	This micro-service is used to extract the number and the year of the legal document.
Text Decomposition	This micro-service is used to decompose the body of the legal document in articles, paragraphs etc.
Extracting Ngrams	This micro-service is used to remove the common words in the legal document and produce the Ngrams.
Extracting Keywords / calculation	This micro-service is used to compare the extracted keywords with the EUROVOC 3rd and 4th level and finally produce the keywords of the legal document.
Correlations Extraction / calculation	This micro-service is used to extract the correlations of the legal document.
Theme extraction / calculation	This micro-service is used to compare the ngrams with the 1st and 2nd level of EUROVOC and find the most common themes.
Publication Date Extraction	This component is used to extract the publication date of the legal document.
Passive Correlations	This component is used to read all produced xmls for Austrian, Greek and European legal documents and update the necessary xmls with the passive correlations.

Even if the research is in pilot phase, the extraction text mining part and passive correlations of law in several pilot countries is working.

7 Conclusion

Automatic end-to-end law analysis conducted by decomposition and semantic annotation, by using the high-performance-computing and government open data can be presented through the processing model and micro-services already developed. and have potential to become a driver of the digital transformation. Even if results could be seen as pilot research on law data from Greece law texts, as open data, they possess the potential to unlock digital innovation and transformation capacity.

This paper presents research baseline for automation of decomposition and annotation with process and service elements developed for utilization on high-performance-computing infrastructure based on government laws open data.

Possible added-value services could be described as follows: i) research through legal corpora, analyzing the alignment of national legislation with EU legislation, ii) comparing national laws iii) analyzing the references to European legislation by national laws, iv) analyzing related laws within the same Member State, v) timeline analysis for all legal acts, vi) visualization of the progress and current status of a specific national or European piece of legislation and vii) sentiment analysis towards new legislation.

Some limitations exist regarding the added-value services that could be offered in the future. There might be a new conceptualised added-value service in the future that could not be supported by the developed and aligned data model. To solve this, we might need to create new data fields or incorporate more fields from the current ontologies. Finally, this solution needs to be tested to more countries to prove its efficiency and effectiveness to different legal environments.

Acknowledgments

Research in this paper is part of larger projects activities in one way connected to use and reuse of Open Data. This research paper is partially supported by Twinning Open Data Operational project (grant agreement No. 857592) and the ManyLaws project (grant agreement. No. INEA/CEF/ICT/A2017/1567 047) and do not necessarily reflect the opinion of the European Union. Both projects received funding from the European Union.

References

Avila-Garzon, C. (2020). Applications, methodologies, and technologies for linked open

data: a systematic literature review. International Journal on Semantic Web and Information Systems (IJSWIS), 16(3), 53-69.

Casanovas P., Palmirani M., Peroni S., van Engers T., & Vitali F. (2016). Semantic web for the legal domain: The next step.Semantic Web, 7(3), 213–227. 10.3233/SW-160224

Charalabidis, Y., Zuiderwijk, A., Alexopoulos, C., Janssen, M., Lampoltshammer, T., & Ferro, E. (2018). Open Data Interoperability. In: The World of Open Data. Public Administration and Information Technology, vol 28. Springer, Cham. https://doi.org/10.1007/978-3-319-90850-2_5

Delen, D., & Crossland, M. D. (2008). Seeding the survey and analysis of research literature with text mining. 34(3), pp.1707-1720.

Erhardt, R. A., Schneider, R., & Blaschke, C. (2006). Status of text-mining techniques applied to biomedical text. Drug discovery today, 11(7-8), pp. 315-325

Fulton C. (2011). Web accessibility, libraries, and the law.Information Technology and Libraries, 30(1), 34–43. 10.6017/ital.v30i1.3043

Gupta, V., & Lehal, G. S. (2009). A survery of text mining techniques and applications. Journal of emerging technologies in web intelligence, 1(1), pp. 60-76.

Hassani, H., Beneki, C., Unger, S., Mazinani, M. T., & Yeganegi, M. R. (2020). Text mining in big data analytics. Big Data and Cognitive Computing, 4(1), 1.

Hrustek, L., Furjan, M. T., & Pihir, I. (2020). ENABLING OPEN DATA PARADIGM FOR BUSINESS IMPROVEMENT. Economic and Social Development: Book of Proceedings, 174-183.

Hrustek, L., Tomičić Furjan, M. & Pihir, I. (2019). Influence of Digital Transformation Drivers on Business Model Creation. In Proceedings of the 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics; Croatian Society for Information and Communication Technology, Electronics and Microelectronics, Opatija, Croatia, 20–24 May 2019; pp. 1509–1513.

Justicia De La Torre, C., Sánchez, D., Blanco, I., & Martín-Bautista, M. J. (2018). Text mining: techniques, applications, and challenges. International journal of uncertainty, fuzziness and knowledge-based systems, 26(04), 553-582.

Lachana, Z., Loutsaris, M. A., Alexopoulos, C., & Charalabidis, Y. (2020). Automated Analysis and Interrelation of Legal Elements Based on Text Mining. International Journal of E-Services and Mobile Applications (IJESMA), 12(2), 79-96. http://doi.org/10.4018/IJESMA.2020040105

Merkl, D., & Schweighofer, E. (1997). En route to data mining in legal text corpora: Clustering, neural computation, and interational treaties. In Database and Expert Systems Applications, 1997. Proceedings., Eighth International Workshop on, IEEE, pp. 465-470.

Nguyen L. M., Bach N. X., & Shimazu A. (2011). Supervised and semi-supervised sequence learning for recognition of requisite part and effectuation part in law sentences.Proceedings of the 9th International Workshop on Finite State Methods and Natural Language Processing, Association for Computational Linguistics, pp. 21-29.

Tomičić Furjan, M., Tomičić-Pupek, K. & Pihir, I. (2020). "Understanding Digital Transformation Initiatives: Case Studies Analysis", *Business Systems Research*11(1), 125-141.

Zeleti, F. A., Ojo, A., & Curry, E. (2016). Exploring the economic value of open government data. Government Information Quarterly, 33(3), 535–551.

Zuiderwijk, A., & Janssen, M. (2014a). Open data policies, their implementation and impact: A framework for comparison. Government Information Quarterly, 31(1), 17–29 https://doi.org/10.1016/j.giq.2013.04.003

A comparison of business platforms used by SMEs to digitalize management activities

Olivia Doina Negoita, Mirona Ana-Maria Popescu
Faculty of Entrepreneurship, Business Engineering and Management
University POLITEHNICA of Bucharest
313 Spl. Independentei, 006042, Bucharest, Romania
{olivia.negoita, mirona.popescu}@upb.ro

Albena Antonova
Faculty of Mathematics and Informatics
Sofia University "St. Kliment Ohridski"
25 James Bourchier Blvd, Sofia 1164, Bulgaria
a_antonova@fmi.uni-sofia.bg

Dominik Bork
Faculty of Informatics, Business Informatics Group
TU Wien
Favoritenstrasse 9-11, 1040 Vienna
dominik.bork@tuwien.ac.at

Abstract. *The current economy is based on innovation, and this is the engine that fuels the ambition to gain market share and make a difference in the product. It also helps minimize operational costs and improve operational efficiency, thus maximizing profitability. This requires different skills than we have now. Especially in the context in which technology can replace many manual operations, in many economic fields. The authors aim to present a static in terms of digitalization in Romanian business environment. Through secondary research of current scientific literature there are presented software applications developed for companies to improve their efficiency. Using an online tool Similar web there are extracted up to date data to describe the client perspective towards different tools to help organizations reach digitalization.*

Keywords. Digitalization, software applications, business environment

1 Introduction

The need and access to new technologies and digital tools have grown with the context of the pandemic and the health crisis that has moved both most employees, but also pupils or students, together with their teachers, to the online environment.

Following this challenge, one of the pressing questions was whether all those who need it have access to digital tools, the Internet, and other technologies, and use them. But, with the passage of time, people realized that the right question is whether those who had access to all the necessary tools also had sufficient knowledge and skills to use them.

Access to the opportunities offered by the digital world should be equal for all. The most recent studies show that Romanians are in the last places in terms of daily internet use, compared to other regions in the European Union.

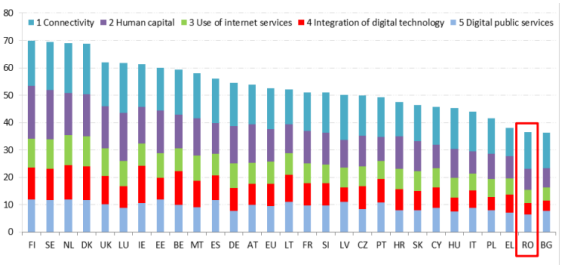


Figure 1. Digital economy and society index (DESI) 2019 ranking. Source: Digital economy and society index, accessed at 8 May 2022

According to the 2019 Digital Economy and Society Index (DESI) by the European Commission, Romania is ranked 27th out of the 28 EU member states.

Romania's performance in virtually all the DESI categories assessed indicates small improvements, but considering that the overall gain was moderate, its ranking remained steady. Romania scored highest in the connectivity category since there are so many fast and ultrafast fixed broadband networks available there (especially in urban areas). The economy is still not fully digitalized, more than a fifth of Romanians have

never used the internet, and less than a third possess even the most fundamental digital skills.

Romania performs the worst among the member states when it comes to digital public services, despite the high percentage of e-government users (7th in the EU). On the other hand, Romania has the third-highest percentage of ultrafast broadband subscribers in the EU at 45% of homes. Romania is well-positioned in terms of female ICT professionals, ranking 16th, with 1.3 percent of Romanian women employed.

Any manager is constantly looking for new ways to become even more productive and increase the business and team results. Without an electronic task management system, it is difficult to check and track tasks assigned to colleagues and desired deadlines. By using a task management program leaders have a complete picture of tasks and processes; they can track deadlines and check the status of various tasks.

The aim of this paper is to present the most used web applications in the business environment and to make an analysis of them for further use of organizations. As methodology, the authors analyze latest management software adopted in organizations through a review of scientific papers. Similarweb analytical tool is used to determine current trends for similar sites and their rank in terms of user's usage and appreciation. Following the extracted data, the authors compare the most popular software for management activities. Through this paper the authors pursued three main objectives:

O1. Determine the current trends in terms of business applications

O2. Establish the main types of software applications used by organizations to digitalize their current activity

O3. Determine the user perspective towards software applications used by SMES

2 State of the art

Software solutions are constantly adapted by companies to automate their processes and allocate their time to other activities that lead to high performance. By adopting digitization, efficiency increases, and they remain competitive at the market level.

The management can combine all corporate operations and processes thanks to the effective deployment of the ERP system, giving the firm a competitive edge in the current competitive market. Companies opt to devote a substantial amount of time and money to implementing ERP systems after considering the possible advantages. However, a lot of ERP system installation projects have still fallen short due to other crucial success elements that have been mentioned in the literature. It is crucial to analyze the state of ERP system use in the Republic of Croatia since the use of ERP systems is one of the major indicators emphasized by the European Commission under the Digital Agenda for Europe (EU 2020), which establishes the European Union's (EU) growth objectives by 2020. The purpose of this paper is to identify which companies in the Republic of Croatia consider the most significant critical success factors, considering the peculiarities of business in relation to developed countries and potential differences in attitudes among large companies and medium- and small-sized businesses. This is because critical success factors of the successful ERP implementation are investigated in the Republic of Croatia to a lesser extent. An online questionnaire survey was used for this aim, and the Mann-Whitney U test was used to compare major corporations to medium-sized and small businesses (Hornung, 2020).

Prior studies have emphasized the digitalization of innovation processes and results, particularly in relation to the effects of information technology (IT) on the creation of new products (NPD). This paper's goal is to investigate the influencing variables on the adoption of certain software tools to support innovation management methodologies, known as innovation management software (IMS), and their unique capabilities, as opposed to evaluating general IT usage in the NPD. To investigate which tool functionalities, such as assisted innovation techniques and activities, impact the adoption of IMS, the authors employed an online questionnaire to collect information from 99 innovation managers of German industrial enterprises.

Contrary to earlier hypotheses made in the literature, the findings of this study suggest that IMS adoption by innovation managers is favorably affected if the IMS tools include capabilities for concept appraisal and portfolio management but declined for idea creation and scenario management. Because of this, this article may demonstrate that the digitalization of the innovation process using IT tools is more finely nuanced than a "the more - the better" logic that is frequently supported in the context of digitalization and suggested in the literature on IMS earlier.

These findings contribute to innovation managers' understanding of organizational and technical factors that are transforming the innovation process toward digitalization. Innovation managers want to do this by successfully integrating and utilizing IMS within their NPD. Additionally, IMS providers receive insightful empirical data from the user perspective that may aid in their marketing and sales efforts. From a theoretical standpoint, this study advances previous research on IT use and digitalization in the innovation process by focusing on a segment of digitization for the NPD activities that has received less attention: the perspective of the innovation manager (Huesig, 2018).

The purpose of this research is to examine the current state of knowledge (1998–March 2019) about the effects of implementing enterprise resource planning (ERP) in small and medium-sized businesses (SMEs). Both important success and failure variables are present. The method relies on creating surveys to acquire optional information. It describes the terms that are used to get research articles from databases and

sophisticated studies on the adoption of ERP in SMEs. It also includes consideration and restriction criteria to strengthen the validity of papers. At that stage, a thorough review of the available publications is done to investigate the impact of ERP adoption in SMEs. The real impact of ERP installations on SMEs and the critical criteria that determine their success or failure are revealed (add actual success and failure factors here besides impact). Because only recent research publications were examined, the study presented in this paper has several limitations. Its failure to consider empirical research in favor of only concentrating on the state of the art in the research field is an inherent restriction. However, its revelations will likely be useful and provide guidance for further work. This study adds to the body of knowledge about the beneficial or negative effects that ERP deployment may have on SMEs. Using secondary data collecting techniques, it identifies important success, failure, and impact aspects. The insights will enable SMEs, SMEs' stakeholders, and ERP service providers to better understand the causes of success or failure and to choose the appropriate course of action (Kiran, 2019).

This study evaluates how various digital technology adoptions impact corporate productivity. In an empirical approach that takes business heterogeneity into account, it combines cross-country firm-level data on productivity with industry-level data on the use of digital technologies. The findings offer solid proof that a sector's use of digital technology is linked to increases in firm-level productivity. Effects are more pronounced in routine-intensive industries like manufacturing. As a result of the complementarities between digital technologies and other forms of capital, they also tend to be stronger for more productive enterprises and weaker in the case of talent shortages (e.g., skills, organization, or intangibles). Digital technology may thus be partly to blame for the increasing disparity in productivity performance between enterprises. Therefore, efforts to encourage digital adoption should go hand in hand with building the framework necessary for trailing enterprises to catch up, particularly by facilitating access to skills (Gal, 2019).

With the goal of discovering BPM adoption success criteria, several research on the successes and failures of business process management (BPM) have been carried out. Because of BPM's complexity and breadth, there is currently no widely acknowledged paradigm for its successful deployment. Using business process management software is one way to ensure BPM success in general (BPMS). It is challenging to specify the requirements for choosing BPM software since there is presently no agreement on a universally acknowledged description of the program. This is due to several factors: Because (i) the scope and complexity of the industry, and (ii) the sometimes-complex nature of company demands, additionally, (iii) the features and capabilities of BPM software vary widely amongst suppliers due to the complexity of the industry. In this post, we look at the technical and contextual aspects of BPMS adoption and the associated key success criteria (CSF). This study's objective was to provide decision-makers with BPMS selection criteria that considered the organizational, environmental, and technological CSFs of BPMS adoption. With order to do this, we used a multiple-case study methodology and conducted a series of interviews in businesses that had either completely or partially adopted BPMS. Semi-structured interviews were utilized to collect qualitative contextual (organizational and environmental) CSFs important for the success of BPMS adoption as well as quantitative data for those issues that can be measured statistically (Bosilj, 2018).

The performance and accomplishment of corporate goals have improved for organizations that have adopted the BPM management method. To reap the rewards of BPM, organizations must properly implement the BPM management methodology across the whole enterprise. There is very little empirical evidence to support the claim that having an organizational culture that is supportive of and in line with the BPM culture is a key component in the success of BPM adoption. In many strategic BPM efforts, an IT solution is put into place. The effectiveness of BPM adoption with firms based in the United States executing BPM efforts including IT solutions varies across different types of organizational cultures inside IT, which was the focus of this non-experimental quantitative correlation study. A BPM project including an IT solution that was developed and implemented during the previous two years was the focus of the target demographic, which was made up of IT personnel who work in the US at companies with at least 50 workers. To assess the impact of organizational culture on BPM adoption, the study involves gathering a sample of 157 IT professionals from the target demographic of at least 223,000 utilizing an online survey company's marketing panel. ANOVA, Tukey-Kramer post-hoc, and Pearson's r correlation parametric tests were employed in the study to examine the sample data. In contrast to market and hierarchical culture types, the results showed that adhocracy culture type had the highest level of BPM adoption success. The adoption of BPM as determined by BPO is significantly positively correlated with the adhocracy cultural type. According to BPO and PPI measurements, there is a sizable inverse relationship between the market culture type and BPM adoption. The association between PPI and BPO is quite positive. The findings are different from previous studies, which may be due to other factors not considered in the current study, such as a distinct national culture, a population that solely includes IT resources, or a mix of both. The knowledge gathered from the current study can aid practitioners in making well-informed choices about how to embrace BPM within their IT community and offer academics knowledge for next

studies on the connection between organizational culture and BPM (Letts, 2019).

This essay investigates how adoption of digital technology is impacted by globalization. The paper's goal is to clarify how globalization influences the uptake of new technologies. Using cutting-edge panel data modeling, the authors leverage country-level data from the globalization index (KOF), digital adoption index (DAI), global competitiveness index (GDI), and total factor productivity (TFP) on 183 nations. According to empirical evidence, globalization has a big impact on how quickly technologies are adopted worldwide. The study's findings demonstrate that globalization has a favorable impact on technology spillovers and transfers, especially when employing digital technology. Countries experiencing considerable technological transition acquire digital technologies at an ever-increasing rate. Our study's data is based on a global viewpoint, with a sizable sample of 183 nations that account for nearly 80% of the transfer process (Skare, 2021).

There are many possibilities for classifying e-business applications. When it comes to business applications, there are two main categories.

1. Intra-business applications, which include all e-business applications that exist in the internal domain of the company / organization. The applications take place within the boundaries of the organization and are connected to the internal activities of the business.
2. Interface applications, which include all applications that exceed the limits of an organization and ensure interaction with external entities such as: customers, a business partner company, or a financial institution.

Most enterprise solutions are based on an intranet infrastructure for sharing information and computing resources between employees. This class of applications is commonly known as B2E (Business-to-employee). In addition, an intranet can be used to facilitate group work or teleconferencing.

Content can be retrieved from internal databases and publications, external databases and publications, and new information from employees. E-business maintains new types of relationships between organizations: automating the exchange of information because of integrating business applications and processes and collaborating by sharing information and knowledge between businesses for mutual benefit (Figure 2).

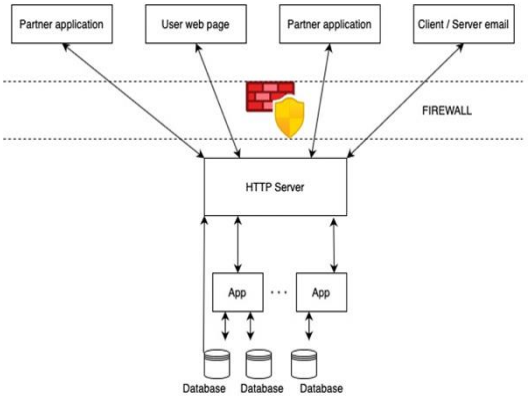


Figure 2. Data share. Adapted after source: Classification of e-business applications, accessed at 11 May 2022

Software applications are those programs that are designed for users to use the computer in a specific problem and to perform a specific task (word processing, billing, DTP, graphics applications). There are different programs each with a specific function, for example:

- Communication programs - Yahoo Messenger, Outlook Express (with these programs a person can send messages and communicate with different people regardless of their geographical location).
- File handling and management programs - Apple OS 9, Linux, Windows Explorer (with these programs a person can create, delete, or rename files).
- Web browsing programs - Netscape Navigator,
- Internet Explorer (with these programs a person can access various Internet pages).
- Word processing programs - WordPro, StarOffice Document, Microsoft Word (with these programs a person can access information in the form of text, with the ability to edit, save and print the document).
- Spreadsheet programs - Microsoft Excel, Lotus 1-2-3, StarOffice Spreadsheet (these programs allow the manipulation of existing numerical data in spreadsheets).
- Database management programs - Filemaker Pro, Microsoft Access, Appleworks (this program organizes large collections of data so that information is available to the user by performing queries and retrieving data).
- Others, used in various fields, specialized - Adobe Illustrator, Quark Express.

The main categories of software and their utility are summarized in the table 1.

Table 1. Software classification. Source: Software classification based on criteria, accessed at 8 May 2022

Software type	Utility
Operating systems	- Basic software packages - Essential for computer operation - The interface between the computer and the user, respectively between the computer and the applications used.
Office packages	- Application packages that help the user to perform the usual tasks related to office work - Includes: text editors, spreadsheets, presentations, databases, calendar.
Text editor	writing letters and other documents
Tabular calculation	working with tables in general, data tables in particular
Database	managing large collections of data, such as information for population records
Presentations	creating printable or projectable presentations to help give lessons, speeches, etc.
Agendas	- replace traditional agendas, - allow the recording and organization of contact details, meetings / meetings, tasks to be performed
Emails	- receiving and viewing emails, - composing and sending new messages, - organizing messages
Internet browser	Internet browsing
Graphics	create, scan, edit and correct images

3 Analysis of software applications used in SMEs

Now there are countless applications with multiple functions to help the process of digitalization (Parviainen, 2017). What a company needs to do is establish its needs, prioritize them, and based on these aspects, choose the best applications that cover its needs and help it grow further (Yoon, 2018). There are companies which prefer custom made applications and some which consider that a better fit are the ones already developed on the market (Koutsabasis, 2008).

According to Risco, the most used software applications are divided in 3 main categories, listed below. Each software platform was analyzed using Similarweb tool from a traffic and engagement perspective, audience perspective and competitors' perspective. Using the online software tool Similarweb there are presented statistics in terms of audience, age range and similar platforms. This is done to highlight that young people have better digital skills are eager to learn and apply software in their daily work activities.

Applications for overall productivity

Evernote is a personal assistant and is an application for organizing. The main feature of the application is the creation of notes and to-do lists, but Evernote has other extremely interesting features: scanning documents, saving favorite web articles and keeping them for later reading and sharing ideas. , notes with team members. Evernote has a basic package that is free and contains only the part of creating notes and other premium and business packages, at prices starting from 11 RON / month. Compared to last month the traffic of this application has decreed by 8.18% on mobile and desktop environment according to Figure 3.



Figure 3. Traffic and engagement, Total visits last 3 months. Source: Similarweb, accessed at 10 May 2022

Audience composition may state the current market share of people from different age intervals.

This application is used by male audience at a percentage of 61.03% and 38.97% by females as presented in Figure 4. The largest age group is formed by people who have the age between 25-34 years old (only on desktop).



Figure 4. Gender, Age distribution. Source: Similarweb, accessed at 10 May 2022

In Table 2 are presented the top 5 competitors for Evernote. The most visited is Toodledo with 920.9K, but it is not ranked as good as Trello from the user's perspective. If it is taken into consideration the afinity, the first two positions are taken by getpocket and todoist with a percentage of 100%. After Trello, if it Zoho seems to be the best alternative based on these three characteristics.

Table 2. Competitors and Similar sites. Adapted after source: Similarweb, accessed at 10 May 2022

Competitors	Affinity	Monthly visits	Category rank
getpocket.com	100%	13.0M	134
todoist.com	100%	20.4M	78
Toodledo.com	97%	920.9K	2,015
Trello.com	95%	93.6M	13
Zoho.com	92%	50.9M	22

Project Management Applications

Trello is excellent for organizing and tracking any project within the company. The tables, lists, and charts that Trello provides help you better visualize the status of a project and communicate more effectively with team members. Each task is assigned a "card" that can be moved using drag-and-drop depending on the stage of the project. You can add comments, deadlines, everything in a very simple to use and very useful application. The packages are basic (free), Business Class and Enterprise starting at $ 9.99 per user, per year. This month the traffic of Trello has decreased by 6.91% according to Figure 5. Also, it is revealed if the site meets the clients' expectations and if they find it useful.

The audience for Trello is formed mostly by males 67.60% and only 32.40% by females as in Figure 6. The percentages are like Evernote. Also, the main market is formed by people with age between 24-34 years old, meaning young people that work and have digital skills.



Figure 5. Traffic and engagement, Total visits last 3 months. Source: Similarweb, accessed at 10 May 2022

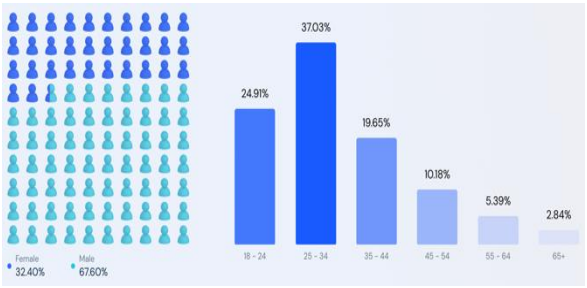


Figure 6. Gender, Age distribution. Source: Similarweb, accessed at 10 May 2022

In Table 3 are listed the main competitors for Trello and based on the date Github and Stackoverflow are the best ranked with an affinity of 100% and 97%. Also, on these two platforms are registered the best data in terms of monthly visits, due to their useful content and community.

Table 3. Competitors and Similar sites. Adapted after source: Similarweb, accessed at 10 May 2022

Competitors	Affinity	Monthly visits	Category rank
Github.com	100%	412.9M	4
Stackoverflow.com	97%	274.4M	8
Pivotaltracker.com	89%	404.3K	3,813
Miro.com	87%	25.1M	75
Atlassian.com	86%	25.6M	58

Asana is another web and mobile application used to create and manage projects on a shared table board. It can be integrated with Slack, Google drive and Dropbox to facilitate communication between members without sending more emails. Based on figure 7 the traffic has decreased by 10.88% from past month, but the users still spend a significant time in this platform, around 11 minutes.

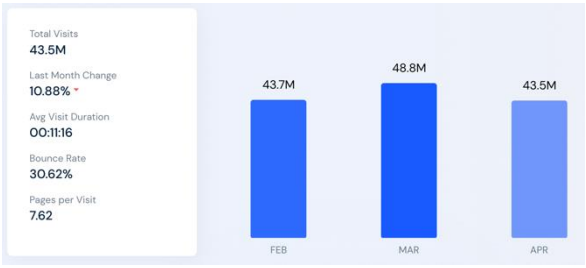


Figure 7. Traffic and engagement, Total visits last 3 months. Source: Similarweb, accessed at 10 May 2022

The female percentage (44.70%) is under the male (55.30%) in terms of audience, but the difference is not that high, only 10.6%. This application is used generally by young people having between 25 and 34 years old.

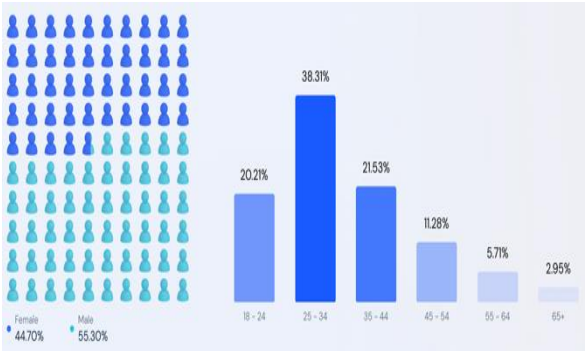


Figure 8. Gender, Age distribution. Source: Similarweb, accessed at 10 May 2022

In Table 4 there are listed the top competitors for Asana. The most monthly visits are on Notion competitor, that is better ranked than Trello.

Table 4. Competitors and Similar sites. Adapted after source: Similarweb, accessed at 10 May 2022

Competitors	Affinity	Monthly visits	Category rank
Trello.com	100%	93.6M	13
Notion.com	99%	113.8M	8
Figma.com	95%	49.8M	10
Calendly.com	93%	32.1M	56
Airtable.com	92%	18.4M	91

Customer Relationship Management (CRM) Applications

Nimble is a CRM application that helps to associate customer contacts with social media profiles. Through it, common interests can be found with them, so that everyone can create a customer profile and thus prospect exactly the customers they need. It can be easily integrated with other applications and services such as MailChimp, Evernote, Magento and many more. The packages provided by Nimble start at $ 19 per user per month.

According to Figure 9 the traffic has increased on Nimble with 137%.

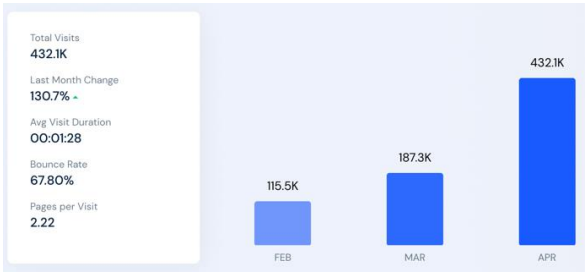


Figure 9. Traffic and engagement, Total visits last 3 months. Source: Similarweb, accessed at 10 May 2022

The audience is in the same range as for Trello and Evernote meaning 58.62% males and 41.38% females as presented in Figure 10. Also, 32% of the audience is composed by young people with age from 25 up to 34 years old.



Figure 10. Gender, Age distribution. Source: Similarweb, accessed at 10 May 2022

In Table 5 there are listed top 5 competitors of Nimble software. The best ranked is Nimblestorage, followed by Agilecrm. Agilecrm has the most monthly visitors, followed by Sendible. As from an affinity perspective, the top two positions are taken by Sendible and Gravlab with 100%. As it can be seen the values are for affinity are close to each other, but the other two categories have values very different as Nimblestorage has quite a low audience but is better ranked.

Table 5. Competitors and Similar sites. Adapted after source: Similarweb, accessed at 10 May 2022

Competitors	Afinity	Monthly visits	Category rank
Sendible.com	100%	225.6K	29,32
Gravlab.com	100%	<5K	
Nimblestorage.com	88%	61.6K	16,27
Agilecrm.com	85%	268.7K	19,30
Commence.com	85%	10.3K	228

In Figure 11 is presented a comparison made by authors between these 3 top software applications used by companies all around the world. It was taken into consideration the visit average from last 3 months, the duration of the visits and the pages accessed per visit, illustrating the interest of end users. Nimble is by far the most visited, but the users only visit 2 pages and the average time spent is 1 minute. In comparison Trello has an average of 7 pages per visit and users spend around 4 minutes.

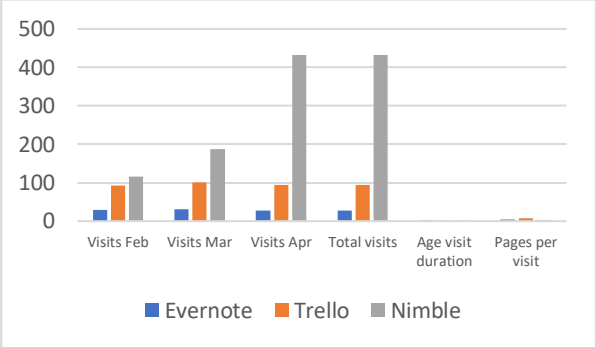


Figure 11. Comparison of software applications for management by visits

4 Conclusions

The business digitalization process is gaining momentum in our country. More and more funds are already being offered for business digitization. Digitization is becoming a necessity in business because of the speed with which the world is developing. Employees are evolving, jobs are changing, and business is becoming more and more automated. Mental work is gaining ground in the face of physical work, with employees' role becoming to think about processes, enter data, and check systems and robots. Digitization is a process, and it is different for each company. It can be different even for companies that have the same field of activity.

The younger generation is now also known as Generation C (connected), because people up to the age of 40 prefer to interact with businesses and services online, but also using software applications, (Cirilli, 2019). From websites, payment platforms and even online counters, all these become easier to access, which is why those who grew up with technology prefer these means. A business that wants to grow quickly and well must work hard to digitize all services and invest in employees and a work environment equipped with tools designed to facilitate their work. Thus, companies must invest in employees, but also in the necessary software solutions to achieve operational excellence, (Sokiyna, 2020).

The study conducted on ERP Systems in Croatian Enterprises analyzed the state of adoption this type software solutions to better understand what are the factors that generate a successful deployment (Hornung, 2020).

Another paper investigated the factors that affect how software tools that enable innovation management approaches are adopted (Huesig, 2018).

The evolution of knowledge regarding enterprise resource planning software between 1998 and 2019 with the objective of establishing beneficial or negative effects of ERP deployment inside SMEs (Kiran, 2019).

Regarding digitalization and productivity assesses the effects of adopting various digital technologies on company productivity (Gal, 2019).

A case study was done on business process management software was entailed to extract both accomplishments and setbacks found in real world scenarios (Bosilj, 2018).

A survey enacted upon 157 IT professional was conducted to obtain relevant date pertaining to how company culture affects the adoption of BPM (Letts, 2019).

An essay was done with the intent to present the correlations between globalization and digital technology adoption (Skare, 2021).

All the scientific research made by authors through literature review and online analytical tool Similar web are supporting the objectives of this paper, enhancing the importance of digitalization and software applications in SMEs.

Funding

The DEMO project has received funding from the European Union's ERASMUS+ Program under the grant agreement number 2021-1-RO01-KA220-HED-000027576, project title: Digital Platform Enterprise (Acronym: DEMO).

References

Bosilj Vukšić, V., Brkić, L. and Tomičić-Pupek, K., 2018. Understanding the success factors in adopting business process management software: Case studies. Interdisciplinary Description of Complex Systems: INDECS, 16(2), pp.194-215.

Best applications for business, Risco, https://www.risco.ro/suport/practici-in-afaceri/cele-mai-bune-aplicatii-pentru-companii-de-care-ai-nevoie-pentru-a-ti-creste-afacerea-1855, accessed at 3 May 2022

Classification of e-business applications, http://informatica.ase.ro/site/A140212/clasificarea_aplicatiilor_de_eb.htm, accessed at 3 May 2022

Crittenden, W.F., Biel, I.K. and Lovely III, W.A., 2019. Embracing digitalization: Student learning and new technologies. Journal of Marketing Education, 41(1), pp.5-14.

Cirilli, E., Nicolini, P. and Mandolini, L., 2019. Digital skills from silent to alpha generation: An overview. In Edulearn19: 11th International Conference on Education and New Learning Technologies; Chova, LG, Martinez, AL, Torres, IC, Eds (pp. 5134-5143).

Digital economy and society index, https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-index-desi-2019, accesed at 8 May 2022

Eurostat, https://www.edu.ro/sites/default/files/SMART.Edu%20-%20document%20consultare.pdf, accessed at 7 May 2022

Gal, P., Nicoletti, G., Renault, T., Sorbe, S. and Timiliotis, C., 2019. Digitalisation and productivity: In search of the holy grail–Firm-level empirical evidence from EU countries.

Hornung, K. and Hornung, M., 2020. ERP systems in croatian enterprises. Tehnički vjesnik, 27(4), pp.1277-1283.

How are the Romanian employees in terms of digitization in the pandemic, https://www.startupcafe.ro/afaceri/digitalizare-romania-firme-angajati.htm, accessed at 15 May 2022

Huesig, S. and Endres, H., 2018. Exploring the digital innovation process: The role of functionality for the adoption of innovation management software by innovation managers. European Journal of Innovation Management.

Kiran, T. and Reddy, A., 2019. Critical success factors of ERP implementation in SMEs. Journal of Project Management, 4(4), pp.267-280.

Koutsabasis, P., Stavrakis, M., Viorres, N., Darzentas, J.S., Spyrou, T. and Darzentas, J., 2008. A descriptive reference framework for the personalisation of e-business applications. Electronic Commerce Research, 8(3), pp.173-192.

Letts, B., 2019. Relationship between IT Organizational Culture Type and Business Process Management Adoption in Organizations Implementing IT Solutions–A Correlational Study (Doctoral dissertation, Capella University).

Parviainen, P., Tihinen, M., Kääriäinen, J. and Teppola, S., 2017. Tackling the digitalization challenge: how to benefit from digitalization in practice. International journal of information systems and project management, 5(1), pp.63-77.

Similarweb, https://www.similarweb.com, accessed at 9 May 2022

Skare, M. and Soriano, D.R., 2021. How globalization is changing digital technology adoption: An international perspective. Journal of Innovation & Knowledge, 6(4), pp.222-233.

Software classification based on criteria, https://ldportofoliu.wordpress.com/2016/09/22/clasificare-softuri-criterii-de-clasificare/, accessed at 8 May 2022

Sokiyna, M. and Aqel, M., 2020. The role of e-business applications software in driving operational excellence: Impact of departments collaboration using sustainable software. Sustainable Computing: Informatics and Systems, 28, p.100445.

Vasilev, V.L., Gapsalamov, A.R., Akhmetshin, E.M., Bochkareva, T.N., Yumashev, A.V. and Anisimova, T.I., 2020. Digitalization peculiarities of organizations: A case study. Entrepreneurship and Sustainability Issues, 7(4), p.3173.

Yoon, C.Y., 2018. Measurement of firm e-business capability to manage and improve its e-business applications. In Advanced Multimedia and Ubiquitous Engineering (pp. 101-107). Springer, Singapore.

Digital skills gap: Industry demand vs. Higher Education Institutions offerings

Clara Viganego, Anna Akhmedova, Marta Mas-Machuca
Faculty of Economics and Social Sciences
Universitat Internacioanal de Catalunya
Immaculada, 22 08017 Barcelona, Spain
{cviganego,aakhmedova,mmas}@uic.es

Abstract*. The digital era has brought about the automation of the past routine tasks, as well as the development of soft skills such as problem solving or creativity. Every citizen is now required to handle basic digital knowledge to fully participate and live in modern society.*

In this context, a fundamental question arises: What do digital skills mean and how is it different from digital competence or digital literacy? Are universities training students properly so that they can achieve the level of digital skills that will be demanded in their future workplaces? Is there a gap between the companies' expectations and the graduates' abilities in terms of digitalization?

The responses point out that although mastery of basic automation tools is an essential requirement in any job description (e.g., advanced Excel), the use of more specific tools (transactional tools such as SAP, Salesforce, Power BI, etc.) is closely linked to the type of business and sector. For this reason, the focus should be placed not so much on the ability to execute digital tools but instead on the development of specific soft skills.

Keywords. Education, employment, digital skills, gap

1 Introduction

Every citizen is now required to handle basic digital knowledge to fully participate and live in modern society (Spante, 2018; Mattar et al., 2022; Zhao et al., 2021). The European Commission is fully aware of this call and because of that is actively trying to find ways to motivate the digitalization among its citizens to increase European competitiveness. The action of the European Commission can be classified into three main categories including investigation, funding, and education (Lopez-Meneses et al. 2020).

Many articles have been published regarding this area of knowledge (Crawford-Visbal, 2020; Guzman-Simon et al., 2017; Handley, 2018; Kim et al., 2018); however, this topic has not been approached from an enterprise point of view (Zhao et al., 2021). Instead, most of the articles focus on the university side and how can teachers improve when it comes to their digital preparation (Mengual-Andrés, 2016; Jiménez-Cortés et al., 2017). That is why the object of this research are digital skills specifically applied to the business level and recent BBA graduates, hoping that it can contribute to a better understanding of these two realities (He and Zhu, 2017; Leon-Perez, 2020).

The competency-based pedagogical model is having a significant influence on the debate about aligning Higher Education to labor market requirements. For that reason, each discipline needs to continuously monitor job prospects and adjust the university's competence curricula accordingly (Wild and Schulze, 2020; López-Gil, 2020).

Following this idea, the output of this research consists of a digital competency framework that can serve as a reference both for students when entering the labor market, and for businesses when determining their needs in terms of digital preparation.

There are three main research questions (R.Q.) that this investigation aims to answer:

R.Q 1: What do digital skills consist of according to the companies' point of view. What are the elements that conform to this definition?

R.Q 2: Is there a gap separating the demands of the industry and the supply of the University in terms of digital skills?

R.Q 3: If so, how can this gap be reduced so that the expectations of the sector are in line with the offerings of recent graduates?

2 Literature review

The search was developed through the Web of Science, including terms related to digital skills or competencies.

2.1 Criteria and selection

The criteria used in the search presented the following structure:

Table 1. Search strategy, title restrictions

1ST string "OR"		2ND string "OR"
Digital skill* Digital competence* Digital capabilit* ICT skill*	"AND"	Higher education HEI (higher education institutions) Universit*

1. First stage; or automatic selection: The records identified in the initial search included 161 results. From that, an automatic filtration was done excluding review articles, book chapters, book reviews, books, and editorial materials. Conversely, the number of articles fell to 148.
2. Second stage; manual selection: A final manual selection was done based on the abstracts, excluding the results that were irrelevant to the topic. After an exhaustive reading of the 148 abstracts, 77 articles were dismissed for being out of the scope. This meant that the final review contained 71 documents, including both proceedings papers and articles.
3. Criteria of inclusion and exclusion of the articles: Articles that pictured the level of digital skills among graduates or pointed out the impact of better digital capabilities on employability were included. Instead, those that focused too much on university teachers, studying methods or schools were dismissed, since the study focused exclusively on higher education institutions.

2.2 Bibliometric analysis

The analysis of the final articles was developed manually and with the help of VOS viewer, a software tool for building and visualizing bibliometric networks. Thanks to it, information can be clustered and classified based on common characteristics and patterns. That is why most of the figures in this research come from this source.

The elements conforming to the bibliometric analysis were time frame, geography, authors, journals, and content analysis.

The samples of the different studies included mainly students in different universities, who were assessed on their self-perceived digital competencies and other matters surrounding the topic.

Some articles combined several research methods, meaning that they relied on both qualitative and quantitative methodologies for their studies.

2.3 Content analysis

The five main descriptive tags conforming the articles were clustered and organizes. These were: "digital skills", "digital tools", "teaching", "students" and "digital natives and Covid-19".

After that, a distinction between three fundamental terms was presented. These terms are digital competence, digital skills, and digital literacy. The first two concepts can be considered synonyms since they refer to the "abilities necessary for proper inclusion and participation in our digital society" (UNESCO). Digital literacy, instead, represents the ability to understand the context that surrounds digitalization (Bali, 2016).

The E.U has created a tool to identify the basic elements that conform to digital competence. This tool is known as the Digital Competence framework (or Dig Comp 2.0), and it includes a set of 21 competencies in 5 different areas. It was used as a reference to develop the digital competency framework of this research.

3 Methodology

The sample of this research consists of 14 interviews: Eleven from business realities and three from the academic sector.

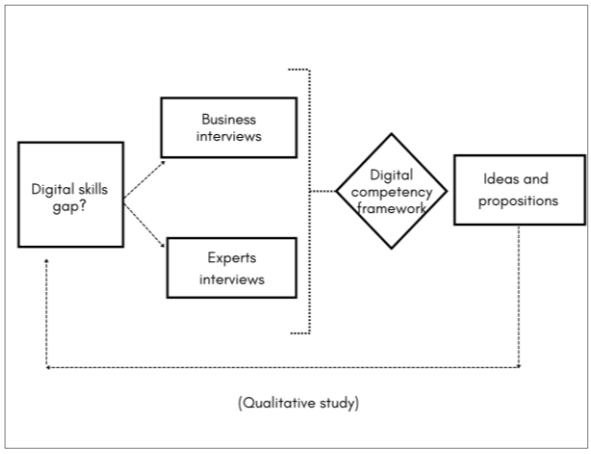


Figure 1. Structure of the research

3.1 Data collection and sample

To answer the three R.Q. we conducted a qualitative study based on semi structured survey. For data collection, people from both the business world and the academic sector (that is, university professors or researchers) were consulted. Each of them was individually contacted to schedule an interview according to their availability. Once the date was set, the interview proceeded, either in person or online (through zoom, teams or meet platform) following a pre-established thread of questions. All interviews were recorded, with prior consent, to allow a more accurate transcript of the responses.

The duration of the interviews was an average of 21 minutes each. The total duration of the recordings adds up to 274' and 45", that is, a total of 4:57 hours. The development of the interviews was online 85% of the time. Only once it was developed in person or through a written response that was provided via email.

On the one hand, experts from the university world were approached, specifically from three educational institutions (UIC Barcelona, IESE and Faculty of organization and informatics in Croatia) since their presence at the University gave them a clearer vision of the digital skills held currently by students. In addition, profiles such as Bea Lucaya (internship coordinator at UIC Barcelona) or Carlos Cosials (Industry 4.0 consultant and coordinator of the Master in Big Data, UIC Barcelona) represent a point of connection between university and business, given that they connect these two realities daily, both through teaching and service providing. In short, the profiles called "experts" were those that provided information about the HEI's offering, or what is the same, about the level of digital skills that a recent BBA graduate has today.

On the other hand, people working in companies were interviewed. The business realities were diverse, but they shared a common point: all of them hired recent business administration graduates, more or less frequently. That is why it was especially relevant to know their perspective. Ten was the number of analyzed companies, but eleven were the interviews since in the case of KPMG it was possible to contact two partners (one in the transaction services and deal advisory and the other in technology, risk, and cybersecurity). Of the ten companies analyzed, two are devoted to production (Quad pack in the cosmetics sector and Fluidra in the wellness sector) while the other 9 provide services. These go from consulting (KPMG, Deloitte, Industry 4.0 and Selliger and Conde), to courier services (Glovo), urban mobility (Yego) and financial services (Caixabank, Abacum).

3.2 Semi-structured interviews

For the data collection, an interview was developed both for businesses and experts in the field. The interview was carried out thoughtfully, to identify the existence or non-existence of a digital skills gap, and the enumeration of the skills necessary for any recent graduate coming from BBA studies.

To write these questions, it was necessary to attend to previous studies on the subject, where the style of questions could be inferred. It was not easy, because in most cases the previous studies had been quantitative using questionnaires and tests. Besides, the recipients of the questions were not normally companies but students. However, articles such as (Llorens, 2013) distinguished which generic skills were the key to enabling newly graduated engineers, basic computer, and telecommunication engineers to enter the information and communication technology (ICT) sector. It was an important reference that helped understand how to approach businesses.

The following table contains the structure of questions that was followed in all the interviews.

Table 2. Semi-structured interviews

Presentation questions	How many years have you been in the company? What has been your career and what is your current position? Do you manage a team? If yes, how many people do you hire per year and what studies do they normally come from?
Questions about skills	What do you understand by digital skills? What are the most important digital skills that a recent graduate must have to fit in a job like this? Can you give an example of digital competence that is especially demanded? Have you seen an evolution in this demand over the years you have been in the company?
Gap	Do you understand that greater preparation of recent graduates in these skills would be necessary? Do you have any examples that can illustrate this need? Have you ever developed "training plans" for the adaptation of employees to emerging technologies?
Relationship with universities	What relationship does this company have with universities? Do you see a need for more fluid university-company communication?
Future prospects	What are your future prospects in this regard? Do you think that this demand for digital skills will increase?

3.3 Data analysis

The interpretation of the data is qualitative since the information is not given by numbers or statistical data but is expressed through the words used by each respondent to the questions formulated in the interview.

Qualitative research has been highly criticized due to its “alleged incapability” of presenting empirical evidence in its conclusions and findings (Gioia, 2012) However, this work has used qualitative study tools that have been recognized precisely because they provide credibility to qualitative analysis, given their rigorous methodology and techniques for organizing and interpreting information.

Mainly two tools were used: The Gioia method and the NVivo software. On the one hand, the Gioia method is an approach to qualitative research that consists of the organization of data into 1st and 2nd order categories to render easier the assembly of the information into a more organized form. It brings as a result an inductive model which shows the dynamic relationships among the concepts obtained and finally a narrative of the theory is developed including quotes that prove the relation between the model and the fieldwork. In conclusion, the overall objective of this methodology is to build new concepts and ideas based on fieldwork, structure them and give them a meaning that can guide future research (Gioia, 2012).

On the other hand, to classify the "raw answers" gathered from fieldwork, a software called NVivo was applied. It constitutes a data management package to support the researcher during the data analysis process which the researcher must always remain in control of (Zamawe, 2015). This tool automated the organization of the information using markers and graphics that provided even greater rigor to the analysis.

4 Results

The presentation of the results follows the order of the research questions:

- R.Q 1: What do digital skills consist of according to the companies' point of view. What are the elements that conform to this definition?

This concept has been defined within a specific scope. This research, as explained in the introduction, studies digital skills applied to the business reality. In other words, we have tried to create a conceptual framework in which the most required skills could be classified according to a logic that is consistent with the requirements of organizations (see Figure 2).

To develop this framework, the answers given by the interviewees have been used, as well as the existing model of competencies that are presented in the literature review of this same work (DigComp 2.0). Thanks to the interview with Carlos Cosials, the triple distinction between the tools was identified, and it was very useful to systematize the information and concepts that the respondents pointed out. He also highlighted the concept of "lifelong learning" as an indispensable summary of the framework.

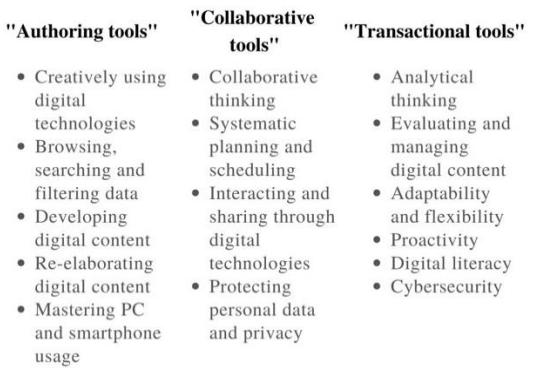


Figure 2. Digital skills, industry demands (BBA)
Source: Own elaboration.

First, authoring tools are those used for content creation (for example, the development of a PPT presentation with the results of a certain project). These tools are mainly represented by the well-known Office package (Word, Excel, PPT…) but there are also similar programs that carry out the same functions. As mentioned by Carlos Cosials: “*In the office, according to me, you work and produce because you manipulate digital tools to create content”.* The competencies associated to these tools were:

1. Creatively using digital technologies
2. Browsing, searching, and filtering data
3. Developing digital content
4. Re-elaborating digital content
5. Mastering PC and smartphone usage

Secondly, collaborative tools constitute the necessary instruments to discuss and share results with colleagues, always through digital means. Within this classification, the most frequently cited examples have been mail, the cloud, or videoconferencing platforms such as Google Meet, Teams, Zoom, etc. In other words, these are all programs that make it possible to create a "workflow" among employees where all the results are known and shared in an agile way. This category also includes programs designed ad hoc by consulting firms, which develop them after understanding the company's business model and way of working. A recent graduate can't know, let alone master, the management of this type of ad hoc designed programs, so technical knowledge of these tools is not expected but rather a package of soft competencies. The competencies associated to these tools were:

1. Collaborative thinking
2. Systematic planning and scheduling
3. Interacting and sharing through digital technologies
4. Protecting personal data and privacy

Finally, transactional tools are those where all the information related to business activity is centralized. An ERP (enterprise resource planning) refers to the

tool that allows the centralization of the information regarding all the company areas (such as HR, SCM, CRM, etc.). According to Judit Gispert, an ERP is the *"backbone of the company where all its information is dumped"*. Most of the time we talk about SAP, SAJE or Oracle platforms, but in any case, there are also ad hoc tools for each company that are highly adapted to their business model. The skills associated to these tools were:

1. Analytical thinking
2. Evaluating and managing digital content
3. Adaptability and flexibility
4. Proactivity
5. Digital literacy
6. Cybersecurity

All in all, the digital competence that will be required for a recent graduate can be identified always according to their ultimate utility. That is why this framework serves as a benchmark where the demands of the industries are properly organized and systematized.

- R.Q 2: Is there a gap separating the demands of the industry and the supply of the University in terms of digital skills?

To answer this question, it is necessary to look at the interviews. Most of the time, the ideas were repeated but formulated through different words. Therefore, it was necessary to regroup all the answers that referred to each question to find a common pattern. This was done through the NVivo software, which apart from classifying the citations, also offers the possibility to classify the categories of response ("nodes") according to the weight given to each of them. All in all, it was possible to obtain a hierarchical map showing each category with its respective weight, which is directly correlated to the frequency of references inside that group.

1. Analytical thinking competency was the most mentioned. This competency refers to the ability to *"interpret the information provided by our (digital) tools"*, Maite Poyos. In the opinion of most of the interviewees, this was a competency that although it has not been in demand to date, it is now that information is provided through large amounts of data. Some even believe that this will be the job of the future *("I think that at some point the market will require a change and analysts will be the most sought-after figure in a company"*, Matteo Saita). It is also referred to as "critical thinking".
2. Adaptability and flexibility also had a considerable weight. This skill is understood as the *"ability to adapt to all the changes in a more agile way than the rest"*, Sergi Gil. It has been very often cited since any recent graduate when he/she finishes university, is immersed in a completely new reality, where he/she will also have to work with new tools, especially digital. In this context, the willingness to learn or the flexibility to embrace change is important.
3. Another very frequently cited competency was mastering PC and smartphone usage. Most of them referred to the fact that the main work tool with which a worker spends the whole working day with is the computer. In this sense, the idea was repeated that a recent graduate cannot be unaware of how to use it. *"The first thing they will give you when you start in a position is a computer, so from that moment on you can't be a stranger to the use and handling of it",* Jorge Lluch.
4. The ability to develop digital content, on the other hand, has been constantly linked to Excel mastery. Any recent graduate is expected to know how to develop content through this tool not only at a basic but also at an advanced level. *"Right now, I wouldn't hire anyone who doesn't know how to use Excel"*, Beatriz Gisbert.
5. Among the least cited competencies there are collaborative thinking, cybersecurity, interacting and sharing through digital technologies, protecting personal data and privacy or systematic planning and scheduling. Most of them are competencies belonging to the category of collaborative tools.

- R.Q 3: If so, how can this gap be reduced so that the expectations of the sector are in line with the offerings of recent graduates?

Formation could be seen as the first solution. Most companies, especially the bigger ones, already offer courses for the digital education of their employees. In these courses, they can get in touch with the specific software that the employees manage on a day-to-day basis ("*We train based on what we need",* Sergi Gil). Some companies even have platforms where all the courses are offered for the employee to choose *from ("LinkedIn learning is that platform where everyone can organize their training as they want"*, Judit Gispert).

Some training is developed hand in hand with a senior employee that has experience in the job and can teach the recent graduate how the system works ("*When we hire, the person leaves the career, we put him/her together with seniors and we train him/her by participating in projects",* Sergi Gil). These training plans can be mandatory or not, always depending on the needs of the company.

Other ideas were shared such as the importance of mixing university and companies, the need for a more digitalized education at university level, the need for "hybrid profiles", that can combine business knowledge with the mastery of essential digital programs.

Lastly, the idea of "lifelong learning" was presented as an innovative approach to university education. It consists of the reduction of university timings that allow for faster insertion of students in the

working field, ensuring that their education is monitored over the years.

5 Discussion

After a thorough analysis of the interviews, the answer to the first research question is given in the form of a competency framework. It is not a closed classification since can be expanded and changed over time and depending of external situations as COVID (Batez, 2021; Heidari et al., 2021). This framework is intended to reflect the current reality of digital skills, according to the opinions of various experts in the field. It can serve as a basis both for companies (to detect the ideal profile for their needs) and for recent graduates (to have a reference of what can be expected of them in terms of digital skills). All in all, it represents a relevant and up-to-date contribution.

On the one hand, three types of tools have been distinguished: authoring tool (where the output is created from scratch), collaborative tools (through which employees work together) and transactional tools (where all the information regarding the business activity is centralized). Each of these tools has been associated with a series of competencies, fifteen in total, which have been analyzed according to the importance attributed by the interviewees.

It is worth noting the high weight given to competencies such as analytical thinking, adaptability, and flexibility to the detriment of others such as cybersecurity or systematic planning and scheduling (He and Zhu, 2017). The fact that the most frequently cited competencies have to do with transactional tools may lead one to think that these represent the main driver of value creation for the company today. In any case, it is understood that the interpretation and management of large databases represent an indispensable requirement for today's business reality.

On the other hand, it has been observed that companies are not so much interested in the executive aspect, but rather the soft skill preparation of students at the university level. That is, they do not expect recent graduates to know how to handle digital tools created ad hoc and specific to their business models, but rather they expect them to be able to adapt quickly to the digital reality of their company.

However, it should be noted that there are some tools of which mastery is expected, including video conferencing tools, the cloud, google workspace and the classic Office package (especially relevant at this point is the mastery of advanced Excel).

All in all, it could be said that there is a strong awareness on the part of companies of the importance of hiring people with adequate digital skills, according to Llorens (2013). But above all, they search for people that can adapt and dynamically embrace change. All this is because we are in a constantly changing reality, where tools evolve, and people are expected to evolve at the same pace. However, we must not forget that technology should not become the end of the organization, but the means to achieve business objectives.

6 Final remarks and implications

This study is highly practical for both academia and the business sector. It is in line with two of the thirty sustainable development goals proposed by the United Nations: "Quality education" as well as "Decent work and economic growth".

It differs from other studies in that it focuses on the opinion of companies, and not so much on the perspective of professors or professionals in the University. In this way, the results can represent a point of reference both for companies when hiring personnel and for recent graduates when entering the labor market. It should not be forgotten that one of the objectives of the university, especially in a faculty of business studies, should be to bring students as close as possible to the business reality.

Apart from the implications and practicality of this study, there are also several limitations, including the fact that it is a purely qualitative perspective. This study has analyzed the responses of 14 companies and experts to an interview. These results could be enriched by the quantitative study of data that could reinforce the ideas expressed above.

On the other hand, it should be noted that the companies were interviewed at the national level. That is, they are Spanish professionals except for the case of Ivan Malbasic, from the Faculty of organization and informatics in Croatia. This means that the conceptual span is narrower and that opinions from countries other than Spain cannot be reflected.

Finally, the number of interviews has been reduced given the limited availability of resources. A larger number of interviewees with a greater variety of companies could allow for a more rigorous analysis.

Nevertheless, this topic should be the object of further research to try to clarify how the reality of digital skills evolves in the future and how recent graduates should prepare themselves accordingly.

Acknowledgments

I would like to thank my thesis tutors Marta Mas (Dean) and Ana Akhmedova (teacher and researcher of my faculty), as well as all the interviewees who, despite their occupied agendas, offered some minutes of their time to respond to the interview questions (Beatriz Lucaya, Xavier Brossa, Sergi Gil, Beatriz Gisbert, Mattteo Saita, Sivlia Vizcaíno, Jorge Lluch, Maite Poyos, Carlos Cosials, Ivan Malbasic, Judit Gispert, Nuria Cobo, David Agost, Jose Luis Marcó).

References

Bali, M. (2016, 02 03). *Literacy now*. Retrieved from International literacy association: https://www.literacyworldwide.org/blog/literacy-now/2016/02/03/knowing-the-difference-between-digital-skills-and-digital-literacies-and-teaching-both

Batez. (2021). ICT Skills of University Students from the Faculty of Sport and Physical Education during the COVID-19 Pandemic. Sustainability, 10.

Crawford-Visbal, C.-T. O.-Z. (2020). Competencias Digitales en estudiantes de Comunicación a través de cuatro universidades latinoamericanas. Education in the Knowledge Society (EKS), 5-10.

Gioia, K. G. (2012). Seeking Qualitative Rigor in Inductive Research: Notes on the Gioia Methodology. *Organizational Research Methods*, 17.

Guzmán-Simón, F., García-Jiménez, E., & López-Cobo, I. (2017). Undergraduate students' perspectives on digital competence and academic literacy in a Spanish University. Computers in Human Behavior, 74, 196-204. https://doi.org/10.1016/j.chb.2017.04.040

Handley, F. J. L. (2018). Developing Digital Skills and Literacies in UK Higher Education: Recent developments and a case study of the Digital Literacies Framework at the University of Brighton, UK. PUBLICACIONES, 48(1), 109-126. https://doi.org/10.30827/publicaciones.v48i1.7327

He, T., & Zhu, C. (2017). Digital informal learning among Chinese university students: The effects of digital competence and personal factors. International Journal of Educational Technology in Higher Education, 14(1), 44. https://doi.org/10.1186/s41239-017-0082-x

Heidari, E., Mehrvarz, M., Marzooghi, R., & Stoyanov, S. (2021). The role of digital informal learning in the relationship between students' digital competence and academic engagement during the COVID-19 pandemic. Journal of Computer Assisted Learning, 37(4), 1154-1166.

Jiménez-Cortés, R., Vico-Bosch, A., & Rebollo-Catalán, A. (2017). Female university student's ICT learning strategies and their influence on digital competence. International Journal of Educational Technology in Higher Education, 14(1), 10. https://doi.org/10.1186/s41239-017-0040-7

Kim, H., Hong, A., & Song, H.-D. (2018). The Relationships of Family, Perceived Digital Competence and Attitude, and Learning Agility in Sustainable Student Engagement in Higher Education. Sustainability, 10(12), 4635. https://doi.org/10.3390/su10124635

León-Pérez, F., Bas, M.-C., & Escudero-Nahón, A. (2020). Self-perception about emerging digital skills in Higher Education students. Comunicar, 28(62), 91-101. https://doi.org/10.3916/C62-2020-08

Llorens, L.-A. (2013). The ICT skills gap in Spain: Industry expectations versus university preparation. *Computer Applications in Engineering Education*.

López-Gil, B. B. (2020). Teaching in the Network Society: analysis of the digital competences of students in Education at the University of C diz. International journal of educational research and innovation, 53.

López-Meneses, E., Sirignano, F. M., Vázquez-Cano, E., & Ramírez-Hurtado, J. M. (2020). University students' digital competence in three areas of the DigCom 2.1 model: A comparative study at three European universities. Australasian Journal of Educational Technology, 69-88. https://doi.org/10.14742/ajet.5583

Mattar, J., Ramos, D. K., & Lucas, M. R. (2022). DigComp-Based Digital competence Assessment Tools: Literature Review and Instrument Analysis. Education and Information Technologies, 1-25.

Mengual-Andres, S., Roig-Vila, R., & Mira, J. B. (2016). Delphi study for the design and validation of a questionnaire about digital competences in higher education. International Journal of Educational Technology in Higher Education, 13(1), 12. https://doi.org/10.1186/s41239-016-0009-y

Spante, M., Hashemi, S. S., Lundin, M., & Algers, A. (2018). Digital competence and digital literacy in higher education research: Systematic review of concept use. Cogent Education, 5(1), 1519143. https://doi.org/10.1080/2331186X.2018.1519143

Wild, S., & Schulze Heuling, L. (2020). How do the digital competences of students in vocational schools differ from those of students in cooperative higher education institutions in Germany? Empirical Research in Vocational Education and Training, 12(1), 5. https://doi.org/10.1186/s40461-020-00091-y

Zamawe, F. C. (2015). The Implication of Using NVivo Software in Qualitative Data Analysis: Evidence-Based Reflections. *Malawi Medical Journal*, 3.

Zhao, Y., Llorente, A. M. P., & Gómez, M. C. S. (2021). Digital competence in higher education research: A systematic literature review. Computers & Education, 168, 104212.

Education and Learning Analytics

Barbi Svetec, Lea Oksanen, Blaženka Divjak, Damir Horvat
Digital Teaching in Higher Education during the Pandemic: Experiences in Four Countries

Francisco Ortin, Juan Ramón Pérez-Pérez, David Cabielles-Hernández, Miguel Sánchez-Santillán, MPuerto Paule-Ruiz
Mining Interaction Patterns of Children with and without Communication Disorders in the use of Tablets Apps

Stefan Svetsky, Oliver Moravcik
Design of All-In-One Technology-Enhanced Learning Software for Supporting Teachers' Personal Activities

Lourdes Guardia Ortiz, Maite Fernandez-Ferrer, Nati Cabrera, Federica Mancini, Marcelo Maina
Enhancing Self-regulation Strategies through Peer Feedback in Online Higher Education

Marcelo Maina, Lourdes Guàrdia, Nati Cabrera, Maite Fernandez-Ferrer, Maria Luisa Malerba, Carolina García
Evaluation of a Study Programme for the Implementation of Blended Learning at the Maldives National University

Melita Milić, Blaženka Divjak
Digital Maturity of Schools - Explanation, Literature Review and Analysis

Gordana Borotić, Gordan Gledec, Tomislav Jagušt
Exploring the Usage of an Adaptive Learning System for Elementary School Math Classes

Antonela Čižmešija, Goran Bubaš, Igor Balaban
A Preliminary Investigation of the Effects of Communication Quality and Use of Communication Platforms on Students' Perception of Virtual Team Performance

Snježana Babić, Darko Etinger, Monika Lenić
Implementing Educational Humanoid Robots in Croatian Secondary Schools: Potential Challenges and Benefits

Dragana Kupres, Nati Cabrera Lanzo, Albert Sangrà Morer
Digitally Mature Education: The Role of Digital Maturity and Educational Leadership in Meaningful Digital Transformation of Schools

Digital Teaching in Higher Education during the Pandemic: Experiences in Four Countries

Barbi Svetec
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, Varaždin, Croatia
basvetec@foi.hr

Lea Oksanen
Lappeenranta-Lahti University of Technology LUT
Yliopistonkatu 34, Lappeenranta, Finland
lea.oksanen@student.lut.fi

Blaženka Divjak, Damir Horvat
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, Varaždin, Croatia
{bdivjak, damir.horvat1}@foi.hr

Abstract. *This paper presents a study conducted within an international project to investigate the experiences of higher education (HE) teachers with the shift to digital teaching during the COVID-19 pandemic. The study was based on a survey conducted among HE teachers (n = 167) in autumn 2021 in four European countries: Spain, Portugal, Finland and Croatia. Analysis of Variance (ANOVA) was used to investigate the differences in responses with respect to the country and the study field, with more significant differences found among the countries than among the study fields. The results will contribute to a MOOC for teachers' professional development.*

Keywords. higher education, online learning, hybrid learning, blended learning, teachers, COVID-19 pandemic

1 Introduction

Online learning took full speed with the onset of the COVID-19 pandemic in spring 2020, when nearly all teaching and learning was shifted online, including in higher education (HE) (OECD, 2021). The prediction for the future of education online came true (Powell et al., 2015).

Challenges of online education were met differently by different countries, and even institutions, depending on their digital maturity. Well-prepared countries, with advanced IT strategies, infrastructure, and human potentials, as the key elements of digital maturity of educational institutions (Ifenthaler & Egloffstein, 2020; Đurek et al., 2017), shifted to the online learning mode nearly instantly, while for others the challenges were more overwhelming. There were also countries not performing the best in relation their overall digital maturity, but "fit" in terms of ongoing reforms and readiness for change (Svetec & Divjak, 2021). Yet, similar challenges were widely identified in shifting to online education; first is the ICT accessibility, affordability, and flexibility, followed by pedagogical competence, combined with online teaching and learning methodologies and modes (Murgatroyd, 2020). In other words, the pandemic not only raised technical and resource issues, but also brought forward the question of teachers' and students' digital competences. It raised questions on pedagogical practices in teaching, learning and assessment, by taking away the face-to-face interactions between learners and teachers.

Some higher education institutions (HEI) had already introduced practices contributing to the flexibility and accessibility of HE, including blended learning and hybrid learning (Bashir et al., 2021). However, in many countries, HEIs were not well prepared for the shift to predominantly online education, and there was often not enough experience and time to conceive new formats of teaching and learning (OECD, 2021). The pandemic made many HEIs speed up the developments related to e-learning, which was done in emergency, without structured long-term planning (Müller et al., 2020). After the initial closures of HEIs and mainly online learning, in some countries, possibilities for hybrid learning were introduced (OECD, 2021). In terms of future developments, research has shown that HE teachers appreciated blended and hybrid learning, allowing them to use the strengths of both face-to-face and online learning (Müller et al., 2020).

The change from traditional face-to-face learning to an online learning environment during the pandemic has been experienced differently by teachers and

students (OECD, 2021). Teachers have had an essential role in the continuation of learning during the pandemic (Council of the EU, 2020), and how teachers cope generally affects many aspects of the learning process (Jelińska & Paradowski, 2021). Some research in secondary and primary education has indicated that teachers' frequency, confidence and self-perceived proficiency in using digital technologies for teaching, as well as their motivation for improving their digital skills and using digital technologies have increased during the pandemic (Beardsley et al., 2021). However, many HE teachers were not fully satisfied with the transition to online teaching and learning (OECD, 2021), and some research has distinguished between teachers who were more engaged and better coping with the challenges of online teaching, and those who were struggling in both aspects (Jelińska & Paradowski, 2021). It has also suggested that teachers with prior experience with online teaching were the most engaged and the best coping (Jelińska & Paradowski, 2021). Similarly, it has been found that HE teachers who had used innovative pedagogies like flipped classroom before the pandemic, in the face-to-face or blended mode of delivery, had more success with their online implementation during the pandemic (Divjak et al., 2022). In this respect, it should be noted that the use of online tools by HE teachers (academics) before the pandemic was not widespread (OECD, 2021), which could have considerably affected their experiences with the shift to online teaching and learning. Moreover, digital competences for teaching and possibilities for professional development aimed at developing digital competences were found essential in adapting to online teaching (König et al., 2022).

In order to explore the experiences of HE teachers with digital teaching before and during the COVID-19 pandemic, and compare the perspectives of teachers in different HE systems and different study fields, we conducted a study within the international Erasmus+ project entitled *Digital and Entrepreneurial Skills for European Teachers* (eDesk). The project includes partners from Spain, Portugal, Finland and Croatia. The aim of the project is to develop a hybrid teaching methodology and encourage the development of teachers' competences by supporting their continuous professional development. In this context, our research was conducted as the basis for the preparation of a MOOC for HE teachers' professional development.

The study was focused on the following research questions (RQ):

RQ1: What were the experiences of HE teachers with the shift to digital teaching during the COVID-19 pandemic?

RQ2: Were these experiences different with respect to the country and the field of study?

Our study encompassed various modes of digital teaching, including fully online, blended, and hybrid. Though blended and hybrid teaching are distinguished from each other in this study, it needs to be noted that they have frequently been used interchangeably. Therefore, in discussing the two teaching modes, it is important to point out the difference within this study, as explained in the Methodology section.

2 Methodology

Our study included HE teachers in four European countries: Spain, Portugal, Finland, and Croatia. It was based on a survey carried out in the period between 24 August 2021 and 21 September 2021 (further: autumn 2021).

2.1 Instrument

The questionnaire was developed by the international project team. The first version was developed by the second author, at the LUT University, Finland, based on daily discussions with teachers, and subsequent discussions with IT staff. The questionnaire was then further discussed with the researchers from the international project team, primarily the third author, from the University of Zagreb, Croatia, and the final version was prepared.

The final version of the questionnaire included 28 questions, with the majority of questions being multiple-choice, in form of Likert-type items (from 1 - *Strongly disagree* to 5 - *Strongly agree*, or adjusted to a specific question). The remaining were multiple-choice questions with multiple answers possible, *Yes/No/I don't know* questions, and open-ended questions.

The questions were focused on teachers' experiences related to the shift to digital education, including their practices before and during the pandemic. Several groups of questions were included, focused on (1) perceived benefits of digitalization for teaching and learning, (2) use of digital technologies in teaching and learning, (3) teaching delivery modes, (4) teaching strategies and methods, (5) assessment, (6) self-assessment of competences, (7) organizational support, (8) interaction with students and peers, (9) entrepreneurial teaching. In this paper, due to the foreseen length, we analyse all the groups except for entrepreneurial teaching. As for the included groups, we present the key results, but not details related to all the questions.

To ensure focused responses, the questionnaire included the essential definitions, including those of the modes of delivery of teaching, defining:

- *fully online mode*, in which students complete courses entirely online;
- *blended mode*, in which the learning environment includes both online environment and face-to-face teaching;
- *hybrid mode*, in which students are simultaneously present in the same classroom, either physically or remotely.

Additionally, the questionnaire was designed to collect background data on participants' country,

gender, age, teaching experience (years), employment (permanent or not, full-time or part-time), type of a HEI (public or private), and the study field.

2.2 Participants

The questionnaire was distributed electronically by each project partner among their HEIs, and filled in by HE teachers teaching in various study fields. A total of 167 fully completed responses were collected: 36 in Spain, 73 in Portugal, 27 in Finland, and 31 in Croatia.

2.3 Data analysis

To respond to RQ1, related to the teachers' experiences with the shift to digital teaching during the pandemic, we used descriptive statistics to analyse the sample and the overall responses to the multiple-choice questions. Moreover, we used qualitative analysis to analyse the open-ended questions.

To respond to RQ2, we conducted one-way Analysis of Variance (ANOVA) tests, taking into account unbalanced designs (considering the differences in group sizes) on a subset of chosen questions, to determine whether there are statistical differences between the groups related to the independent variable. The first independent variable was the *country*, with four groups of data (Spain, Portugal, Finland, Croatia). The second independent variable was the *study field*, with six groups of data (Arts & Humanities, Social Sciences, Technical Sciences & Engineering, Science & Mathematics, Health Sciences, Other).

As ANOVA assumes that the observations within each group are normally distributed and have equal variances (De Veaux et al., 2012, pp. 727 – 729), additional tests were conducted before the ANOVA. To determine whether the groups have a normal distribution, histograms were analysed. To determine whether the variances are similar, the Levene test was conducted. For several questions in which the Levene test pointed to unequal variance, Welch's ANOVA was conducted (Dalgaard, 2008, p. 133). After the ANOVA, for variables (questions) that demonstrated significant differences between the groups, the Tukey test was additionally conducted, to investigate which specific groups' means are different in comparison to each other. The analyses were conducted in R, with *tidyverse* and *car* as essential packages.

With respect to both RQs, in this paper, we present only a sample of the most interesting results.

3 Results

3.1 Sample analysis

In the total sample (167 participants), including all the four countries, there were 85 female HE teachers (50.9%), 79 male (47.3%), and three who declared themselves as other or preferred not to say (1.8%). Their age ranged from 25 to 69 years, with teaching experience from one to 45 years. A majority of teachers held a permanent position (76%). Moreover, a vast majority worked full-time (93.4%). Finally, all the participants worked at public universities, except for one, who worked at a private university. As for the field of study, the distribution of the participants was as follows: Social Sciences (41), Technical Sciences & Engineering (41), Arts & Humanities (40), Science & Mathematics (29), Health Sciences (13), Other (3).

3.2 Experiences with the shift to digital education (RQ1)

In terms of the **perceived benefits** of digitalization for teaching and learning, a vast majority of the participants (strongly) agreed that digitalization *supports* teaching and learning (88%), as well as that it *enhances* teaching and learning (70.6%).

When it comes to the **use of digital technologies** in teaching before the pandemic, only a small portion of the participants (13.8%) reported having ICT fully embedded in their practice, but a majority (59.8%) reported either using ICT as a significant or regular feature (31.1%) or frequently using ICT with a varying impact (29.3%). Some of the participants also reported never (2.4%) or only occasionally (23.4%) using ICT in their practice. In autumn 2021, the proportion of the participants having ICT fully embedded in their practice increased almost three times (40.7%), while a majority (56.9%) still used ICT as a significant or regular feature (39.5%) or frequently with a varying impact (17.4%). While there were still some of the participants who reported using ICT only occasionally (2.4%), none responded they never used ICT in their practice. Similarly, a vast majority of the participants (78.4%) (strongly) agreed that the pandemic had had a positive effect on the use of digital technology in their institutions.

In relation to the **modes of delivery**, regarding the period before the pandemic, a vast majority of the participants (89.2%) reported using face-to-face teaching always or frequently. On the other hand, a vast majority reported using other delivery modes (fully online, blended, hybrid) occasionally, rarely, or never. In autumn 2021, the proportion of the teachers who reported using the face-to-face mode of delivery always or frequently decreased (59.9%), while the portion of the teachers who used other delivery modes (fully online, blended, hybrid) always or frequently increased. Details are presented in Fig. 1.

As for the **teaching methods** used in digital teaching, a majority of the teachers reported always or frequently using questions and answers (63.5%), while around a half of the participants reported always or frequently using problem-based (45.5%), project-based (45.5%), work-based (45.5%), collaborative and peer-learning (41.9%). Inquiry-based learning, flipped classroom, and learning based on cooperative models

(such as social economy or cooperative values) were reported as less frequently used. Details are presented in Fig. 2. In terms of other methods, in their comments, the teachers mentioned traditional lessons, individual lessons, presentations, guest speakers, debates and discussions, game-based learning, team-based learning, peer-learning, case studies, reading scientific texts, brainstorming sessions, field trips, and other.

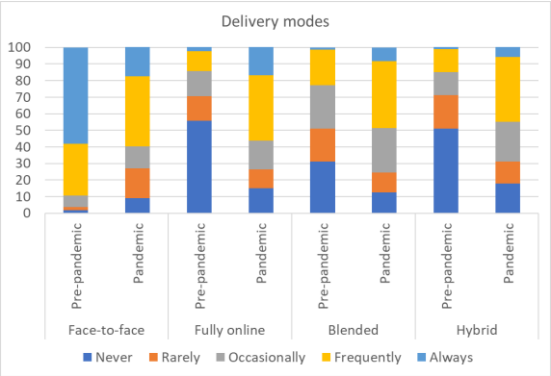


Figure 1. Delivery modes before the pandemic and in autumn 2021 (responses in percentages)

In relation to **assessment during the pandemic**, slightly more than a half of the participants (strongly) agreed that they tried to replicate the approaches used before (53.3%) and that they adapted assessment approaches used before to the online environment (55.7%). The participants expressed the highest level of agreement with the claim that they had to rethink assessment and introduce considerable changes, with which two thirds (strongly) agreed (67.7%). In terms of concrete assessment methods and tools, results suggested that more than a half of the participants applied more problem-solving exercises, while fewer participants introduced group reflection and peer-assessment, or used examination databases and self-assessment in a digital environment. Details are presented in Fig. 3. However, in their comments, some of the participants explained that they did not introduce peer-assessment and group reflection, as they had already used them before. They also mentioned using online formative quizzes, online tests, oral assessment, and interactive and metacognition activities, as well as rubrics supporting the assessment of complex tasks. Finally, a majority of the participants (60.5%) (strongly) agreed that they used continuous, formative digital assessment methods.

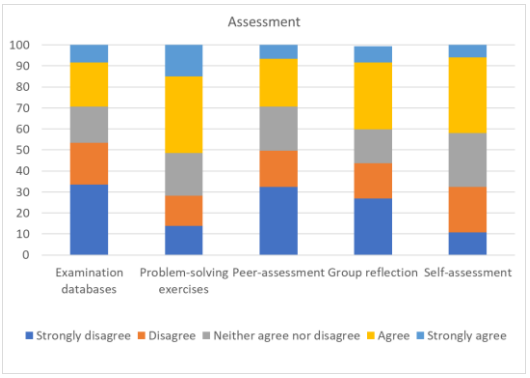


Figure 3. Assessment tools and methods (responses in percentages)

As for the participants' **self-assessment of competences**, more than a half (59.3%) (strongly) agreed that their digital teaching competence is very good, while around two thirds (70.7%) (strongly) agreed that their pedagogical competence is very good.

In terms of **organizational support**, about a half of the participants (strongly) agreed that their organization offers effective support for digital teaching (50.3%), but nevertheless, more than a half (56.3%) found their organization needed to offer more support to improve online teaching.

In terms of **available digital resources** at their institutions, the participants pointed out online platforms (162) and LMSs (151). These were followed by editing tools (58), polling tools (57), add-on software for including questions in Moodle videos (49), and game-based tools (49). As far as the **forms of ICT support** are concerned, the teachers preferred videos (102), followed by short courses (81), tutorials (81), manuals (79), and workshops (76).

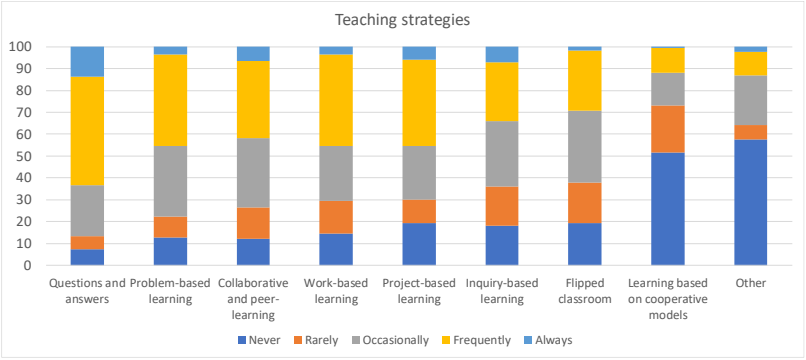


Figure 2. Teaching strategies used in digital teaching (responses in percentages)

As for **interaction** during the pandemic, only a minority (strongly) agreed that they interacted with students (21%) or peer teachers (16.2%), or noticed students interacted with their peers (15.6%) more than before. However, around a half of the participants (48.5%) (strongly) agreed with having noticed that students interacted more with available digital material.

In their responses to open-ended questions related to changes, successes, and challenges of digital teaching, teachers reported different experiences. While some of the participants mentioned that not much had changed with the pandemic as they had used digital tools and resources before, others had less successful experiences. Several of the participants pointed out that they made more digital materials, introduced new tools, prepared new course designs, and had to rethink their practices and enhance their own competences. Several mentioned more creativity in teaching and better access to information for students. But while some found that digitalization enriched and improved their teaching practice, others found it impoverished their practice and was overwhelming for students, and pointed out that digital teaching cannot replace face-to-face. Several of the participants also reported that digital teaching required more effort, workload, and preparation, but some pointed out that digitalization made their teaching more effective. While some reported that digitalization made teaching more impersonal, others stressed it made it more inclusive and open. Moreover, while some thought there was less contact, communication, and interaction with students, others mentioned good interaction and better relationships with students.

3.3 Differences between the countries and study fields (RQ2)

The ANOVA test pointed to a number of variables (questions) that showed significant differences with respect to the country, and fewer that showed significant difference with respect to the study field. (In further text, when the countries and study fields are compared, the one with a higher mean is marked with an asterisk.)

In relation to the **perceived usefulness of digitalization** in terms of supporting and enhancing teaching and learning, there were no significant differences neither between the countries nor between the study fields.

However, the results showed significant differences among the countries with respect to the **extent to which digital technology was used** for lessons before the pandemic ($p = 1.48e-06$, $F = 10.89$), but also in autumn 2021 ($p = 1.44e-05$, $F = 9.03$). With respect to both questions, the differences were the most significant between Portugal and Croatia*, and Portugal and Finland*. Regarding the use of digital technology in autumn 2021, there was a significant difference between the study fields as well ($p = 0.005$, $F = 3.48$), in particular between Social Sciences* and Arts & Humanities.

With respect to the statement that the pandemic has had a **positive effect on the use of technology** at the participants' institutions, the difference between the countries was not found to be significant. However, the difference was significant between the study fields ($p = 0.008$, $F = 3.22$), in particular between Health Sciences* and Arts & Humanities, and between Technical Sciences & Engineering* and Arts & Humanities.

Significant differences were also found between the countries in relation to all the **modes of delivery** before the pandemic: face-to-face (Welch: $p = 1.089e-09$, $F = 20.89$), fully online (Welch: $p = 3.113e-07$, $F = 13.87$), blended ($p = 7.85e-05$, $F = 7.68$), and hybrid (Welch: $p = 0.005$, $F = 4.65$). In autumn 2021, there was a significant difference with respect to face-to-face (Welch: $p = 9.8e-10$, $F = 20.86$) and fully online learning (Welch: $p = 1.649e-07$, $F = 14.51$), but no significant difference in relation to blended learning and hybrid learning. In relation to fully online learning, differences were the most significant between Spain and Croatia*, and Spain and Finland*. When it comes to the study fields, the only significant difference was found regarding the use of fully online teaching before the pandemic (Welch: $p = 0.0007$, $F = 6.83$).

In relation to **teaching and learning strategies**, a significant difference between the countries was found in problem-based learning ($p = 0.002$, $F = 5.03$) and work-based learning (Welch: $p = 0.02$, $F = 3.53$). For problem-based learning, there were significant differences between Portugal and Croatia*, and Spain* and Portugal. For work-based learning, there was a significant difference between Spain* and Finland. As for other teaching strategies, including flipped classroom, project-based learning, collaborative and peer learning, learning based on cooperative models, questions and answers, inquiry-based learning, there were no significant differences among the countries. Regarding the study fields, the only significant difference was found in relation to work-based learning (Welch: $p = 0.02$, $F = 3.44$), in particular between Technical Sciences & Engineering and Arts & Humanities*.

As for **assessment**, there were significant differences in relation to all the variables (questions), except for two, related to whether the participants tried to replicate assessment approaches used before and adapted their approaches to the online environment. Conversely, from the study field perspective, no significant difference was found in relation to the vast majority of variables.

In relation to the participants' **self-assessment of digital teaching competences**, a significant difference was found both among the countries (Welch: $p = 0.0002$, $F = 7.62$) and the study fields (Welch: $p = 0.02$, $F = 3.51$). In particular, in relation to the countries, there was a significant difference between Portugal and

Croatia*, and in relation to the study field, between Science & Mathematics* and Arts & Humanities.

When it comes to **organizational support**, there were significant differences among the countries both in whether the participants found their organization needed to offer more support to improve online teaching ($p = 7.9e\text{-}05$, $F = 7.67$) and whether it offers effective support for digital teaching ($p = 0.03$, $F = 2.99$). In relation to the former, there were significant differences between Croatia and each of the three countries*. In relation to the latter, there was a significant difference between Spain and Croatia*. Moreover, there was a significant difference between the study fields in relation to the need for more support from the organization ($p = 0.03$, $F = 2.65$), in particular Science & Mathematics and Health Sciences*, and Social Sciences and Health Sciences*.

In relation to **interaction**, there was no significant difference with respect to either the country or the study field.

4 Discussion

Our research, even though conducted on a relatively small sample, presents an overview of HE teachers' experiences with the shift to digital education in four countries with different educational traditions. The results mainly confirm the already existing knowledge and assumptions related to the shift to digital education during the pandemic (OECD, 2021), but also offer some interesting and specific insights, especially useful for the preparation of the project's MOOC for professional development of teachers. It should be noted that the study was conducted in autumn 2021, which was a year and a half into the pandemic, when certain teaching and learning practices had already been adapted to the pandemic circumstances. This was no longer the emergency response as the one during the onset and the peak of the pandemic. Nevertheless, the results still suggested a decrease in face-to-face teaching and learning, and an increase in other digital modes of delivery, including fully online, hybrid and blended. This suggests that the pandemic has fostered the digital transformation of HE in a sustainable way.

The results show that HE teachers widely recognize that digitalization *supports*, but are not so much aware that it *enhances* teaching and learning. This may be related to the less positive experiences of some teachers with the shift to digital teaching, and the fact that some found that digitalization impoverished their practice. This may, in turn, relate to the fact that not a negligible proportion of teachers (40.7%) found their digital teaching competence not very good. In this respect, it is not surprising that teachers in Science & Mathematics found their digital teaching competence significantly better than teachers in Arts & Humanities.

Moreover, all of this could also be linked with the level of organizational support, as more than a half of the participants (56.3%) found that their institutions should offer more support for online teaching. In this respect, the ANOVA showed that Croatian teachers reported the highest level of the existing support and the lowest level of the need for additional support. The high level of the need for more support is consistent with the small proportion of teachers (13.8%) who reported having ICT fully embedded in their practice before the pandemic. In relation to innovative pedagogies, previous research reviews have found that HE teachers who had used innovative pedagogies such as flipped classroom before the pandemic were more successful in their implementation during the pandemic (Divjak et al., 2022). However, the causality of this assumptions has not been tested in this research and should be investigated further.

Among the study fields, the highest level of the need for more support was reported in Health Science, which can be related to the fact that this study field includes a lot of practical work, which has also been reported as challenging during the pandemic (OECD, 2021).

It is no surprise that the proportion of teachers who reported having ICT fully embedded in their practice increased almost three times during the pandemic (from 13.8% to 40.7%), and that in autumn 2021, none reported never using ICT in their practice. However, it is interesting to note that some teachers (2.4%) still reported they used ICT only occasionally.

Expectedly, the results confirm that the use of face-to-face teaching decreased, while the use of all other modes of delivery (fully online, blended, hybrid) increased during the pandemic. Here, it is interesting to note that in autumn 2021, there was a bigger emphasis on fully online learning in comparison to blended and hybrid learning (between which there was no significant difference between the countries). The fact that hybrid learning was the least used may be related to the fact that its implementation is also quite complex and requires substantial effort on behalf of the teacher (Goodyear, 2020), which confirmed the need for guidance and methodology, such as those provided by the eDesk project.

It is commendable that the teachers often used innovative strategies like problem-based, project-based, collaborative and peer-learning, as well as work-based learning, which might have been challenging during the pandemic. Some teachers mentioned laboratory classes (practical work) as particularly challenging, but they also mentioned a solution in the form of virtual laboratories. The only significant differences in relation to the teaching and learning strategies were found between the countries by the ANOVA regarding problem-based learning (most frequently used by Croatian and Spanish teachers) and work-based learning (most frequently used by Spanish teachers). Unfortunately, flipped classroom and inquiry-based learning were not widely used, even though they provide substantial opportunity for students' active learning and engagement in online learning (Divjak, 2022; Zhang et al., 2021).

Another interesting result also refers to the extent to which digital technology was used before and during the pandemic. Namely, the result showing significant differences between the same groups of countries (Portugal and Croatia, Portugal and Finland) before and during the pandemic may indicate that a country's response to the pandemic was linked with its digital readiness before the pandemic (similar was indicated in relation to pre-tertiary education by Svetec & Divjak, 2021). This can also be related to the digital maturity of HEIs (Đurek et al., 2017). Similarly, when it comes to the perceived positive effect of the pandemic on the use of technology, significant differences between the study fields, in particular between Arts & Humanities and Technical Sciences & Engineering, may also suggest that the extent to which the pandemic changed the teaching practices depended on how much a study field is technology-related, with a stronger effect reported in Technical Sciences & Engineering.

Finally, interaction with (and among) students and peers was confirmed as one of the most challenging aspects of the shift to digital teaching, without significant differences between either countries or study fields. This is in line with previous research, and corresponds with the results suggesting that active learning strategies such as flipped classroom and inquiry-based learning were not widely used.

4.1 Limitations and future research

One of the limitations of our study is the size of the sample, and including a wider range of participants may have elicited different results. However, it should be noted that the population included in the study were HE teachers, who are not as numerous as students, and might not always have the capacity to take part in extensive surveys as this one. Moreover, the sample was not balanced, as more responses were collected in some countries than in other. The sample was also not randomly assigned, which was due to the specific context and purpose of the project (learning design of the MOOC). It should also be noted that the study was not longitudinal, and the teachers were asked to report on their practices before the pandemic retrogradely. Having already changed their practices during the pandemic might have influenced their answers about the pre-pandemic period. In terms of further research, it would be worth exploring the experiences of HE teachers in other countries in other to make additional comparisons, as well as broaden the sample in the already included countries. Moreover, the study indicates a difference in the implementation of technology in teaching, and self-assessment of digital competences, with respect to specific study fields. However, to make more definite conclusions, further original research or meta-analyses are needed. It would be especially interesting to relate teachers' perceptions with the digital maturity of their HEIs. Finally, it would be useful to conduct longitudinal research, to explore whether and how the pandemic has changed the practices of HE teachers in long term. This paper presents only a summary of the key findings due to the prescribed length. However, more detailed insights into the findings of the conducted analyses may provide deeper insights into the topic.

5 Conclusion

We presented a survey design research (n = 167) aimed at exploring the experiences of HE teachers with the shift to digital teaching during the COVID-19 pandemic in four countries: Spain, Portugal, Finland and Croatia.

The study found that in autumn 2021, the portion of teachers having ICT fully embedded in their practice was almost three times higher than before the pandemic. In autumn 2021, the portion of teachers who reported using the face-to-face mode of delivery always or frequently decreased, while the portion of teachers who used other delivery modes (fully online, blended, hybrid) always or frequently increased. As autumn 2021 was no longer the period of the most serious emergency, this suggests that the pandemic has triggered a sustainable digital transformation of HE. The results also indicated that some innovative pedagogies, like problem-based, project-based, collaborative and peer-learning, as well as work-based learning, were often used, whereas others, like flipped classroom and inquiry-based learning were not used in their full potential.

We analysed the results to determine whether there are significant differences in teachers' responses with respect to the country and the study field. Among the countries, there were significant differences related to the extent of use of digital technologies in teaching before and during the pandemic, delivery modes, some teaching and learning strategies, assessment, teachers' self-assessment of digital competences, and organizational support. Among the study fields, there were fewer significant differences, and those were primarily related to the use of digital technologies during the pandemic and the positive effect of the pandemic on the use of digital technologies, teachers' self-assessment of digital competences, and organizational support. In relation to teaching and learning strategies, there was a significant difference in the use of work-based learning.

It would be valuable to conduct further research with a broader sample and including other countries, as well as longitudinal research. It would also be worth digging deeper into the established differences between the countries and study fields. Finally, future research could consider the relation between teachers' perceptions and the digital maturity of their higher education institutions.

Acknowledgments

This study was conducted within the project *Digital and Entrepreneurial Skills for European Teachers*, financed from the Erasmus+ Programme of the European Union. The sole responsibility for the content of this article lies with the authors. It does not necessarily reflect the opinion of the European Union.

References

Bashir, A., Bashir, S., Rana, K., Lambert, P., & Vernallis, A. (2021). Post-COVID-19 Adaptations; the Shifts Towards Online Learning, Hybrid Course Delivery and the Implications for Biosciences Courses in the Higher Education Setting. *Frontiers in Education, 6*. https://doi.org/10.3389/feduc.2021.711619

Beardsley, M., Albó, L., Aragón, P., & Hernández-Leo, D. (2021). Emergency education effects on teacher abilities and motivation to use digital technologies. *British Journal of Educational Technology*, *52*(4), 1455–1477. https://doi.org/10.1111/bjet.13101

Council of the EU. (2020). Council conclusions on European teachers and trainers for the future. *Official Journal of the European Union.* (2020/C 193/04)

Dalgaard, P. (2008). *Introductory Statistics with R*. Springer New York. https://doi.org/10.1007/978-0-387-79054-1

De Veaux, R. D., Velleman, P. F., Bock, D. E. (2012) *Stats: Data and Models (3rd ed.).* Pearson Education, Inc.

Divjak, B., Rienties, B., Iniesto, F., Vondra, P., & Žižak, M. (2022). Flipped classrooms in higher education during the COVID-19 pandemic: findings and future research recommendations. *International Journal of Educational Technology in Higher Education*, *19*(1), 9. https://doi.org/10.1186/s41239-021-00316-4

Đurek, V., Begičević Ređep, N., Divjak, B. (2017). Digital Maturity Framework for Higher Education Institutions. *Proceedings of the Central European Conference on Information and Intelligent Systems.* September 27-29, 2017, Varaždin, Croatia.

Goodyear, P. (2020). Design and co-configuration for hybrid learning: Theorising the practices of learning space design. *British Journal of Educational Technology*, *51*(4), 1045–1060. https://doi.org/10.1111/bjet.12925

Ifenthaler, D., & Egloffstein, M. (2020). Development and implementation of a maturity model of digital transformation. *TechTrends, 64*, 302–309. doi:10.1007/s11528-019-00457-4

Jelińska, M., & Paradowski, M.B. (2021). Teachers' engagement in and coping with emergency remote instruction during COVID-19-induced school closures: A multinational contextual perspective. *Online Learning Journal*, *25*(1), 303-328. https://doi.org/10.24059/olj.v25i1.2492

König, J., Jäger-Biela, D. J. & Glutsch, N. (2020). Adapting to online teaching during COVID-19 school closure: teacher education and teacher competence effects among early career teachers in Germany. *European Journal of Teacher Education*, *43*:4, 608-622, DOI: 10.1080/02619768.2020.1809650

Murgatroyd, S. (2021). Insight: A Cancelled Future. What's Next for Universities and Colleges?. In: Burgos, D., Tlili, A., Tabacco, A. (eds) *Radical Solutions for Education in a Crisis Context.* Lecture Notes in Educational Technology. Springer, Singapore. https://doi.org/10.1007/978-981-15-7869-4_22

Müller, A. M., Goh, C., Lim, L. Z., & Gao, X. (2021). COVID-19 Emergency eLearning and Beyond: Experiences and Perspectives of University Educators. *Education Sciences, 11*(1), 19. https://doi.org/10.3390/educsci11010019

OECD. (2021). *The State of Higher Education: One Year into the COVID-19 Pandemic*. OECD Publishing, Paris. https://doi.org/10.1787/83c41957-en

Powell, A., Watson, J., Staley, P., Patrick, S., Horn, M., Fetzer, L., Verma, S. (2015). *Blending Learning. The Evolution of Online and Face-to-Face Education from 2008–2015.* International Association for K–12 Online Learning (iNACOL). Retrieved from https://www.inacol.org/resource/blending-learning-the-evolution-of-online-and-face-to-face-education-from-2008-2015/

Svetec, B. & Divjak, B. (2021). Emergency responses to the COVID-19 crisis in education: a shift from chaos to complexity. *European Distance and E-Learning Network (EDEN) Proceedings*, 2021 Annual Conference, Madrid, 21 – 24 June 2021. doi: 10.38069/edenconf-2021-ac0051

Zhang, L., Carter, R. A., Qian, X., Yang, S., Rujimora, J., & Wen, S. (2022). Academia's responses to crisis: A bibliometric analysis of literature on online learning in higher education during COVID-19. *British Journal of Educational Technology*, *53*(3), 620–646. https://doi.org/10.1111/bjet.13191

Mining Interaction Patterns of Children with and without Communication Disorders in the use of Tablets Apps

Francisco Ortin, Juan Ramón Pérez-Pérez, David Cabielles-Hernández, Miguel Sánchez-Santillán, MPuerto Paule-Ruiz

Computer Science Department

University of Oviedo

c/Federico García Lorca 18, 33007, Oviedo

{ortin, jrpp, uo209910, sanchezsmiguel, paule}@uniovi.es

Abstract. *A communication disorder (CD) is an impairment in the ability to receive, send, process, and comprehend concepts or verbal, nonverbal, and graphic symbol systems. Different research works pursue the early detection of communication disorders, because its treatment at early ages shows significant benefits. In a previous study, we developed two applications for tablet devices that help children in the process of learning the sounds and writing of letters and words. Using the Montessori educational method, our applications showed important benefits, for both children with and without communication disorders, in the learning process of letters, words, and their corresponding sounds. In this article, we use the interaction information produced by our applications in the learning sessions. The purpose of our research is to see if there exist particular interaction patterns of children with and without communication disorders, for the two given applications. We use different data mining techniques and algorithms to process the interaction data generated by 353 children through at least nine sessions. There exist statistically significant differences in 7 of the 36 interaction variables measured. For some sessions of both applications, children with CD made more mistakes than those without CD. The only significant interaction pattern retrieved from the data is that the children with the 12% lowest number of taps over letters (3 taps at most) have CD, fulfilled by 28.1% of the children with CD that used one of our applications. This group of children might be representing children with receptive language disorders.*

Keywords. Communication disorder, interaction pattern, tablet applications, data mining

1 Introduction

Communication disorders (CDs) limit the individual's ability to receive, send, process, and comprehend concepts or verbal, nonverbal, and graphic symbol systems (Collins, 2011). CDs include hearing loss, voice and speech disorders, and, on the most essential level, language disorders. It has been estimated that 5% to 10% of Americans may have communication disorders, costing approximately $154 to $186 billion annually (Ruben, 2000). CDs are associated with an unemployment rate of 41.9%, and the population income with CD is 45% of the income of the non-impaired ones (Ruben, 2000). Likewise, the Bureau of Labor indicates that at least 92% of US employment requires good communication skills (Herman & Abraham, 2000). This shows how, nowadays, CDs represent a substantial public health issue (Ruben, 2000).

The early detection and treatment of CDs have been pursued to provide early intervention, a better quality of life, and inclusion in society. Detection is performed at the earliest stages of life with different educational strategies (Paul, 2016). For example, the First Words Project uses the CSBS Developmental Profile (Wetherby & Prizant, 2002) to detect CDs in children less than 24 months old (Wetherby et al., 2003). Infants are first screened with a brief parent-report checklist, followed up with a more in-depth parent report tool and face-to-face evaluation. The findings support the importance of prelinguistic predictors and the role of the family in the early identification of CDs.

The Montessori educational method for children seeks to develop natural interests and activities rather than use formal teaching methods (Montessori, 2021). Created by the physician Maria Montessori, it emphasizes independence, since children are naturally eager for knowledge and capable of initiating learning in a well-prepared environment. Different applications for mobile devices have been developed to apply the Montessori method through different stages in children's learning process. Such applications cover learning letters and words, recognizing phonemes and word sounds, up to reading and creating phrases and sentences (Jones, 2016; McKenzie, Spence & Nicholas, 2018).

The Montessori method has already been successfully used with people with CDs, including

children (Pickering, 2017) and even elders with communication disabilities (Douglas, Brush & Bourgeois, 2018). To support that capability, we developed two tablet applications that, following the Montessori method, were designed to help children in the process of learning the sounds and writing of letters and words (Pérez-Pérez et al., 2021). In a former study, our two applications were successfully used in a school with children between 3 and 5 years old, under the supervision of their teachers but with no intervention (Pérez-Pérez et al., 2021). There were children diagnosed with CD (48.3%) and without CD (51.7%) and, for both cases, our tablet applications represented important benefits in the learning process of letters, words, and their corresponding phonemes and sounds (Pérez-Pérez et al., 2021).

Our software records data about how children interact with the tablets to solve the educational activities proposed by the two tablet applications. Such data might represent valuable information to be mined and analyze whether there are differences in the way children interact with the applications. In particular, we search for (combinations of) variables representing interaction patterns that are present in the recorded data to be mined, and see whether that could be used to identify a child has (no) CD.

Therefore, in this article, we raise the following research question:

Are there interaction patterns capable of identifying children with and without communication disorders (for the two given educational tablet applications)?

To answer the research question, we mine the interaction data produced by our two programs with 6 different data mining techniques, 19 algorithms, and two different use-case scenarios of both tablet applications, utilized by 353 children.

The rest of the article is structured as follows. The next section details related work, while the two tablet applications are briefly described in Section 3. Section 4 presents the methodology of our study, and Section 5 discusses the results. Conclusions are drawn in Section 6.

2 Related Work

Considering the DSM-5 taxonomic and diagnostic manual for mental disorders, there are two main classes of CDs (DSM-5, 2014). Language disorders represent difficulties in learning and using language, mainly caused by issues with vocabulary, grammar, and sentence construction. Language disorders can both be receptive (understanding the language) and expressive (synthesizing language constructions). The second class of CDs is speech sound disorders, when there exist issues with pronunciation and articulation in the native language.

The First Words project is aimed at building an evaluation model for identifying children at risk for CDs (Wetherby et al., 2003). The model is focused on children younger than two years old, because of the significant benefits of early intervention. The First Words project uses the Communication and Symbolic Behavior Scales - Developmental Profile (CSBS DP) to measure prelinguistic communication (Wetherby & Prizant, 2002). First, children are screened with a brief form filled in by the parents. Then, the children have a face-to-face evaluation and a more in-depth parent report. Two studies with 232 and 246 children showed that the First Words project can be used as a prelinguistic predictor, emphasizing the important role of the family in using the method to identify CDs at the early stages of life.

There are different software applications to aid children with CDs. Frutos et al. developed a videogame for children and teenagers to help them in the learning and enhancement of habitual language tasks (Frutos et al., 2011). The game encourages the children to repeat the words pronounced by the system for the shown pictogram. The voice signal of the child is analyzed to check whether the word has been pronounced correctly. The correct pronunciation makes the pictogram be visually stored in a success box. The game ends when all the boxes are filled. Frutos et al. do not present a statistical analysis of the performance of their game in the learning process.

Toki and Page developed a web application for children with language learning and speech articulation problems (Toki & Pange, 2010). The pedagogical model of their application is based on the Nearest Neighbor Learning (NNL) method, where students work in groups formed by themselves and are free to move from one group to another (Toki & Pange, 2006). Their web application is based on pronunciation activities in a talent show game. 12 children between 5 and 6 years old were selected to evaluate the software. The usage of the program showed an average benefit of 7% for speech articulation tasks. For language activities, the performance benefit was 4%.

The Montessori method has been included in different technological systems for distinct learning scenarios. TriPOD is a prototype app for tablets, addressed for educational applications (Di Fuccio, Siano & De Marco, 2017). The app was designed to exploit the psycho-pedagogical practices aimed at focalizing the centrality of the manipulation of real objects in the learning processes. They use different tangible objects, equipped with capacitive pins throughout their different shapes. Those objects are recognized by the software when they placed are onto the tablet surface. Children can play with the system autonomously, without the supervision of a teacher or tutor, following the Montessori method. TriPOD recognizes 24 objects, the Dienes' logic blocks, supporting the implementation of different didactic games. To our knowledge, no study to use TriPOD for learning purposes has been undertaken.

There are some other applications for tablets that have been used with the Montessori method for multi-

sensory literacy development intervention (Smith, 2012). Examples of such applications are WordBingo, Montessori Crosswords, and ABC Spelling Magic. The interactivity of the tablets allows children to use their fingers to manipulate the screen, creating an emotional experience and enhancing memory. Educators have realized the effectiveness of these devices, together with the impact on learning, for different intervention programs including dyslexia, dyscalculia, and attention deficit hyperactivity disorder (Drigas & Gkeka, 2017).

3 Tablet Applications

Our two tablet applications deal with reading and writing skills during the first educational stages of children. The first application, Matching Cards, gives a lexical awareness of words. It is not necessary to have previous knowledge of word spelling, since the task is focused on comparing shapes. The second one, Cards & Sounds, develops phonological skills by discovering how words are pronounced, comparing them, and finding out which sound each word begins with. Instructions are given orally by the applications at every step.

The Matching Cards is based on the Picture Exchange Communication System (PECS), using cards with pictograms and written words (Bondy & Frost, 2008). Cards about the same item should be matched by drag and drop actions. Tapping on the cards produces a sound related to the tapped item. Correctly matched cards are grouped, whereas wrong matches produce moving cards to their original position.

The left-hand side of Cards & Sounds shows boxes with all the letters, which play the corresponding phoneme when they are tapped. Four images (pictograms) appear on the right-hand side of the screen (they are also pronounced when tapped). The child should drag each image and drop it over the corresponding letter. When a letter is matched, its background color becomes gray and the card disappears to indicate that the action is correct; otherwise, the drag-and-drop action is revoked.

The interaction data stored by both applications are:

- `TotalTime` (real number). How long it took the child to solve the activity, expressed in minutes.
- `AverageTime` (real number). The average number of seconds measured between each interaction (tap or drag and drop action) of the child with the application.
- `Taps` (integer number). Number of taps performed by the student in each session. For Matching Cards, it is just one variable; but two variables are captured for Cards & Sounds:
  - `LetterTaps` (integer number): Number of taps on letters.
  - `WordTaps` (integer number): Number of taps on words.
- `DragDrops` (integer number): Number of drag and drop actions undertaken by the children during the session.
- `Mistakes` (integer number): Number of wrong matches performed by the student during the session.
- `ConsecutiveMistakes` (integer number): The greatest number of consecutive mistakes performed by the child during the session.

Each session requires the child to match the items (two for Matching Cards and four for Cards & Sounds) of four different panels. The data stored is anonymous so no personally identifiable information can be obtained. For a more detailed description of both applications, please consult (Pérez-Pérez et al., 2021) and (Cabielles, 2021).

4 Methodology

4.1 Context

We contacted Montessori associations and public schools in the north of Spain, where nursery education is addressed to children between 3 and 5 years old with and without CD. The children with CD included in our study had been previously diagnosed with language disorder and speech sound disorder. The diagnosis is supplied by the schools and is based on DSM-5 (DSM-5, 2014). Caregivers were responsible for installing the specific activities for the children and, for such purpose, they gave explicit consent. All the caregivers and teachers involved in the study were delivered a face-to-face seminar about how to use the applications and avoid adult intervention during sessions.

The number of children who participated in our study, together with their ages and genders, are depicted in Table 1. In total, 166 children participated in Matching Cards (48.8%/51.2% with/without CD and 45.2%/54.8% males/females) and 186 in Cards & Sounds (47.8%/52.2% with/without CD and 46.8%/53.2% males/females). The average age of students was 5.38 (standard distribution of 1.24).

Table 1. Number (N), arithmetic mean and standard deviation (SD) of ages, and genders of the children who participated in our study.

	Matching Cards		Cards & Sounds	
	Children with CD	Without CD	Children with CD	Without CD
N	81	85	89	97
Age mean	5.5	5.4	5.3	5.4
Age SD	1.2	1.1	1.3	1.2
Males	43	39	43	44
Females	38	46	46	53

4.2 Sessions

The study took place in the first trimester of the academic course. The two case studies (Matching Cards and Cards & Sounds) were carried out in 12 sessions (one session per week). The maximum duration of each session was 15 minutes. As mentioned, each session consisted in matching the items of four different panels. Each child worked independently. No interaction existed among children or between children and teachers or caretakers, following the Montessori method.

Since children of that age often get ill, we considered those who at least attended 75% of the sessions (i.e., 9 sessions out of 12). We considered the data gathered by our two applications in both the first and the last sessions. The purpose is to extract the interaction patterns for the first contact with the learning activity (no former interaction with the application) and once they have experience with the program. This is because differences between children with and without CD might exist in one scenario but not in the other one.

4.3 Data Mining

We mined the four datasets gathered from the log information produced by the two applications for the first and last sessions. We used the following data mining techniques.

4.3.1 Anomaly detection

Anomaly or outlier detection aims to identify unusual data records that may be interesting to analyze. If all the records with anomalous values for a given variable (do not) have CD, that would provide valuable information regarding our research question. Anomalous records may also be caused by measurement errors and, in that case, they should be removed.

For univariate outlier detection, we used Tukey's fences, which is a non-parametric, robust and widespread method to detect outliers. With Tukey's fences, an outlier is defined as an instance that does not belong to the following interval, where Q_n represents the n quartile:

$$[Q_1 - 3\times(Q_3 - Q_1),\ Q_3 + 3\times(Q_3 - Q_1)] \qquad (1)$$

For multivariate anomaly detection, we used the isolation forest algorithm that identifies outliers by considering how far a data point is from the rest of the data. We selected this algorithm because it has been measured to provide better performance than other methods such as ORCA, LOF and Random Forests (Liu, Ting & Zhou, 2008).

The contamination hyperparameter specifies the proportion of outliers in the dataset. We found 1% (0.01) as the contamination value that identified outliers in our datasets the best.

4.3.2 Decision tree classification

Decision tree learning is a supervised learning technique to predict values from previous observations. Decision trees are used to foresee one target value from the known features of a given sample. When the values to predict are discrete, decision trees act as classifiers; for continuous values, they become regressors.

A good characteristic of decision trees is that they are easy to understand and interpret. That is why they are useful for both machine learning and data mining. For example, the decision tree in Figure 1, which classifies children after the first interaction with Cards & Sounds, can be used to analyze that dataset. The children who tapped three or fewer times on a letter are diagnosed with CD. 25 out of 186 children fulfilled this condition (13.4%). If `LetterTaps` is greater than 3, other conditions should be checked to know whether the child has a CD. Each path from the root node to a leaf (i.e., a tree branch) represents a classification rule.

An important hyperparameter in decision trees is the maximum depth of the tree (the tree in Figure 1 has a max depth of 3). For deep trees, the predictive model overfits since one branch per sample could be created. Such over-complex models do not generalize well for the training data, and hence they do not produce valuable information. For this reason, we created decision trees with maximum depths from 2 to 10 and analyzed all the trees obtained (see Section 4). Classification rules are then extracted and analyzed to see if they represent an interaction pattern to identify children with(out) CD. We used the CART algorithm to build the decision trees because it supports numeric features and tree pruning as a regularization mechanism.

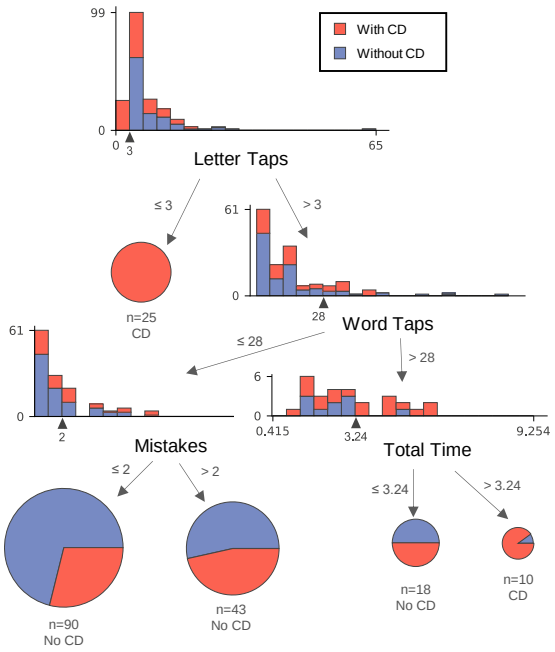


Figure 1. Decision tree for the first session of Cards & Sounds (max depth = 3).

4.3.3 Classification rules induction

Rule induction is a technique in which formal rules are obtained from a set of observations. They can be used to create predictive models (machine learning) and to represent data patterns (data mining). A classification rule is a collection of propositional predicates for a given value of the target feature. For example, the following predicates are used to classify children with CD in the first session of Cards & Sounds:

$$\texttt{LetterTaps} \leq 1 \vee (\texttt{Mistakes} \geq 7 \wedge \texttt{WordTaps} \geq 32)$$

The previous rule[1] states that, if a child performs one or no letter taps, or they have more than 7 mistakes and more than 31 word taps, then they have a CD; otherwise, the child has no CD.

There exist different algorithms for classification rule induction. We used the RIPPER*k* and IREP algorithms. The former usually obtains error rates lower than the C4.5 algorithm, scales nearly linear to the number of training instances, and is able to efficiently process noisy datasets (Cohen, 1995). The latter includes an incremental reduced error pruning process that avoids overfitting with noisy data and provides good generalizations (Fürnkranz & Widmer, 1994).

4.3.4 Association rules learning

Association rule learning is a method to discover relations between variables in existing datasets. Let $I=\{i_1, i_2, \ldots i_n\}$, be a set of n binary features, then an association rule is defined as:

$$X \Rightarrow Y \quad where\, X, Y \subseteq I \quad and \quad X \cap Y = \emptyset \qquad (2)$$

One important requirement of the i_n features in the dataset is that they must be binary. To meet this prerequisite, we discretized the numeric features into five intervals, following a quantile strategy to define the widths of the bins. In this way, each numeric feature is converted into one of the following five: very low (first quintile), low, average, high, and very high (last quintile). With this discretization, the following example rule was extracted from the last session of our Matching Cards dataset:

$$\texttt{Mistakes} = \mathit{VeryLow} \wedge \texttt{CD} = \mathit{False} \Rightarrow \texttt{ConsecutiveMistakes} = \mathit{VeryLow}$$

It tells us that, when children have very low mistakes (within the 25th percentile) and they do not have been diagnosed with CD, the number of consecutive mistakes is also very low. This shows how association rules can be used to find interaction patterns that could provide information to answer our research question.

Support and confidence are two widespread metrics to measure the performance of association rules. They are defined as:

$$Support(X \Rightarrow Y) = \frac{number\ of\ instances\ containing\ X\ and\ Y}{total\ number\ of\ instances} \qquad (3)$$

$$Confidence(X \Rightarrow Y) = \frac{number\ of\ instances\ containing\ X\ and\ Y}{number\ of\ instances\ containing\ X} \qquad (4)$$

The previous example rule has 28.9% support and 100% confidence.

For our purpose of finding interaction patterns of children with and without CD, we only analyze the association rules with the following criteria. They should have a minimum support threshold of 5% to be considered as representative rules, and at least 90% confidence to tolerate some noise in data. Rules must have the CD binary attribute as the only consequent, since we want to find different patterns for children with and without CD. We used the FP-Growth (Frequent Pattern Growth) algorithm instead of the traditional and widespread Apriori method, because FP-Growth efficiently represents the item sets as a tree of frequent patterns, reducing execution time and memory consumption (Han, Kamber & Pei, 2012).

4.3.5 Dimensionality reduction

Dimensionality reduction is the transformation of data from a high-dimensional space into a low-dimensional one, retaining meaningful properties of the original data. The four datasets used in our study have six and seven dimensions (Section 3). By applying dimensionality reduction techniques, the high-dimensional datasets can be embedded in a low-dimensional space for visualization. In those visualizations, points represent instances (children in our study) in such a way that similar instances are modeled by nearby points and dissimilar instances are represented by distant ones. By plotting children with and without CD with different colors, it is possible to visually identify different patterns for the two groups. Visual groups (clusters) made up of close points with the same color mean that there exists an interaction pattern associated with CD or the absence of it (our research question).

We used eight different algorithms: Principal Component Analysis (PCA), Non-Negative Matrix Factorization (NMF), Factor Analysis (FA), Sparse PCA, Truncated Singular Value Decomposition (SVD), t-distributed Stochastic Neighbor Embedding (t-SNE), Fast Independent Component Analysis (FastICA), and Kernel PCA. The first five algorithms are linear and the three last ones are nonlinear, thereby covering a wide set of alternatives for dimensionality

[1] The rule consists of the mentioned predicates as the antecedent and "CD = *True*" as the consequent.

Table 2. Number of instances (N), arithmetic means, standard deviations (SD), and min and max values of the four datasets (bold font represents variables with significant differences between children with and without CD).

			With CD (N=81)			Without CD (N=82)		
			Mean	SD	Range	Mean	SD	Range
Matching Cards	First Session	TotalTime	1.55	1.10	[0.43 – 7.37]	1.40	0.86	[0.45 – 6.75]
		AverageTime	2.41	1.75	[0.77 – 13.81]	2.22	1.40	[0.77 – 11.91]
		Taps	39.49	7.68	[32 – 78]	38.99	7.05	[32 – 76]
		DragDrops	20.57	6.00	[15 – 51]	19.82	5.02	][15 – 46]
		Mistakes	1.93	2.39	[0 – 11]	2.16	2.40	[0 – 13]
		ConsecutiveMistakes	0.99	1.23	[0 – 8]	1.16	1.04	[0 – 6]
	Last Session	TotalTime	1.16	0.71	[0.45 – 4.84]	1.33	1.21	[0.42 – 7.69]
		AverageTime	1.86	1.03	[0.64 – 6.29]	2.14	1.66	[0.75 – 11.18]
		Taps	38.07	5.80	[32 – 59]	37.26	7.06	[32 – 87]
		DragDrops	18.95	4.22	[15 – 38]	18.79	6.34	[15 – 68]
		Mistakes	**2.12**	**2.16**	**[0 – 8]**	**1.47**	**1.52**	**[0 – 8]**
		ConsecutiveMistakes	1.15	1.15	[0 – 7]	0.86	0.80	[0 – 3]
			With CD (N=89)			Without CD (N=97)		
			Mean	SD	Range	Mean	SD	Range
Cards & Sounds	First Session	TotalTime	1.88	1.55	[0.41 – 9.25]	1.55	0.99	[0.47 – 5.92]
		AverageTime	2.25	1.74	[0.53 – 11.81]	2.00	1.08	[0.71 – 6.84]
		LetterTaps	**5.61**	**5.00**	**[0 – 25]**	**7.96**	**7.69**	**[4 – 65]**
		WordTaps	**24.44**	**8.91**	**[16 – 63]**	**21.02**	**5.93**	**[16 – 51]**
		DragDrops	49.52	13.53	[33 – 103]	46.48	12.19	[36 – 116]
		Mistakes	**3.47**	**3.54**	**[0 – 14]**	**1.51**	**2.38**	**[0 – 11]**
		ConsecutiveMistakes	**1.49**	**1.53**	**[0 – 7]**	**0.89**	**1.22**	**[0 – 6]**
	Last Session	TotalTime	1.30	1.44	[0.41 – 11.78]	0.97	0.71	[0.41 – 4.09]
		AverageTime	1.79	2.14	[0.57 – 18.13]	1.33	0.77	[0.55 – 5.58]
		LetterTaps	**4.57**	**4.15**	**[0 – 25]**	**5.82**	**3.79**	**[4 – 31]**
		WordTaps	**21.55**	**5.78**	**[16 – 46]**	**19.67**	**5.58**	**[16 – 48]**
		DragDrops	44.89	9.44	[34 – 81]	43.02	10.68	[36–103]
		Mistakes	**2.76**	**3.38**	**[0 – 15]**	**1.53**	**2.50**	**[0 – 13]**
		ConsecutiveMistakes	1.12	1.06	[0 – 5]	0.98	1.51	[0 – 9]

reduction. We reduce the datasets to two-dimensional data, and the resulting data are visualized to graphically analyze the existence of different patterns related to CD.

4.3.6 Clustering

We also apply clustering algorithms to detect groups of children (clusters) that show similar interaction patterns. Clustering algorithms are unsupervised machine learning techniques that find similar groups of instances from unlabeled datasets. We automatically gather the clusters of children from the datasets of our study (suppressing the `CD` variable). Then, we check whether any cluster is made up of children with or without CD. In such a case, the cluster represents an interaction pattern of children with or without CD. If so, we analyze the features of the cluster to document the interaction pattern associated with or without CD.

Most clustering algorithms are sensitive to outliers. Thus, once outliers are detected and documented (Section 4.3.1), we do not include them in the dataset passed to the clustering algorithm. We also perform a z-score normalization of the data before running the clustering algorithms, because that commonly improves the efficiency of the algorithms.

We ran the following clustering algorithms due to their good performance and the variety of approaches they represent (Berkhin, 2006): k-means, spectral clustering, DBSCAN (Density-Based Spatial Clustering of Applications with Noise), agglomerative clustering, and OPTICS (Ordering Points To Identify the Clustering Structure). For those algorithms that require us to specify the number of clusters, we executed them with 2 to 10 clusters. Out of the clusters found, we analyze those with at least 10 children where 90% of them either have (or do not have) CD.

5 Results and Discussion

Table 2 shows the arithmetic means, standard deviations, and maximum and minimum values of the features of the four datasets. We used the unpaired student's *t*-test and Wilcoxon statistical hypothesis tests to see if there are significant differences between children with and without CD. Rows in bold font in Table 2 indicate that, for that feature, the null hypothesis (means of two populations are equal) can be rejected (p-value<0.05).

For Match & Cards, the number of mistakes in the final session is the only feature that shows significant differences. Children with CD show more mistakes than children without CD in the last session. That could imply that, after practicing, children without CD are

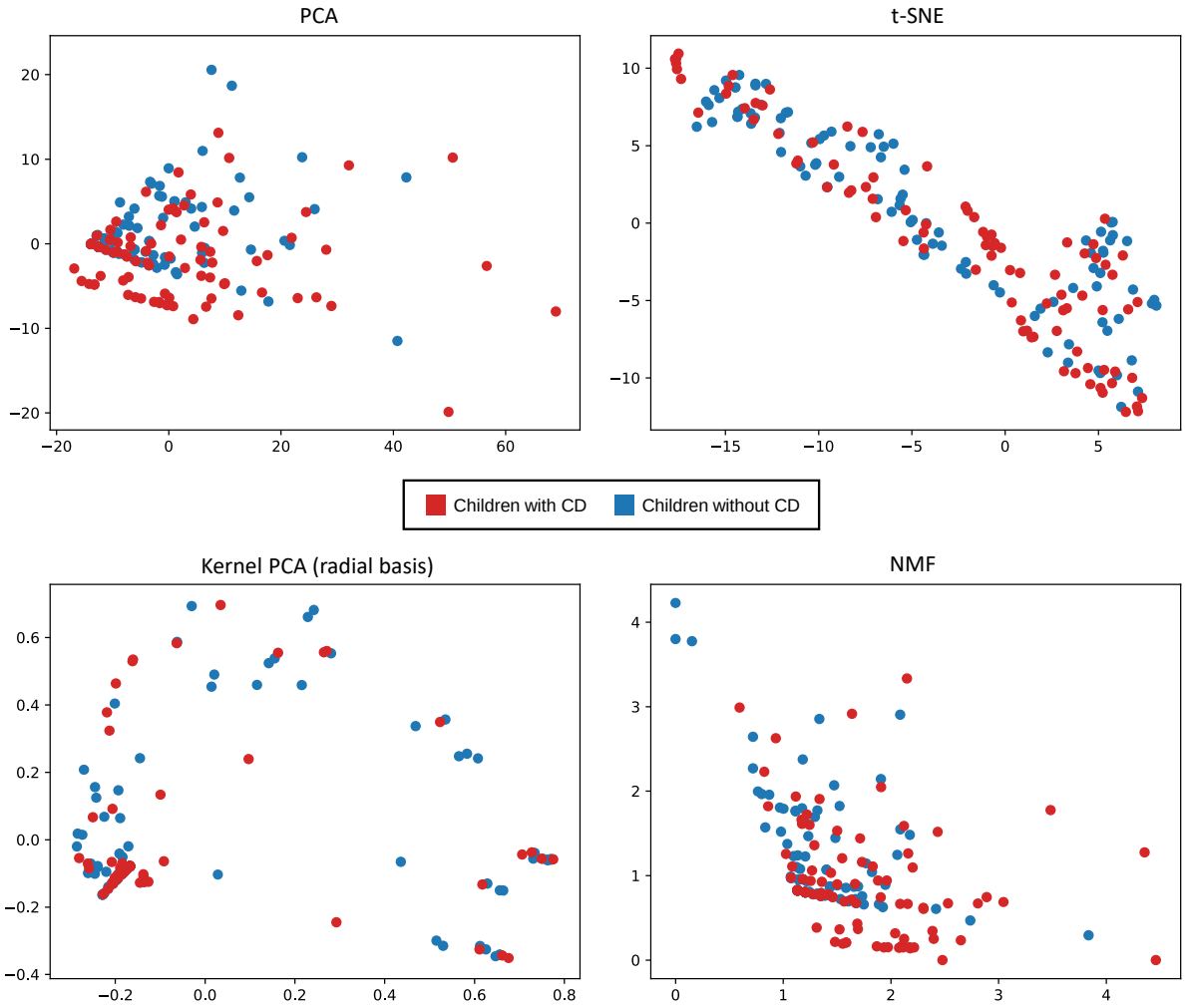


Figure 2. PCA, t-SNE, Kernel PCA (radial basis function), and NMF visualization of the first session of Cards & Sounds.

able to solve the activity with fewer mistakes than those with CD.

In Cards & Sounds, mistakes are significantly different in both sessions, and children with CD have significantly higher values. The same occurs for the consecutive mistakes, but only for the first session. Therefore, after practicing, children with CD no longer show higher consecutive mistakes than children without CD.

Letter and word taps also show significant differences for both the first and last sessions of Cards & Sounds. For word taps, children with CD require, on average, more taps than those without CD. It is the opposite for letter taps.

A feature with significantly different values indicates that the means of the two groups are not the same, but it does not necessarily mean that that feature can be used to successfully identify whether a child has been diagnosed with CD (our research question). Hence, we apply different data mining algorithms to see if those significant differences represent sufficient dissimilarity to find interaction patterns that identify children with or without CD.

5.1 Anomaly Detection

We used the outlier detection algorithms described in Section 4.3.1. For the Matching Cards dataset, three outliers were found. One child with CD that, for the first and last session, showed very high `TotalTime`, `Taps`, and `Mistakes`. Another one with CD presented, in the last session, very high values of `TotalTime`, `Taps`, and `DragDrops`. The third outlier had no CD, and also showed high `Taps`, `DragDrops`, and `Mistakes`. Therefore, regarding our research question, it cannot be stated that outliers with high values of any feature are associated with (no) CD.

For the Cards & Sounds application, we have similar results. Five outliers were found, three for the first session and two for the last one. All of them have very high values for at least two features. There are outliers with and without CD in both sessions, so we cannot conclude that atypical feature values are associated with (no) CD diagnosis (our research question).

Table 3. Association rules found (both obtained from the Cards & Sounds application).

Rule	Confidence	Support	Session
`LetterTaps` = *Very Low* ∧ `Mistakes` = *Average* ∧ `WordTaps` = *High* ⇒ `CD` = *True*	100%	5.4%	First
`Mistakes` = *Low* ∧ `ConsecutiveMistakes` = *Very Low* ⇒ `CD` = *True*	100%	5.4%	Last

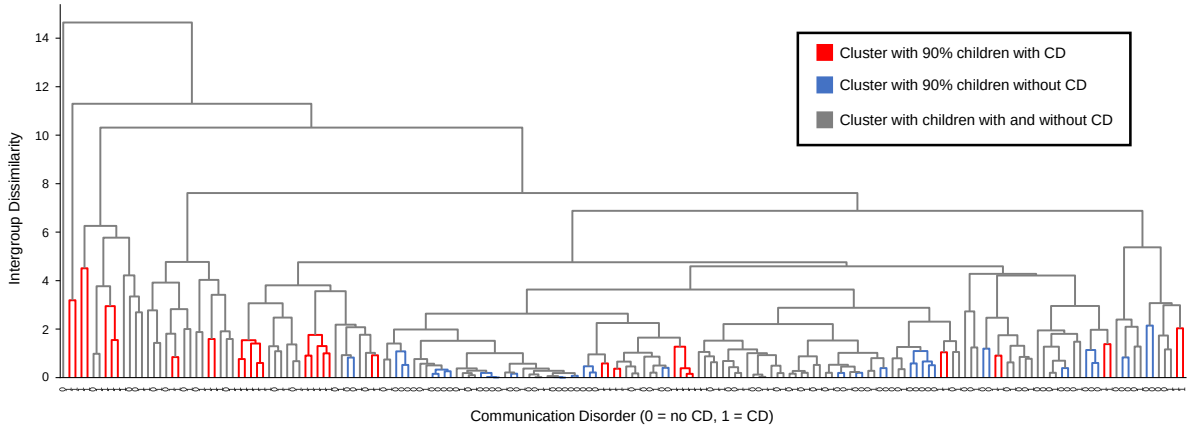


Figure 3. Dendrogram showing the agglomerative clustering for the last session of Cards & Sounds.

5.2 Decision Tree Classification

We created decision trees for the four datasets, with maximum depths from 2 to 10 (40 decision trees in total). Then, we analyzed the classification rules (tree branches) in those trees that at least classify 10% of the population with a minimum accuracy (correct classifications) of 90%. Only one single rule, retrieved from the Cards & Sounds trees (both sessions), met these criteria (one of such trees can be seen in Figure 1).

The classification rule found for both sessions was "`LetterTaps` ≤ 3 ⇒ `CD` = *True*". It has 100% accuracy and classifies 13.4% (first session) and 12.4% (last session) of all the children. Thus, the "`LetterTaps` ≤ 3" condition is only met by children with CD in the Cards & Sounds application, both in their first interaction and after at least nine sessions. Concerning our research question, this is the only interaction pattern extracted from the decision trees built.

5.3 Classification Rules Induction

The RIPPER*k* and IREP algorithms obtained 16 classification rules. The classification accuracies obtained ranged from 50.6% to 70.4% (average 60.9%). These classification accuracies are very low, considering that we have a binary classification problem (i.e., 50% accuracy is obtained by a random classifier).

The antecedent of each classification rule is expressed as disjunctions of conjunctions of propositional predicates ($(p_{11} \vee \ldots \vee p_{1n}) \wedge \ldots \wedge (p_{m1} \vee \ldots \vee p_{mn})$) , where the consequent is one value of the `CD` target variable. Therefore, each disjunction may be interpreted as a pattern or subrule aimed at classifying part of the population (e.g., $p_{11} \vee \ldots \vee p_{1n} \Rightarrow$ `CD` = *True*). We analyzed all the subrules, selecting those with a minimum accuracy of 90%.

The only rule meeting the 90% accuracy criterion was "`LetterTaps` ≤ 1 ⇒ `CD` = *True*", for the two sessions of Sounds & Cards (100% accuracy and 11.8% and 11.3% support). This rule is a subrule of the one found with decision trees (Section 5.2), so a new interaction pattern is not found.

5.4 Association Rules Learning

We ran the FP-Growth algorithm against our four datasets, obtaining 3,852 association rules. To gather the patterns asked in our research question, we then filtered those rules with a minimum accuracy of 90% and 5% support, having CD as their unique consequent.

Table 3 shows the two rules found, both for the Cards & Sounds application. Both define a pattern for children with CD, with 100% accuracy but representing only 5.4% of the children. As with the previous rules, the first rule in Table 3 keeps identifying a very low number of letter taps, combining it with average mistakes and a high number of word taps, but with much lower support (5.4%).

The second rule associates low and very low number of mistakes and consecutive mistakes to children with CD, after using the application in all the sessions. However, this result cannot be generalized to "children with CD have fewer mistakes" because of the following reasons:

a) There are no significant differences in the `ConsecutiveMistakes` variable (Table 2).
b) Significant differences exist for `Mistakes`, but the mean of children with CD is higher than those without CD.

c) The association rule just represents 5.4% of the children.

Therefore, the association rules extracted do not represent a sufficiently significant interaction pattern to identify children with or without CD.

5.5 Dimensionality Reduction

We reduced the dimensions of the four datasets to two, using the eight algorithms described in Section 4.3.5. The explained variances with two dimensions for the PCA algorithm ranged from 94.3% to 95.4%. This means that the distribution / visualization of the data in two dimensions only loses around 5% of the variability of data, so little information is dropped.

Figure 2 visualizes the first session of Cards & Sounds, using the PCA, t-SNE, Kernel PCA (radial basis as the kernel function), and NMF algorithms. In the representations, we cannot clearly identify a group of points (clusters), close among them, distant from the rest of the points, and with a substantial number of instances. Kernel PCA is probably the visualization that groups more instances (children) together (e.g., points around x=0.7 and y=-0-1). However, no cluster holds children only with or without CD (i.e., points with the same color). After analyzing all the plots, we did not identify a group of children (points) with or without CD (similar color) representing a similar interaction pattern.

5.6 Clustering

By running the five different clustering algorithms described in Section 4.3.6, we obtained 506 clusters of different sizes that represent distinct interaction patterns. Then, we considered those clusters whose instances account for at least 90% of children either with or without CD. All the clusters meeting that requirement hold at most 10 children, representing only 5.38% of the population. The only cluster with 10 children is the already discussed interaction pattern for the Cards & Sounds activity, where the value of `LetterTaps` is very low.

The disassociation between the clusters found and the CD diagnosis classification can be seen in Figure 3. That figure hierarchically displays agglomerative clusters from one single child (bottom of the figure), comprising them up to one single cluster with all the children (top of the figure). The height of clusters is proportional to the value of the intergroup dissimilarity. Red and blue colors indicate clusters with at least 90% of the children with and without CD, respectively. Gray color indicates mixed values. It can be seen how only clusters with at most five children have 90% of their instances with the same CD value. Responding to our research question, the clustering algorithms utilized found no representative interaction patterns to identify children with or without CD.

6 Conclusions

In a previous study, we showed how two tablet applications, which follow the Montessori method, can help children with and without CD in the process of learning the sounds and writing of letters and words. In this article, we study whether, using those two applications, it could be found interaction patterns capable of identifying children with and without communication disorders. The only significant pattern found, followed by 12.4% of the children in the Cards & Sounds activity, is that the children with the 12% lowest number of taps over letters (3 taps at most) have CD. This is fulfilled for both the first interaction with the application and the last one, after at least eight sessions of training. The pattern is also supported by the significant difference in the variable that measures the number of taps on letters, where children with CD show lower values than children without CD. The pattern may be representing the fact that some of the children with CD do not find it useful to hear the starting phoneme of one word to match it with the corresponding word, due to their CD. According to the DSM-5 classification of CDs, the interaction pattern found might be representing some of the children with receptive language disorders, since hearing the starting phonemes does not help in solving the activity. CDs include a wide range of diagnosis categories and the pattern found is only followed by 28.1% of the children with CD involved in the study.

All the datasets used in our study, the source code implemented to mine all the information, and the data mining models created are freely available for download at
https://reflection.uniovi.es/download/2022/com-dis/

Acknowledgments

This work has been partially funded by the Department of Science, Innovation and Universities: project RTI2018- 099235-B-I00. We have also received funds from the Principality of Asturias, through its Science, Technology and Innovation Plan (FC-GRUPIN-IDI/2018/000199) and University of Oviedo (GR-2011-0040). We thank the teachers and caregivers of the educational institutions for their support.

References

Berkhin, P. (2006). A Survey of Clustering Data Mining Techniques, In *Grouping Multidimensional Data*, Springer, Berlin, 25–71.

Bondy, A.S., & Frost, L.A. (1998). The picture exchange communication system. *Seminars in Speech and Language*, 19(4), 373–389.

Cabielles, D. (2021). *Mobile device interaction and intervention for users with autism spectrum disorder and communication disorder*. PhD thesis, Computer Science Department, University of Oviedo.

Cohen, W. W. (1995). Fast Effective Rule Induction. Machine Learning, In *Proceedings of the 12th International Conference on Machine Learning*, ICML'95, 115–123.

Collins, J. W. (2011). *The greenwood dictionary of education*. Greenwood, ISBN 978-0-313-37930-7.

Di Fuccio, R., Siano, G., & De Marco, A. (2017). TriPOD: a prototypal system for the recognition of capacitive widget on touchscreen addressed for Montessori-like educational applications. In *Recent Advances in Information Systems and Technologies*, 2, 664–676.

Drigas, A., & Gkeka, E. G. (2017). ICTs and Montessori for Learning Disabilities. *International Journal of Recent Contributions from Engineering, Science & IT*, 5(3), 77–84.

Douglas, N, Brush, J., & Bourgeois, M. (2018). Person-Centered, Skilled Services Using a Montessori Approach for Persons with Dementia. *Seminars in Speech and Language*, 39(03), 223–230.

DSM-5. (2014). *Diagnostic and Statistical Manual of Mental Disorders.* American Psychiatric Association Publishing, 5th edition.

Frutos, M., Bustos, I., Zapirain, B. G., & Zorrilla, A. M. (2011) Computer game to learn and enhance speech problems for children with autism. In *16th International Conference on Computer Games*, CGAMES'11, 209–216.

Fürnkranz, J., & Widmer, W. (1994). Incremental Reduced Error Pruning. In *Proceedings of the International Conference on Machine Learning*, ICML'94.

Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. Oxford, England: Morgan Kaufmann.

Herman, A. M., & Abraham, K. G. (2000). Occupational Outlook Handbook. *US Department of Labor, Bureau of Labor Statistics*, bulletin 2520.

Jones, M. R. (2016). *Implementing Technology in the Primary Montessori Classroom*. MSc thesis, St. Catherine University.

Liu, F.T., Ting, K.M., & Zhou, Z.H. (2008) Isolation forest. In *Eighth IEEE International Conference on Data Mining*, ICDM'08, 413–422.

McKenzie, S., Spence, A., & Nicholas, M. (2018). Going on Safari: The Design and Development of an Early Years Literacy iPad Application to Support Letter-Sound Learning. *Electronic Journal of e-Learning*, 16(1), 16–29.

Montessori, M. (2021) *The Montessori method*. Mint editions.

Paul R. (2016). *Language disorders from infancy through adolescence: assessment & intervention*. Elsevier Health Sciences.

Pérez-Pérez, J. R., Cabielles-Hernández, D., Sánchez-Santillán, M., & Paule-Ruiz, M. P. (2021). Interaction of children with and without communication disorders using Montessori activities for the tablet. *Personal and Ubiquitous Computing*, 25, 495–507.

Pickering, J. (2017). *Montessori for Children with Learning Differences*. Montessori Life.

Ruben, R. J. (2000). Redefining the Survival of the Fittest: Communication Disorders in the 21st Century. *The Laryngoscope*, 110(2), 241–245.

Smith, J. (2012). iLit: Using iPads for Multi-sensory Literacy Development Intervention. In *Abydos Teachers' and Trainers' Conference*.

Toki, E. I., & Pange, J. (2006). A Comparative Study of Two Learning Methods: Collaborative Learning versus Nearest Neighbor Learning. In *E-Proceedings of the 4th International Conference on Education and Information Systems, Technologies and Applications*, EIST'06.

Toki, E. I., & Pange, J. (2010) E-learning activities for articulation in speech language therapy and learning for preschool children. *Procedia Social and Behavioral Sciences*, 2, 4274–4278.

Wetherby, A., & Prizant, B. (2002). *Communication and Symbolic Behavior Scales Developmental Profile–First Normed Edition*. Brookes Publishing Co.

Wetherby, A. M., Goldstein, H., Cleary, J., Allen, L., & Kublin, K. (2003). Early Identification of Children with Communication Disorders. Concurrent and Predictive Validity of the CSBS Developmental Profile. *Infants & Young Children*, 16(2), 161–174.

Design of All-In-One Technology-Enhanced Learning Software for Supporting Teachers' Personal Activities

Stefan Svetsky, Oliver Moravcik
Slovak University of Technology in Bratislava
Vazovova 5, Bratislava, Slovakia
stefan.svetsky, oliver.moravcik@stuba.sk

Abstract. *Computers cannot support all the activities performed by a teacher. Due to the lack of appropriate Information Technology (IT) tools, a teacher must adapt to existing technologies. The educational software WPad, developed within the framework of the authors' research on technology-enhanced learning (TEL), offers a universal solution. WPad controls the so-called virtual knowledge into which the teacher inserts content to simulate teaching activities. In this paper, TEL design in research and for publishing support are described based on examples from university teaching. The methodology of synchronising teaching and computing activities is also explained using the idea of building a personal information system.*

Keywords. Technology-enhanced learning, educational software, IT tools, synchronisation of pedagogical and informatics activities

1 Introduction

In practice, every university teacher participates in a large number of educational processes, which include not only teaching and learning, but also research, publishing, self-study, communication, international collaboration, e-learning, educational management, promotional recruitment, and various related administrative and economic activities. Thus, the integration of IT into education practically means that some software and IT infrastructure must be available for all these processes.

As far as IT support for education is concerned, learner-centred research is most often emphasised in the academic literature. One rarely encounters a focus on personal support for the teacher, even though the teacher is a key player in education. Even from the point of view of common sense, one can ask whether there can be any learner-centred approach without the involvement of the teacher. The requirement of IT support for the teacher as a key person in education is captured well in the TEL monograph by Goodman (2002). In the context of American universities, it is emphasised that every professor should have a team available to provide IT support. Although this monograph was published twenty years ago, it is still relevant to TEL. In this context, in a discussion on ResearchGate, it was said that while this works at US universities, in the EU, IT support is still left to the professors.

The integration of IT into education is generally covered mainly by research areas such as TEL, educational technology, and learning technology, which, however, are only synonymous from the teacher's point of view. Various studies (Balacheff, 2009; Roblyer, 2013; Bower, 2017), book series (Huang et al., 2021), and literature reviews (Walker et al., 2016) list a large number of IT tools. However, they are rather global and general solutions to which the teacher must adapt; the teacher must determine the suitability of each of these tools for his teaching activities. Examples of these tools are learning management systems (Moodle, Blackboard, academic information systems), Office packages, and tools for specific areas such as learning analytics.

It is necessary to realise that computer technology cannot support all the educational processes performed by the teacher. Therefore, it is logical that in the literature we find the argument that software engineering has little focus on educational software development (Martens, 2014), criticism of IT solutions or research approaches to education (Kinchin, 2012; Lundie, 2016; Kirkwood and Price, 2021), and a lack of theoretical approaches (Oliver, 2013). The importance and use of educational technology in the classroom has been discussed, including the question of 'whether schools and teachers themselves are ready for the use of technology in education and whether they are aware of its benefits' (Stošić, 2015).

Compared to the current situation, the authors' long-term research has brought practical solutions using computers to the classroom and online environments based on the development of their own all-in-one TEL software, which implements the paradigm of the mass processing of educational knowledge on a teacher's personal IT infrastructure. We did not find a similar solution in the literature.

The authors began researching TEL from 2006–2008, when they realised that teachers, researchers, students, and knowledge workers need to be equipped with personal all-in-one software. Users basically should not notice any differences between working on their own computer or online. This vision was presented as part of an e-learning solution for undergraduates at the ICETA conference (Svetsky, 2007). At the time, the authors assumed that as computer technology gradually improved, there would be such universal software on the market. To date, however, this has not happened, and, on the contrary, the authors have shown in recent years that they eventually programmed such a solution themselves.

Following this, the first author of this paper, as a teacher, wrote the source code for all the activities he performed, e.g., the creation of materials and tutorials for the teaching of undergraduates, the automation of multilingual searching, self-study, publishing support, and even grading and assessment, including linking these activities to the university's academic information system (AIS). For these purposes, he used a specific database structure in the form of tables. He processed the content of this structure by writing source codes from which he gradually built the WPad database application. Students had WPad installed on their computers in the classroom and used it in several study programs, e.g., the background of environmental protection, industrial management projects, occupational safety and health, programming languages, and the basics of chemistry. At the same time, several pre-service teachers tested WPad as part of their diploma theses concerning the production of e-learning kits.

Thus, the WPad software acted as a teacher's personal teaching tool and multifunctionally supported several teaching and learning activities. At the same time, WPad has begun to be used in research under the 7th Framework Program for the area of technology-enhanced learning. For this purpose, an advanced BIKE application was used; it ran only on the designer's computer. Its use in TEL research was first published as part of the IGIP-SEFI conference, when WPad template tables were created within the framework of international cooperation to support publishing (Divjak et al., 2010).

At this time, the authors investigated how it is possible for WPad to function as an all-in-one teaching tool. This made it clear that the WPad data structure can simulate educational knowledge if content and related meta-information are inserted into it. Thus, it can function as a virtual knowledge macro-substitution and can be transferred between offline and online environments. The novelty of this approach was confirmed by the Slovak Patent Office (utility model Nr. 7340, 2016).

On this basis, the paradigm of the batch processing of educational knowledge was further developed, TEL software was designed, and a combined offline/online IT infrastructure was developed; within this infrastructure, virtual knowledge is transferred in the form of tables (as computer files). It therefore provides added value compared to the transfer of common computer files, because only reduced educational content is transferred. This significantly reduces information overload and entropy. The authors have continuously presented dozens of application outputs their long-term research on the integration of IT into education and academic research at scientific conferences and journals (e.g., Svetsky and Moravcik, 2014, 2016, 2017, 2019, 2021).

In the following sections, a demonstration of the research approach and examples of the use of the WPad software and tables in academic teaching and research will be described. The area of implementation of learning analytics, which was tested directly during the teaching of undergraduates, will also be mentioned. In a separate section, the principle of the synchronisation of pedagogical (didactic) and informatics approaches to IT support for publishing, which is a universal academic educational activity, is discussed. Such synchronisation requires the collaboration of the educational algorithms and IT algorithms, including adaptation to the operating system and online learning spaces and clouds.

2 TEL Design in Research

In TEL design, it is not enough to solve only content and pedagogical algorithms. Apart from the fact that the output must 'fit' on the computer screen, a separate problem is that computer files with educational content are in dozens of different and incompatible formats (PDF, HTML, DOC, TXT, PNG, MP3, etc.). This also applies to software and hardware and the entire offline and online infrastructure. A comprehensive approach to TEL research therefore requires solutions for various educational and IT categories. In addition, the teacher must deal with the management of the large number of computer files that teachers tend to collect over time.

2.1 Research Approach and Methodology for Teaching and Learning

In programming terms, this means that software adaptation to the operating system, networks, and the web must be addressed as a separate block. Therefore, a separate category of user menus must be programmed. Since educational content has to be selected and tailored for specific teacher and student activities, practically this means that it also has to be transferred between personal computer folders and online environments (virtual learning spaces, clouds, the public web).

The authors' long-term research has shown that it is not technically possible to cover such a huge number of variations of teacher and student activities and sub-activities, because if these activities are not defined in advance by algorithms, the computer does not know how to deal with them. It is important to note that current artificial intelligence can only process data that is hierarchically an order of magnitude lower than the information and knowledge that is formed from this data. Therefore, the challenge is to program an infinite number of TEL applications. In other words, first, a pedagogical algorithm is designed, source codes are written for its steps, and adaptive algorithms have to be programmed for software and hardware, i.e., a certain IT infrastructure, for the sake of compatibility.

The biggest challenge for human knowledge processing, however, is the synchronisation of the pedagogical and informatics steps. However, explaining this using the example of TEL design for teaching is practically impossible, because the reader would have to be familiar with the subject matter being taught. To explain how synchronisation should work, it is more appropriate to approach the design of IT support for publishing activities. This is because this is a general academic activity that involves teachers, researchers, and students. A practical solution to the TEL synchronisation of IT support for writing a scientific paper is described in a separate section.

2.1.1 Personal Learning Analytics

The authors were in the consortium of the 7th Framework Program L3Pulse, which submitted a project in response to a call for learning analytics (LA). As part of this, they tested a personalised solution to find out how to create a practical application to support the needs of a teacher teaching undergraduates. The result of the LA literature survey was that its essence is to analyse the visiting logs of online educational environments to learn how students use these environments. On the basis of calculating statistics for various online activities, educational institutions can then improve their learning strategies.

Such LA statistics, however, have no added value for teachers, because for assessment and evaluation they need to monitor how students receive and work with educational content. In this context, the writing of tacit knowledge in the WPad educational tables by students during lectures or exercises proved to be useful. Because the WPad tables of all students can be combined, it was possible to calculate statistics related to their human knowledge and not only to internet logs.

In addition, the statistics and visiting categories of the faculty educational environment were monitored in parallel by an external service during the testing of LA. Thanks to this, the teacher had an overview that included how many students used the connection to the faculty educational server during and outside of class. At the same time, he knew which WPad tables were being used. This feature was used when students were doing a WPad internet search. Since the Opera browser was in use at the time, it was possible to monitor on each computer which pages the students visited during the lesson and to evaluate the effectiveness of teaching.

In this way, the methodology for using WPad for LA was developed. Its added value was that the LA statistics obtained were directly related to the students' knowledge in the classroom; it is not possible to obtain such statistics from online clicks (compare with Martin and Ndoye, 2016). From a practical point of view, the teacher gained a useful tool for formative assessment and, thanks to the feedback obtained, he was able to better manage his teaching.

2.1.2 Multi-search (online/offline)

WPad has a rich user menu, which also includes internet search items. Since the keywords to be searched by a search engine are entered directly into the text window of a WPad table, the user does not have to open a web browser and enter the keyword. Instead, the user just chooses the menu item of interest, and then the browser opens a window with the search results. There is a specific option in WPad, which conventional software does not have, allowing the user to type several keywords in a text window; closing the text window will cause WPad to open a browser window with search results for each keyword. This multi-search is automated, so that if 50 keywords were typed or pasted at once, the Opera browser would open 50 windows that would cover the entire computer screen.

The literature does not discuss the fact that, after a few years, a huge number of computer files accumulates on teachers' computers. In other words, the teacher also needs to deal with offline search on their own computer, CD-ROMs, DVDs, and USB drives. WPad has some menu items for these purposes. A teacher who uses virtual environments or even virtual machines on clouds needs to search for files in these environments as well (in other cases, file managers are used for this purpose).

2.1.3 Big Educational Knowledge vs Big Data

The purpose of this paper is to point out that a regular teacher has a 'small internet' on his computer. Thus, he needs a way to process such 'big educational knowledge'. From the individual point of view, the teacher has a huge number of files in which learning content is stored on his home and work computers or in online clouds.

Table 1 shows the number of files that the teacher had on disks C and D of his desktop computer in 2017 and on his HP notebook in 2022. The table shows that the total number of computer files is on the order of half a million to one million. Of course, this number also includes all the files in the Windows and

software folders. He personally 'collected' tens of thousands of files in various formats (TXT, DOC, PDF, HTML). There are also about 9000 PNG files, i.e., image files, in the table. In this case, these image files are almost exclusively screenshots that were taken for visualisation modelling and literature searches. This has led to the finding that creating screenshots is easier than making manual notes; in this case, mass image processing is performed.

It should be emphasised that the teacher works on his computer not only with a large number of files but also with folders. From his point of view, this data represents big educational content. Artificial intelligence is currently popular; it can process big data, which basically consists of short pieces of texts or numbers, but it cannot work with the range of content included in big educational content, i.e., it cannot work with knowledge, which is hierarchically higher than data. WPad also allows such 'big educational knowledge' to be processed multifunctionally at the level of the individual, and thus it acts as a container for the individual's information and knowledge.

Table 1. Number of files on the teacher's computer

	C (2017)	D (2017)	HP (2022)
Files	448891	656744	532183
TXT	5831	6410	8569
PDF	8693	13978	14812
HTML	42285	75983	25590
DOC	13568	19059	33141
PNG	-	-	9088

Although both Windows Explorer and file managers (e.g., Total Commander) allow one to work with many files and search for files in an offline mode, they only partially address the teacher's needs. This is due to the fact that these programs do not know exactly what the teacher needs. If a teacher wants to look up something he did years ago, he does not know in which files it is stored. Using WPad enables a teacher to automatically insert a list of files from folders or disks (including USB drives and CD-ROMs) into the WPad table and to filter, search through them, or open them directly as needed. In other words, WPad has several menu items that provide tailored activities related to files and folders similar to the services that file managers provide.

For example, the number of rows in a WPad table could be exactly equal to the number of files in the inserted list, as illustrated in Table 1. For example, on the HP notebook, the table had 532183 rows, and after filtering it for PNG files, it had 9088 rows. The important result of this test is that through purposeful filtering, the teacher could find what he was looking for. Fig. 1 illustrates such a search for PDF files with papers from the British Journal of Education Technology. Importantly, the teacher can also open the files directly from the table without a need to switch to Windows Explorer or Total Commander. In practice, this method was developed as part of the V4+ACARDC project: teachers from the V4 countries and Ukraine modelled the automated creation of educational packages (this is not the subject of this paper).

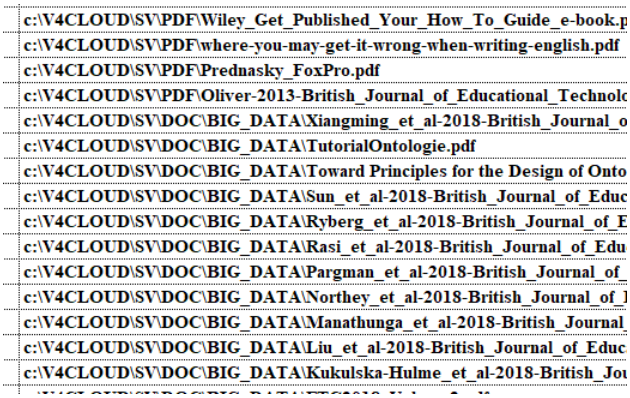


Figure 1. Results of searching for PDF papers in the table with 532183 rows

2.1.4 Personal TEL Infrastructure for Teachers

Fig. 2 illustrates the personal TEL infrastructure for teachers that the authors used in their research. It enables the teacher to combine offline and online activities. The basic TEL level requires only Windows and the installation of the WPad software on the computers of the teacher and students. This level is very user friendly because it does not require the users to have any higher IT skills. The advanced TEL level additionally includes many TEL online solutions, such as the ability to access the web, clouds, virtual learning spaces, virtual machines, academic information systems of universities, and PIKS channels (this requires routine IT skills).

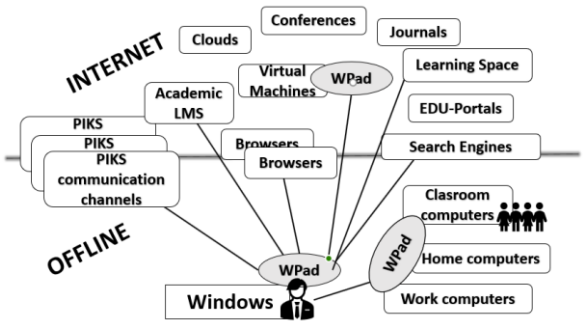


Figure 2. Personal TEL infrastructure for teachers

2.2 TEL Design for Publishing Support

The authors' most recent TEL research focuses on the synchronisation of learning and informatics activities. In practice, this means that the final algorithm that combines the sub-activities of the learning activities is

written simultaneously with the WPad source codes. It is important that the user identifies these learning activities and defines the sequence of steps. As a result, the WPad application menu is extended, and potential research topics are profiled. In this context, it should also be understood that the teacher needs to process knowledge as quickly and easily as possible. As a result, the teacher adopts IT support as a routine activity to automate his teaching processes. Interestingly, the issue of the speed of knowledge processing is not mentioned in the educational literature.

Explaining the synchronisation of pedagogical and IT steps to the ordinary reader is quite difficult, but it is also difficult to explain to IT specialists. The problem is that IT support is related to dozens of subjects and social or technical areas, and each TEL solution is different. Therefore, it seems to be an optimal starting point to describe the principle of synchronisation using the example of applying TEL solutions to writing scientific papers, because this is a general academic activity carried out by teachers, students, and researchers. A teacher who publishes regularly or more frequently will find, over time, that he has certain resources available and that certain activities in publishing are repetitive. The process of writing a scholarly paper can be described as follows:

- choosing the right conference or conferences;
- reviewing the requirements for authors;
- submitting an abstract or extended abstract;
- submitting a paper for review;
- having a native speaker of the language the paper is written in proofread the paper;
- editing the paper according to the reviewers' requirements;
- submitting the final version (camera ready) of the paper;
- registration and payment;
- creating a video presentation for a virtual conference;
- communicating with the organizers throughout the process via emails, automated emails, and online paper submission systems.

2.2.1 Literature Resources

The most time-consuming phase of writing a scientific paper is writing the introduction, which allows readers and reviewers to classify the presented topic into some well-known category. This is due to the fact that the field of IT or computer science is very atomised in the literature. If the research topic is interdisciplinary, it is almost impossible to find similar works or research approaches. This is the case for the topic of this paper. By using WPad, a teacher does not have to use a dozen pieces of software for the same purpose; however, IT journals specialise in the individual software products.

Therefore, the key question is the following: what should the authors' TEL design be compared with, if there is not any software on the market that covers several teaching and learning areas? In this respect, WPad tables can be applied to all of the following categories of resources, which are used by practically every publishing teacher (a teacher can create their own information system for this purpose):

- educational internet and e-learning portals;
- university information systems for publishing;
- a teacher's folders and files on his own computers;
- archived papers on clouds used by teachers;
- online repositories with scientific literature;
- the teacher's own knowledge base (e.g., WPad tables);
- internet search using search engines;
- many other e-media sources (USB drives, CD-ROMs, DVDs);
- archived e-mail communication.

Integrating IT into teaching is also related to teacher support, so part of the TEL design needs to address this support as well. WPad can be used for this purpose. From the WPad tables, the user can switch between the above-mentioned online and offline resources, perform a simple search or multi-search, and save items to the tables. Naturally, for each paper for a conference or journal, the teacher can create separate tables, which can also hold various associated information and data from communication with organizers and instructions for authors.

2.2.2 Paper Writing

After selecting a conference or journal, the teacher must study the instructions for authors and perform the above-mentioned steps, from choosing the conference to creating the final presentation (be aware that each piece of information can be contained in one line of the WPad table). Here, it is important to download the required formatting template. If the research team members publish in the pedagogical, IT and technical areas, they will immediately find that conferences use their own templates and require various reference styles (IEEE, APA, Harvard, ISO 690). This is seemingly not a problem, but for authors who have at least 15–30 references, rewriting them in another style is very time consuming.

In terms of synchronisation, the publishing activities represent a system of algorithms for which it is possible to write source codes. In this context, during the development of WPad, the teacher also began to logically create WPad tables to support the writing of scientific papers. In practice, it was useful to use a separate table for each article; this table served as a manuscript with plain text. Its content was finally transferred to a Word text editor, where it was edited and formatted as necessary. In the same table,

there were also outputs from the literature search, annotations and associated information, and links to the above-mentioned resources. This could be considered one example of a synchronisation solution. The teacher performs some partial steps related to publishing, and at the same time, the needed source codes are being written.

The principle of synchronisation is based on the authors' basic focus on the automation of knowledge-based processes (where teaching and learning come together). This means that the slowest stage of the process is addressed. The use of the function key F6 was programmed for this reason. If an author has 30–50 rows (records) in the WPad table, he must click between them and needs to filter them thematically. However, this will slow down the writing of the paper, even though there are several items in the user menu for this. Therefore, for this activity (similarly, it can be a learning activity), source code for using the F6 key was written. The author can look at a specific record that includes the word he is searching for, and after the author presses F6, only the rows in the table containing that word are shown. For example, if the word CECIIS is in a given row, and it is in four rows of a table, then pressing F6 will show only those four rows of the table (so that the user does not have to go to the user menu).

Using the F6 key is especially useful if the table's content was created automatically and contains thousands to hundreds of thousands of rows. For example, the university's publication system makes it possible to generate a list of teachers' publications in the form of an RTF file. After copying and pasting the content of the RTF file into WPad, a table is automatically created; it contains each item in the list in a separate row. Now, the table can be searched quickly by pressing F6.

To illustrate this, the author included in a table a list of all the papers published by the faculty at his university from 2005–2021 and searched this table in order to evaluate the quality of the publications of the faculty. As the table had tens of thousands of rows, he selected a row that included his name. After pressing the F6 key, only his publications for the period were shown, as Fig. 3 illustrates.

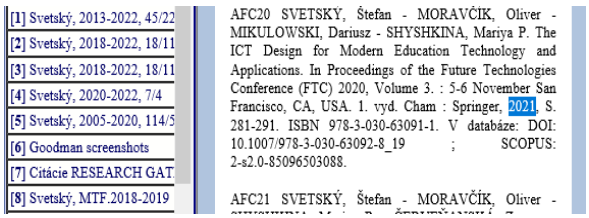


Figure 3. Search results in a WPad table

It should be emphasised that each WPad table can be converted into an HTML table by simply pressing the keys CTRL-F1. The screenshot shown in Fig. 3 illustrates the HTML output of the database of the university publishing system. This is a huge advantage for the teacher or even the students, as they can produce personal or working HTML tables for various educational purposes (e-learning, e-scripts, any content for distance learning, support face-to-face teaching with computers in the classroom, internet search). Users work with WPad tables and/or HTML tables as needed, which are simply copied to the online environment of servers and the web. Moreover, unlike other software, ordinary computer skills are sufficient to work with the tables. Thus, in this case, the technology is adapted directly to the educational needs of the teacher and students and enhances learning.

3 Conclusion

It is clear from the description of writing scientific papers that there are several activities involved. The principle of synchronisation is precisely that the source code is written while the activity is being performed. The most important step is to always address the activities that are part of the slowest stage of the process. This is mainly the writing of the introductory part of the paper, which requires a literature online/offline search. To speed up this process, it is possible to program a simple internet search or multi-search and to create tables with the author's own dictionaries (or containing sentences with frequently used phrases). There is again a large number of other possible procedures. For example, one can choose to use more search engines, such as Bing, Google, IxQuick, DuckDuckGo, or Brave, or perform searches of German, French, and domestic sources.

In principle, one could design a separate software application for each group of activities, but WPad is programmed to be a multifunctional, all-in-one, personal TEL software for teachers. Thus, programming multiple menu items in this way allows the teacher to perform a wide range of IT support tasks. In this context, the programming of a personal information system that can support the publishing activities of individuals will be designed in future research.

Acknowledgments

The research was co-financed by the governments of Czechia, Hungary, Poland, and Slovakia through Visegrad Grants from the International Visegrad Fund. The mission of the fund is to advance ideas for sustainable regional cooperation in Central Europe. The-term research is in general funded by the academic research funds of the Slovak University of Technology in Bratislava (Faculty of Materials Science and Technology in Trnava).

References

Balacheff, N., Ludvigsen, S., Jong, T, Lazonder, A., Barnes, S. (eds.) (2009). *Technology-Enhanced Learning. Principles and Products.* Springer, XXVI, 326 p.

Bower, M. (2017). *Design of Technology-Enhanced Learning: Integrating Research and Practice.* Emerald Group Publishing - Education.

Divjak, B., Begičevič, N., Moravčík, O., Svetský, Š. (2010). Bilateral collaboration for ICT support in engineering education. In *Proceedings of the Joint International IGIP-SEFI: Annual Conference 2010. Diversity Unifies - Diversity in Engineering Education.* Trnava, Slovakia.

Goodman, S.P. (ed.) (2002). *Technology-Enhanced Learning: Opportunities for Change*. Laurence Erlbaum Associates, Mahwah, NJ, USA.

Huang, R., Kinshuk, Jemni, M., Chen, N.-S., Spector, J.M. (Eds.) (2021). Lecture Notes in Educational Technology Series. Available at: https://www.springer.com/series/11777.

Kirkwood, A., Price, L. (2021). Technology-enhanced learning and teaching in higher education: what is 'enhanced' and how do we know? A critical literature review. *Learning, Media and Technology*, 39 (1).

Lundie, D. (2016). Authority, autonomy and automation: The irreducibility of pedagogy to information transactions. *Studies in Philosophy and Education*, 35(3), 279–291.

Martens, A. (2014). Software engineering and modelling in TEL. In: Huang, R., Kinshuk, N.-S.C. (eds.) *The New Development of Technology Enhanced Learning: Concept, Research and Best Practices*, pp. 27–40. Springer.

Oliver, M. (2013). Learning technology: Theorising the tools we study. *British Journal of Educational Technology*, 44, 31–43.

Roblyer, M. D., Doering, A. H. (2013). *Integrating Educational Technology into Teaching*, 6th Edition. Pearson.

Shukor, N. A., Tasir, Z., Van der Meijden, H., Harun, J. (2014). A predictive model to evaluate students' cognitive engagement in online learning. *Procedia – Social and Behavioral Sciences*, 116, 4844–4853.

Stošić, L. (2015). The importance of educational technology in teaching. *International Journal of Cognitive Research in Science, Engineering and Education* (IJCRSEE), 3(1),

Svetsky, S. (2007). Modelling of multilingual e-learning and virtual learning space for R&D staff. In: *ICETA 2007: 5th Int. Conference on Emerging e-Learning Technologies and Applications*. Košice: Elfa, pp. 1–4.

Walker, R., Voce, J., Swift, E., Ahmed, J., Jenkins, M., Vincent, P. (2016). 2016 survey of technology enhanced learning for higher education in the UK. *UCISA TEL Survey Report 2016*. University of Oxford.

Martin, F., Ndoye, A. (2016). Using Learning Analytics to Assess Student Learning in Online Courses. Journal of University Teaching & Learning Practice. 13 (3).

Svetsky, S., Moravcik, O. (2014). The automation of teaching processes based on knowledge processing. Transactions on Machine Learning and Artificial intelligence, 2(5).

Svetsky, S., Moravcik, O. (2016). The utility model Nr. 7340. Implementation of an unstructured data processing system using a specific data structure. Industrial Property Office of the Slovak Republic – WebRegisters.

Available at: https://wbr.indprop.gov.sk/WebRegistre/UzitkovyVzor/Detail/45-2014.

Svetsky, S., Moravcik, O. (2017). The Universal Personalized Approach for Human Knowledge Processing. In *Proceedings of the WEC-ICCSE 2017*, London: GB, pp. 247-251.

Svetsky S., Moravcik O. (2019) Some barriers regarding the sustainability of digital technology for long-term teaching. In: Arai, K., Bhatia, R., Kapoor, S. (eds), *Proceedings of the Future Technologies Conference. FTC, 2018. Advances in Intelligent Systems and Computing*, vol. 880. Springer, Cham.

Svetsky, S., Moravcik, O, Shyshkina, M., Cervenanska, Z., Kotianova, J. The knowledge-based design of educational technology. (2021). In *K. Arai (ed.). Proceedings of the Future Technologies Conference (FTC) 2021*. Volume 3. FTC 2021. Lecture Notes in Networks and Systems, vol 360. Springer, Cham.

Svetsky, S., Moravcik, O. (2021). Human-centered software design for knowledge-based processes. *International Conference on Electrical, Computer and Energy Technologies*, pp. 1–9.

Enhancing Self-regulation Strategies through Peer Feedback in Online Higher Education

Lourdes Guàrdia, Maite Fernandez-Ferrer, Nati Cabrera, Federica Mancini, Marcelo Maina
Psychology and Educational Sciences
Universitat Oberta de Catalunya
Rambla del Poble Nou, 156, 08018, Barcelona
{lguardia, mfernandezferrer, ncabrera, fmancini0, mmaina}@uoc.edu

Abstract. *Specific research addresses the adoption of self-regulation strategies through peer feedback and the use of chatbots for supporting Higher Education institutions to transform online assessment. The purpose of the study presented below was to evaluate the design of a learning activity containing a peer-feedback intervention in order to suggest improvements. A set of data collection instruments was designed and distributed after the first iteration of the experience to gain insights into the students' perception of the peer-review task, designed to enhance their self-regulation capacity and supported by the use of a chatbot throughout the learning activity. The data collected allowed us to identify the learning activity design aspects that should be improved for future iterations to enhance student process performance skills.*

Keywords. Peer feedback, Chatbot, Self-regulated learning, e-Assessment, e-Feedback

1 Introduction

Under the umbrella of the eFeedskills project[1], this paper presents a research experience that took place in the Faculty of Psychology and Education of the Open University of Catalonia (UOC), in the online master's degree in Education and ICT (Information and Communication Technology) and e-learning. The experience consists of an e-learning activity implementation that includes a peer-review task within the course 'Fundamentals of e-learning techno-pedagogical design', which aims to train learning design professionals for technology-mediated learning environments.

The learning activity studied in this research asked students to complete an academic essay writing exercise. The essay format was chosen to encourage students to organise their thinking through a creative writing process, requiring the cognitive effort of describing their ideas and opinions, defending them by citing the sources used and consulted throughout the course, and substantiating them appropriately and adequately (Dysthe, 2015; Klimova, 2012).

In the middle of the essay learning activity, a peer-review task was proposed. This specific task aims to promote student reflection and critical thinking and enhance self-regulatory capacity. Through feedback from their colleagues on their unfinished essays, students are better able to achieve the objectives of the learning activity (Carless & Winstone, 2020; Nicol & Macfarlane-Dick, 2006). In this sense, the feedback provided by peers promotes reflection on what is written, improves progress in academic essay writing and prevents student procrastination (Cerezo et al., 2017).

2 State of the art

The literature recognises the benefit of peer feedback in the development of self-management, evaluative judgement, social skills and self-regulated learning (Nicol, 2009). At the same time, the progressive development of the ability to evaluate the quality and impact of the work produced by others enables students to become more independent and less reliant on the teacher (Tsui & Ng, 2000).

The integration of peer feedback can help to improve the quality of learning as it requires students, firstly, to comprehend the assessment criteria provided by the rubric and learn how it should be applied, and secondly, to read and compare, or question ideas, suggest modifications and even reflect by comparing their learning outcomes with those of others (Liu et al., 2001). In addition, by providing feedback on the work of their peers, students participate in each other's learning and thus achieve greater understanding and appreciation for their peers' experiences and perspectives (Corgan et al., 2004). Existing research also shows a strong focus on providing peer feedback

[1] eFeedskills, https://www.ub.edu/digital-feedback/es/inicio/

as separate from grading because it is so relevant to the formative process (Liu & Carless, 2006; Nicol & Macfarlane-Dick, 2006; Nicol, 2013).

In recent years, the process of providing peer feedback has been increasingly facilitated online. Some authors argue that it may play an even more important role in online learning (Lynch, 2002; Palloff & Pratt, 2001) than in traditional face-to-face learning. A growing emphasis is also placed on the design of online activities providing students with the opportunity to not only reflect on the actual work and performance of their peers (peer feedback), but also to provide them with directions towards achieving the desired goal (Latifi et al., 2021). This type of constructive feedback, known as feedforward, is closely linked to learning activities leading students to implement the suggestions for improvement in further actions (Hattie & Timperley, 2007), and thus feed into the next assignment (Hounsell et al. 2008). Additionally, it is relevant to consider that students' time management impacts on academic performance and can have negative effects on academic achievement (Cerezo et al., 2017), and fostering students' self-regulation capacity could support them in reducing academic procrastination.

In parallel, the development and application of educational chatbots, providing learning content and feedback in an interactive manner (Jia, 2009; Lee et al., 2009), has increasingly gained popularity due to their potential for supporting on-demand interaction with students through dialogue. Scholars have also found that interacting with chatbots generally increases students' learning interests (Johnson, 2001). However, despite the emphasis on the exploration of the possibilities of chatbot technologies for education, their use in the self-regulation of learning is still scarce (Fernández-Ferrer et al., 2021).

Against this scenario, we advocated for fostering self-regulated learning and improving the final outcomes of a complex task (academic essay writing exercise) through peer feedback. The application of a Design-Based Research (DBR) methodological approach (Reeves, 2000) for the design, testing and refinement of the intervention targeted to this purpose aims to overcome some of the shortcomings of traditional research methods in exploring the function of tools and approaches in an educational environment (Amiel & Reeves, 2008). As a result, the DBR iterative process provided pointers for improving the learning activity design, as well as evidence for the elaboration of design principles regarding self-regulation support in academic writing activities.

3 Research context and questions

The learning activity consisting of an academic essay writing exercise comprised several tasks which are detailed below (Figure 1). The self-regulation strategy was supported by a chatbot that played the role of a monitoring and supporting system during the whole process from the forethought to the self-reflection phase.

The chatbot was structured as a decision tree with an initial question (the root of the tree) aimed at identifying the self-regulation phase the student was in before moving the conversation forward. The subsequent closed questions were then calibrated to the corresponding self-regulation phase accompanied by a set of predefined options that the students were asked to select. Through the indicated option, the chatbot was able to guide the students through the entire process. The questions were formulated using student-friendly language.

The chatbot was integrated into the activity sequence presented in Moodle through a direct link to the external application. The instructions of the activity included an explanation of the role of the tool for the development of the learning-to-learn competence, directly related to the students' self-regulation capacity.

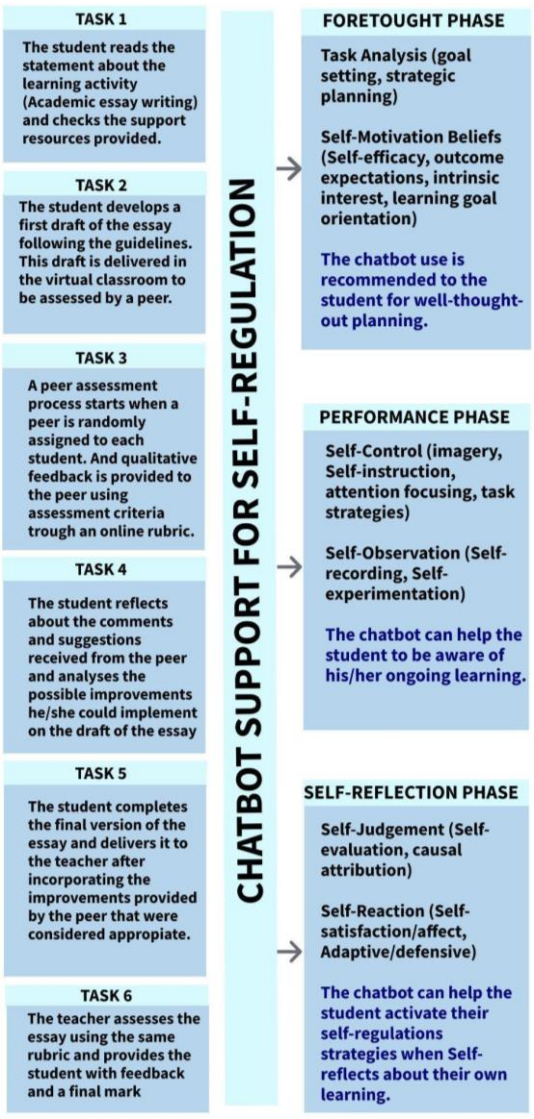


Figure 1. Self-regulation support chatbot for an essay writing activity based on Zimmerman (2002)

The design of the peer-feedback task was based on the three cornerstone principles of the DBR framework (Reeves, 2006) which envisage addressing complex problems in real-world situations, integrating design principles with technological advances to produce realistic solutions, and conducting a reflective inquiry to test and refine innovative learning environments and establish new design principles.

As a result, first we analysed the practical problem to be addressed. The main objective was to improve students' academic performance in essay writing. Being a complex learning task, it requires an already advanced level of writing, creative and critical thinking skills. As such, these skills had to be further improved along with self-regulated learning and the enhancement of the students' time management, goal setting, self-monitoring and motivation. Additionally, the course also aimed to train future professionals in the design of activities promoting self-regulated learning, learning-to-learn skills and autonomous work.

To achieve these objectives, the next step involved developing a peer-feedback task to be performed by the students within the Moodle, and a dedicated forum for communication with the teacher. The design of the solution was informed by existing design principles relating to peer review, self-regulated learning and feedback (Nicol & Macfarlane-Dick, 2006; Zimmerman, 2002). Its integration, therefore, aimed to help students work on the skills needed for the essay writing exercise more effectively, as it promoted self-reflection during the writing process. Such self-reflection on one's own work at an intermediate stage of the process was also encouraged in order to prevent students from doing the whole piece of work shortly before its delivery.

The proposed solution was then submitted to a reflective inquiry organised in iterative cycles to test and refine it in practice. For this purpose, a first iteration was carried out with the aim of testing its impact and gathering information through two data collection instruments. The analysis of the results obtained led us to formulate improvements to the proposed solution before the launch of the next iteration (second semester). The reflection on the design of the solution and its refinement based on the test results also aimed at producing new design principles.

3.1 Research questions

The guiding research questions in the study presented in this paper are the following:

- Have the chatbot and the peer feedback facilitated self-regulation of learning for the development of the academic writing activity?
- Does the design of the learning activity enhance student process performance skills?

4 Methodology

A set of data collection instruments was designed and distributed after the first iteration (first semester) in order to gain insights into the students' perception of the peer-review task and the use of the chatbot to enhance their self-regulation skills. These instruments included:

1. A questionnaire about the students' self-regulation skills.
2. A satisfaction questionnaire about the peer-review task.

Both of them included open-ended and closed-ended questions based on a 5-point Likert scale.

Some of the closed-ended questions from the satisfaction questionnaire were selected to analyse the overall satisfaction with the peer-feedback task and the chatbot that guided them through the process with personalised instructions. Additionally, two open-ended questions (one from questionnaire 1 and one from questionnaire 2) were selected to collect information on the improvements to be made to the design of the second iteration. The question selected from questionnaire 1 involves explaining the benefits and difficulties of the peer-review process, while the one selected from questionnaire 2 asks what aspects of the task could be improved.

Participants in the questionnaires are reported in *Table 1*. Most of them answered both questionnaires, and just a reduced number participated in only one of them. As such, we calculated that the total number of students who provided feedback through at least one of the two questionnaires was 280.

Table 1. Instruments and participants

Instrument	Participants
Questionnaire on self-regulation	174
Satisfaction questionnaire	157

Qualitative content analysis (Schreier, 2012) was applied to the corpus of data from the two selected open-ended questions. Open-ended responses to the two questionnaires coming from the same user were analysed together as a whole.

Atlas.ti was used for coding and analysis. The codebook for this research was developed through open coding techniques. Instead of employing a predefined set of categories or codes, open coding involved selecting bits of text that were of interest and classifying them with a keyword that emerged from the data itself (Creswell & Creswell, 2014). The codebook included definitions (Creswell & Poth, 2018).

In order to generate conclusions from the coded data, the identified codes were ranked with respect to their relevance. Following the idea to focus on the individual as a unit of analysis, such relevance was established on the basis of the number of users who

mentioned a specific conceptual category. Opting for coding at the 'individual level' (Syed & Nelson, 2015), we assigned a code even if it was deemed to be present in only one of the open-ended questions. Finally, we decided to exclude the codes with a frequency lower than 5.

5 Discussion of the results

5.1 General overview of the results of the task

The analysis of the questionnaires shows positive results regarding the overall satisfaction with the peer-review task (see Figure 2). In fact, 83.4% of the students expressed satisfaction with the assignment.



Figure 2. Level of satisfaction with the peer feedback

Data also reveals that more than 50% of students consider that the chatbot played a supporting role by fostering their self-reflection and self-regulation skills along with their 'learning to learn' competence (see Figure 3).

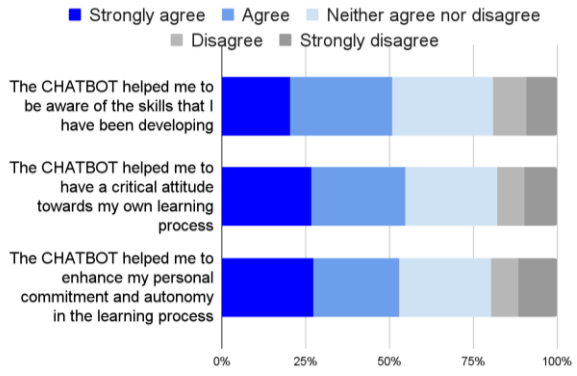


Figure 3. Students' perception of the chatbot

Suggested improvements collected from the questionnaires

Despite the overall positive results, the analysis of the answers provided to the open-ended questions also led to the identification of a number of improvements to be implemented in the second iteration (see Table 2).

Table 2. Suggested improvements (SI) for the second iteration

Nº	Suggested improvements	Frequency	Percent
1	improve the objectivity of the assessment	38	18.7
2	deal with the evaluator's superficial knowledge of the topic covered in the peer's essay	37	18.2
3	remove the inhibition to carry out an honest assessment of peers	26	12.8
4	increase expertise in the use of assessment criteria	25	12.3
5	motivate students to enhancing their commitment	23	11.3
6	enhance the expertise of the students in peer assessment	20	9.8
7	improve workflow and the supporting technology	20	9,8
8	produce an advanced draft of the essay	20	9.8
9	foster a closer relationship with the assessed peer	11	5.4
10	integrate teacher supervision	7	3.4
11	expand options regarding the feedback delivery format	7	3.4
12	improve the understanding of the task	6	2.9
13	receive feedback from more than one peer to compare the results	6	2.9
14	integrate a guide/template on the aspects to be taken into account	6	2.9

The results of the analysis emphasised the subjectivity of the feedback as one of the main concerns to be addressed in the design of the second iteration. In fact, a considerable number of students claimed that their peers might be influenced by their personal taste, preferences and knowledge. The feedback received, therefore, could largely depend on the assigned evaluator and also results in a lack of agreement with it. To face this bias, objectivity should be further fostered by providing clear instructions and supporting instruments (SI1). Another issue raised by the participants concerned the superficial knowledge that the evaluator may have of the subject matter of the peer's essay. The lack of mastery of the subject matter is considered to make the formative assessment intervention unhelpful and generate confusion instead of helping the peer. This result suggests that students should be reminded that the topics of the essays have been covered during the course and, therefore, they should feel confident in providing comments or suggestions on each of the assessment criteria of the rubric (SI2).

Students also emphasised that the evaluations of the peer's work are often insincere. There is a reluctance to criticise colleagues' essays for fear of hurting them and compromising the final results. This attitude inevitably leads students to fall into undesirable social biases. In response to this challenge, the importance of receiving critical feedback should be further nurtured (SI3). Another challenge highlighted by the participants is the interpretation of the evaluation criteria in the rubric and their correct and objective application. As such, they stated that it would be useful to further detail them and involve students in their co-design. Clearer guidelines on how to use them to provide feedback is also perceived as beneficial (SI4).

Moreover, students complained about the lack of commitment of some peers. They also claim that poor communication skills and empathy could reduce the meaningfulness of the formative assessment intervention. A clear strategy would be needed to encourage students to perform the task properly (SI5). Further preparation in providing peer feedback is also mentioned as necessary. According to them, sometimes changes are suggested without having a precise idea of how to assess a piece of work. To face this challenge, it would be advisable to improve the students' practice of giving feedback by enhancing the available instruments, instructions and resources (SI6).

The workflow of the task could be also improved with better coordination of the steps involved in the process and fine-tuning their schedule (e.g. the assessment of the draft should not be requested weeks before the final submission and the delay in receiving feedback should be avoided) (SI7). Furthermore, some difficulties are reported with the technology supporting the process (Moodle) and its monitoring (e.g. accessibility, display and download of the feedback, finding the assigned partner, etc.). Students also revealed that sometimes the work is not sufficiently developed to carry out the evaluation. As such, it is necessary to set clear guidelines on what has to be included in the draft and the required level of development since if the draft is incomplete, the feedback will also be useless (SI8).

The results also reveal that, according to the students, interaction between evaluator and evaluee (e.g. via chat or video call) should be fostered to provide meaningful feedback (SI9). The peer's lack of knowledge and the impossibility of interacting with him/her might incur the risk of providing the formative assessment in an incorrect way and, therefore, of generating misunderstandings. Moreover, not knowing the peer and not having worked side by side with him/her might also produce inputs that are not very relevant. Additional suggestions concerned the need for teacher supervision (SI10) that, according to the participants, could avoid inaccurate and incorrect assessments and ensure the quality of the feedback. The format of delivery could also be improved by providing the evaluator with more options (e.g. integrating a qualitative summary, adding comments to the file containing the essay, and attaching audio files) (SI11).

A reduced number of students also claimed that sometimes understanding the requirements of the activity is subject to the learners' personal interpretation. This challenge could be addressed by sharing detailed guidelines on how the assessment has to be carried out and the expected outcomes of the activity (SI12). Some participants also suggested involving more than one peer in the formative assessment intervention in order to gather more points of view (SI13) and compare the results obtained.

Finally, participants revealed that it would be beneficial to have a guide describing the aspects to take into account in the assessment of the essay (SI14) and a template with clear and visual information (cover, number of pages, sections, etc.) on the work to be delivered. These materials would enable the preparation of essays with the same structure and avoid doubts about the requirements to be met.

5.2 Implementation of the second iteration

With the information gathered from the test of the first iteration, the solution was refined. Improvements to the peer-feedback intervention were designed and implemented before the launch of the second iteration:

a) A video was recorded and made available to students in response to improvement requests 1, 2, 3, 4, 5, 6, 8, 12 and 14 (see Table 2). In order to improve students' expertise in peer assessment, objectivity and competence in the use of the rubric, the video fostered the appropriation of assessment criteria included in the rubric and the general principles to be taken into account in providing good feedback. Among them is the use of an appropriate communication style to express a personal opinion in a constructive, objective and

respectful manner. Furthermore, the video clarified that students are expected to provide feedback on the format of the essay, the clarity of its presentation, the structure and the organisation of content and sections. Although one of the assessment criteria regards the mastery of the subject and the problem being addressed, they are not supposed to be experts on the topic. Similarly, in order to reduce inhibition when providing a critical assessment for their peers, the video also emphasised the usefulness of feedback for improving the peer's essay. The video also strived to stimulate students' motivation to perform the task properly by clarifying the usefulness of this assignment for the development of their learning to learn and self-regulation skills. In response to improvement request 14, the video also shared clear guidelines regarding the script and the structure that an essay should follow.

b) The instructions concerning the task were refined in response to improvement requests 7, 8, 12 and 14 and published within the Moodle. The purpose of these new, more detailed instructions was to solve certain issues that had arisen in relation to the workflow and expected requirements. Besides the instruction, additional references on how to draft an essay were also made available to students within the virtual classroom.
c) A dashboard was implemented in response to improvement requests 5 and 7. This dashboard has been incorporated to enable individual and group monitoring in relation to task accomplishment. Besides providing the opportunity to display relevant information, the dashboard also encourages the student's awareness of his or her own performance and the performance of the group. This should motivate him/her to take the necessary corrective measures to ensure optimal development of the task and to complete the activity in a satisfactory way.

The results of the analysis also identify other improvements that, despite their interest, were not implemented in the second iteration due to time constraints. They refer to improvement requests 9, 11 and 13, and envisage the need for a closer relationship with the peer assessor, the integration of additional options regarding the format of feedback delivery, and the involvement of more than one student in the assessment of an essay in order to compare the assessments obtained. Their implementation could be planned and tested in future tasks.

Finally, we also decided not to consider improvement request 10 because of its lack of alignment with the objectives of the peer-assessment task. Indeed, the participation of the teacher as a supervisor is not fully aligned with the goal of fostering student autonomy along with their critical capacity in evaluating their peers. However, this concept was further clarified in the video along with the list of supporting instruments made available to students to work autonomously.

6 Conclusions

With this study, we conclude that students are satisfied with the peer-review task, because it allows them to activate, with the support of a chatbot, strategies of self-reflection on their assignment as well as self-regulation strategies of their learning process.

It is also shown that providing peer-review tasks in the middle of drafting an academic written essay keeps the student motivated and at their own pace and avoids academic procrastination.

We also conclude that the feedback received, therefore, could largely depend on the assigned evaluator and the maturity of the essay draft. Additionally, there is a reluctance to criticise colleagues' essay assignments for fear of hurting them and compromising the final results, so the relevance of receiving critical feedback should be better addressed.

Refinements emerged from the analysis that are necessary for the second iteration of the learning activity and could be translated into design principles for future iterations. In this sense, the instructional issues of the activity were improved in terms of giving students more detailed specifications aimed at fostering expertise in peer-feedback provision, promoting the objectivity of feedback, and improving understanding of the peer-review task and assessment criteria to improve appropriation. This was provided using a video explanation by the teacher and better guidance about how to write the academic essay and how to apply assessment criteria displayed with a rubric.

Improving students' motivation and engagement also emerged as one of the issues that should be addressed by providing an environment in which participants' assignment completion status can be visualised using a dashboard, or emphasising the usefulness of the task for the development of soft-skills like learning to learn and critical thinking.

Finally, according to the participants, this could avoid inaccurate and incorrect assessments and ensure the quality of the feedback and, in this sense, it is relevant to train students in terms of feedback literacy knowledge.

Acknowledgments

This contribution is made in the framework of the project: *Analysis of the effects of the provision of feedback supported by digital monitoring technologies on transversal competences* (Reference PID2019-104285GB-I00). Ministry of Science and Innovation (MICINN). National Research Agency
https://www.ub.edu/digital-feedback/es/inicio/

References

Amiel, T., & Reeves, T. C. (2008). Design-Based Research and Educational Technology: Rethinking Technology and the Research Agenda. *Educational Technology & Society*, *11*(4), 29-40.

Carless, D., & Winstone, N. (2020). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education*. https://doi.org/10.1080/13562517.2020.1782372

Cerezo, R., Esteban, M., Sánchez-Santillán, M., Núñez, J.C. (2017). Procrastinating Behavior in Computer-Based Learning Environments to Predict Performance: A Case Study in Moodle. *Frontiers in Psycholog*y, doi: 10.3389/fpsyg.2017.01403

Corgan, R., Hammer, V., Margolies &, M., Crossley, C. (2004). Making your online course successful. *Business Education Forum, 58* (3), 51-53

Creswell, J. W., & Creswell, J. D. (2014). *Research Design: Qualitative, Quantitative and Mixed Methods Approaches* (4th ed.). Sage Publications.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry & research design. Choosing among five approaches*. Sage Publications.

Dysthe, O. (2015). Writing Pedagogy in Online Settings-A Widening of Dialogic Space? In M. Deane & T. Guasch (Eds.). *Learning and Teaching Writing Online* (pp. 186-193). Ecclestone.

Fernández-Ferrer, Durall, M., & Cano, E. (2021). Chatbot for self-regulation in complex tasks: co-designing feedback in learning processes, *EDULEARN21 Proceedings*, 9433-9438.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. Re*view of Educational Research, 77* (1), 81-112.

Hounsell, D., McCune, V., Hounsell, J. & Litjens, J. (2008). The quality of guidance and feedback to students. *Higher Education Research & Development*, 27(1): 55–67.

Jia, J. (2009). CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning. *Knowledge-Based Systems*, *22*(4), 249–255. https://doi.org/10.1016/j.knosys.2008.09.001

Johnson, W. L. (2001). Pedagogical agent research at CARTE. *AI Magazine*, *22*(4), 85–85.

Klimova, B. F. (2012). The Importance of Writing. *Indian Journal of Research, 2*(1), 9–11. https://doi.org/10.15373/22501991/JAN2013/4

Latifi, S., Noroozi, O., & Talaee, E. (2021). Peer feedback or peer feedforward? enhancing students' argumentative peer learning processes and outcomes. *British Journal of Educational Technology*, *52*(2), 768–784. https://doi.org/10.1111/bjet.13054

Lee, C., Jung, S., Kim, S., &, Lee, G. G. (2009). Example-based dialog modelling for practical multi-domain dialog system. *Speech Communication*, *51*(5), 466–484. https://doi.org/10.1016/j.specom.2009.01.008

Liu, N., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, *11*(3), 279-290. https://doi.org/10.1080/13562510600680582

Liu, E.Z.-F., Lin, S.S., Chiu, C.H., & Yuan, S.M. (2001). Web-based peer review: The learner as both adapter and reviewer. *IEEE Transactions on Education, 44* (3), 246-251

Lynch, M. M. (2002). *The online educator: A guide to creating the virtual classroom*. Routledge.

Nicol, D. (2009). Assessment for learner self-regulation: Enhancing achievement in the first year using learning technologies. *Assessment & Evaluation in Higher Education*, *34*(3), 335-352. https://doi.org/10.1080/02602930802255139

Nicol, D. (2013). Peer review: Putting feedback processes in students' hands. P*erspectives on Pedagogy and Practice*, *4*(1), 111-123.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Palloff R. M., & Pratt, K. (2001). *Lessons from the cyberspace classroom: The realities of online teaching.* Jossey-Bass.

Reeves, T. (2000). Enhancing the worth of instructional technology research through "design experiments" and other development research strategies. *AERA*

Reeves, T. (2006). Design research from a technology perspective. In J. V. D. Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), *Educational design research* (pp. 52–66).

Schreier, M. (2012). *Qualitative content analysis in practice*. Sage.

Syed, M. & Nelson, S.C. (2015). Guidelines for establishing reliability when coding narrative data. *Emerging Adulthood*, *3*(6), 375-387. https://doi.org/10.1177/2167696815587648

Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of second language learning*, *9*(2), 147-170. https://mite.edu.hku.hk/f/acadstaff/399/Do_Secondary_L2_Writers_Benefit_from_Peer_Comments.pdf

Zimmerman, B. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

Evaluation of a Study Programme for the Implementation of Blended Learning at the Maldives National University

Marcelo Maina, Lourdes Guàrdia, Nati Cabrera, Maite Fernández-Ferrer,
Maria Luisa Malerba, Carolina García
Faculty of Psychology and Education Sciences
Universitat Oberta de Catalunya
Rambla de Poblenou 156, 08020, Barcelona, Spain
{mmaina, lguardia, ncabrera, mfernandezferrer, mmalerba, dgarcialon}@uoc.edu

Abstract. *Little research addresses the adoption of Blended Learning (BL) as a strategy for supporting Higher Education institutions' transformation. The purpose of the study presented below was to evaluate a teacher development Study Programme (SP) focused on the implementation of BL at the Maldives National University (MNU), as part of its digital shift towards becoming a bi-modal institution. This article examines the SP through the lens of Kirkpatrick's model in terms of satisfaction and learning gains. The sample comprised 45 lecturers and data were collected using online surveys and focus groups. The results provide insights into the quality of the SP and its relevance for lecturers adopting BL in their courses and as part of a contribution to a whole institutional strategy.*

Keywords. blended learning, digital transformation, Kirkpatrick's model, capacity building

1 Introduction

1.1 Blended Learning in Higher Education

Universities are increasingly adopting BL as part of their educational offer. Studies have demonstrated its benefits for students and lecturers (Vaughan, 2007; Bohle Carbonell, Dailey-Hebert & Gijselaers, 2012). Previous research has analysed the use of differentiated instruction in BL to address students' individual needs (Boelens, Voet & De Wever, 2018). Research also has examined the decision-making processes behind the adoption of BL programmes and courses as well as the alignment between BL and the institutional strategy (Galvis, 2018).

A BL approach has many advantages, widening access to education, providing new opportunities for career development, optimising teaching and learning time, and it is a successful strategy to promote transformational changes and development in universities (Kaur, 2013). However, staff resistance to change may constitute a major obstacle to the implementation of BL in institutions (Vaughan, 2007; Bohle Carbonell et al., 2012). The adoption of BL demands a high level of organisation and commitment (Garrison & Vaughan, 2013), especially on the faculty and lecturers' side.

For this reason, it is important to provide guidance for the adoption of BL at institutional level and to facilitate a successful transition (Porter, Graham, Spring & Welsch, 2014; Porter, Graham, Bodily & Sandberg, 2015; Graham, Woodfield & Harrison, 2013).

1.2 Kirkpatrick's Model in Educational Settings

It is essential for lecturers to receive adequate training in order to be able to perform their tasks effectively (Rienties, Brouwer & Lygo-Baker, 2013) in BL. Similarly, it is equally important to evaluate the training process itself and its outcomes (Alsalamah & Callinan, 2021) to monitor and improve the action.

Much of the research has highlighted the relevance of Kirkpatrick's model (2006) in the educational field for evaluating both training programmes (Alsalamah & Callinan, 2021; Dewi & Kartowagiran, 2018; Masood & Usmani, 2015; Wu, Hu, Gu & Lim, 2015) and BL environments (Embi, Neo & Neo, 2017). Kirkpatrick's model comprises four levels:

Reaction. This focuses on the initial reaction the training provokes in the participants, and it measures their level of satisfaction. Engagement, motivation and attention are usually indicators and trainees are given questionnaires to analyse the initial impact of the training programme through perceptions of the learning experience and content (Alsalamah & Callinan, 2021; Dewi & Kartowagiran, 2018). Previous authors (Brown, 2007; Alsalamah & Callinan, 2021) have adopted broader categories for this level and have considered it as a multidimensional facet, embracing several aspects of the training process

(learning content, materials, tools, delivery methods, etc.).

Learning. This level refers to "the extent to which participants change attitudes, improve knowledge and/or increase their skills as a result of attending the program" (Kirkpatrick & Kirkpatrick, 2006, p. 22). It takes into account the new skills and knowledge that have been acquired by the participants in the training, through which they have changed their attitudes. In the literature, this level can be assessed by means of self-evaluation surveys (Alsalamah & Callinan, 2021; Ruiz & Snoeck, 2018; Pineda, 2010). This level considers not only the new skills acquired, but also cognitive and attitudinal outcomes connected with the objectives of the training programme (Alsalamah & Callinan, 2021; Mohamed & Alias, 2012).

The third and the fourth levels, *Behaviour* and *Results*, respectively, refer to aspects of actual change and impact in the medium and long term that will be evaluated in a post-project phase some time after the training.

1.3 Purpose of the Study

The aim of this article is to evaluate a Study Programme (SP) designed in the AMED project — *Advancing Higher Education in Maldives through E-learning Development*— using Kirkpatrick's model. One of the main objectives of the AMED project was the co-creation and piloting of a SP for professional development focusing on the implementation of BL. The idea was to contribute to the professional development of MNU academic staff for the transition to becoming a bi-modal university by providing the foundations of e-learning and BL approaches, and training lecturers in the design, planning and implementation of competency-based BL.

This article reports on the results of the pilot that took place at the MNU with a main focus on the perceptions of a group of lecturers. Of the four levels of the model, two are examined: reaction and learning. The guiding research questions of the study were:

RQ1. How do lecturers react to the SP in terms of satisfaction, engagement and relevance?

RQ2. How do lecturers perceive their learning gains, skills and attitudes towards BL after the SP?

In the SP, the lecturers' learning pathway consisted of the following modules:

1. *Foundations of e-learning*. This core course introduced the students to key concepts, principles and practices of e-learning.
2. *Designing digital learning*: Participants of this module were asked to put into practice course design models and principles for technology enhanced and enabled learning, flexible and open learning, and networked learning.
3. *The networked teacher*: In this course, participants explored new forms of educational relationships enabled by technologies: the changing roles of teachers, learner empowerment strategies, collaborative approaches in action, learning scaffolding techniques, e-assessment methods and instruments.

The SP concluded with capstone projects (CP), directly connected to the modules. The CP required lecturers to design a BL activity/experience blueprint together with a learning facilitation plan and, when possible, the development of a prototype. The SP total workload was 140 hours and the Virtual Learning Environment adopted was Moodle.

2 Research design

The pilot took place between March and December 2021. It involved the participation of 45 lecturers.

Different data were collected according to the *reaction* and *learning* levels of Kirkpatrick's model (2006). A survey, designed to evaluate both levels, gathered information regarding the participants' satisfaction, engagement and the relevance of the SP (Dixon, 2010; Hamid & Pihie, 2004). In addition, it enquired about the participants' perceptions of their learning outcomes regarding knowledge, skills and attitudes towards BL. The survey was answered by 43 lecturers after completing the SP, using a Likert Scale of 1-4 (1= strongly disagree and 4= strongly agree). Surveys were analysed through descriptive statistics.

With the aim of gaining insights into the lecturers' perceptions of their satisfaction, acquired knowledge and skills, and what they foresaw for the future in terms of BL, two focus groups with five lecturers each were conducted. The design of focus group protocols was carried out according to Kirkpatrick's (2006) model and the Specific Review Standards from the Quality Matter Higher Education Rubric (2020). A qualitative content analysis (Schreier, 2012) was applied to the corpus of data, using a deductive approach based on the preconceived theoretical concepts and dimensions. *Atlas.ti* was used for coding and analysis.

3 Results

The results indicate the participants' perceptions of the SP and of improvements in knowledge, increased skills and attitudes. Because of the practical approach of the SP, carried out through the CP, perceptions of changes in learning include perceptions of short-term changes in behaviour.

3.1 Reaction

Lecturers valued the SP as a positive experience and found it useful for their professional development. One of the participants stated: "*It was really helpful. I have learnt a lot about Blended Learning*". Overall, participants reacted favourably to the SP, which can be

evidenced in expressions such as the following: "*I really liked it in every sense*". Moreover, the survey showed that lecturers were satisfied with the quality of the SP (67% agree; 16% strongly agree).

Lecturers agreed that the programme structure was clear and helped them gain a better understanding of what was expected in terms of activities. Elements such as flexibility and module progression were considered adequate, particularly task distribution and sequence: "*The task progression was very helpful in each module*". Also, the connections between modules and CP were perceived as a continuum: "*The design and how it was presented to us was developed continuously*".

Participants were satisfied with the materials provided (70% agree; 21% strongly agree). Lecturers found the readings relevant and applicable to their teaching practice: "*I liked the readings. I learnt how to incorporate them into my classes, giving precise and short readings rather than lengthy ones*". However, some participants manifested that they felt overloaded and experienced time constraints due to their regular teaching responsibilities. When asked about their ability to keep up with the workload on the SP, 16% strongly disagreed and 23% of the participants disagreed.

Regarding learning activities, 67% of the lecturers agreed to being satisfied with them (16% strongly agree), concurring that the activities were relevant, engaging and promoted the achievement of the learning outcomes. Similar results were reported for assessment activities (70% agree; 16% strongly agree). In this regard, quizzes allowed lecturers to track their progress: "*they told us how much we know and where we were confused*".

Lecturers also appreciated the guidance and support offered by facilitators (47% agree; 42% strongly agree). The suggestions and examples provided in templates and other methodological instruments were beneficial for their learning processes. However, some of them pointed out that more guidance would have been needed.

In terms of engagement, the survey indicated that most of the lecturers were motivated to learn the content (60% agree; 16% strongly agree). Some of the participants highlighted that they were committed to the SP, commenting that they read the study material (74% agree) and self-organised study groups as a strategy to discuss and clarify ideas.

Despite the fact that the content was new for some participants, they found it interesting, useful and up-to-date (67% agree; 21% strongly agree). Findings indicate that lecturers considered the SP meaningful for their work (63% agree; 30% strongly agree). Also, it contributed to their professional development: "*The AMED project made me open my eyes to the things that I can do. I look forward to improving my teaching subjects. I benefited a lot*".

Most of the lecturers considered that the SP met their expectations (58% agree; 9% strongly agree), and they would recommend it to other colleagues (60% agree; 14% strongly agree), as one of the participants stated: "*Since I already have some experience, this was very helpful. I would recommend it to other lecturers*".

3.2 Learning

Lecturers acknowledged the potential of the SP to face the teaching challenges that the COVID-19 pandemic entailed. According to them, learning outcomes were clear, consistent, and helpful in understanding how BL works. After the SP, they reported being aware of the benefits of using BL in higher education (58% agree; 42 strongly agree). Specifically, lecturers expressed that they understood the concept of BL and its implementation in the classroom: "*Through this programme, I learnt a lot about BL and the different models that we can use. It was very helpful and I try to incorporate it in my classroom as well*".

The participants developed a clear idea of the relationship between competencies and learning outcomes (58% agree; 37% strongly agree). Also, they distinguished the difference between student and teacher-oriented approaches (56% agree; 44% strongly agree).

Participants perceived that not only their knowledge, but also their practical skills in BL improved as a result of undertaking the SP and developing the CP. The survey revealed that lecturers were able to design a coherent BL activity matching the learning outcomes and assessment (72% agree; 28% strongly agree). In addition, lecturers reported that the SP stimulated a better use of technology for teaching purposes, 70% of them agreed that they were able to select the most appropriate technologies to implement in their courses.

They highlighted the role of facilitators in the process of learning how to design and tailor instructional strategies for BL, for instance, the regulation of reading resources: *"they are very good and I'm learning from them the techniques and how I can make it better"*.

Regarding the CP, lecturers explained that it was stimulating and rewarding because they could put into practice what they had learnt by designing or redesigning their courses. The CP allowed lecturers to introduce BL techniques into their courses and progressively change the focus from a teacher-led approach to a student-centred one as shown in the following quote: "*I gave them activities which they could do in their own time after reading the material and then having the face-to-face sessions as discussion or revision time. So, I have changed one of the modules rather than going through the reading material in the face-to-face class*".

Despite some difficulties experienced due to the COVID-19 pandemic and the fact that there was some resistance to BL among faculty, lecturers plan to support learning activities with increased use of technology (65% agree; 35% strongly agree). One of

the participants claimed: "*I am definitely going to use the techniques that I learnt in the course*". They are motivated to participate in the e-learning change and expressed their willingness to apply student-oriented approaches (65% agree; 35% strongly agree).

4 Discussion

Considering the results, it is possible to answer the RQs. Regarding the first one, lecturers reacted positively to the SP as a teacher development strategy for the adoption of BL in the MNU digital transformation. Lecturers, overall, were satisfied and engaged. Secondly, they enhanced their pedagogical and digital competencies, influencing their teaching delivery approach and willingness to adopt BL. The SP played a significant role in their professional development, stimulating teaching practices that promote accessibility to higher education in the Maldivian context.

As Graham et al. (2013) mentioned, the institutional change begins at the faculty level. This study strengthens the importance of professional development offered to faculty to develop new pedagogical and technological skills to implement BL (Porter et al., 2014), demonstrating that the SP fostered the necessary knowledge, skills and attitudes for digital transformation. Lecturers are aware of BL's potential to face challenges such as accessibility to higher education, considering it an innovative solution for supporting the digital transformation of MNU. However, alignment among professional development opportunities, policies and support structures is needed to improve teaching practices and progressively reduce the resistance to new teaching approaches associated with BL.

To sum up, the SP had a positive impact on the faculty's training and professional development, raising awareness of the benefits of adopting BL approaches. It leads to institutional change that may enhance the quality of the teaching and learning in the institution. However, as Garrison and Vaughan, (2013) point out, a transformational institutional change related to BL also requires collaborative leadership and engagement at all institutional levels.

5 Conclusion

The results of the study revealed that the SP and its implementation plays a key role in the process of institutional transformation of MNU towards a strategic integration of BL. The results of the study also showed that the Kirkpatrick's model (Kirkpatrick & Kirkpatrick, 2006) provided a valuable tool to determine the effectiveness of the SP as well as its strengths and improvements for next editions.

Future studies should incorporate the third and fourth levels of Kirkpatrick's model (Kirkpatrick & Kirkpatrick, 2006), *Behaviour* and *Results*, to evaluate not only the implementation of the next editions of the SP, but also additional institutional strategies enacted by the staff to improve the integration of BL at MNU.

Acknowledgments

AMED project (2019-2022) has received funding from the European Union's Erasmus + Capacity Building in the Field of Higher Education, Key Action 2: Cooperation for innovation and the exchange of good practices.

References

Alsalamah, A., & Callinan, C. (2021). Adaptation of Kirkpatrick's four-level model of training criteria to evaluate training programmes for head teachers. *Education Sciences*, *11*(3). https://doi.org/10.3390/educsci11030116

Boelens, R., Voet, M., & De Wever, B. (2018). The design of blended learning in response to student diversity in higher education: Instructors' views and use of differentiated instruction in blended learning. *Computers & Education*, *120*, 197–212. https://doi.org/10.1016/j.compedu.2018.02.009

Bohle Carbonell, K., Dailey-Hebert, A., & Gijselaers, W. (2013). Unleashing the creative potential of faculty to create blended learning. *The Internet and Higher Education*, *18*, 29–37. https://doi.org/10.1016/j.iheduc.2012.10.004

Brown, K. G. (2007). Training evaluation. In *Encyclopaedia of Industrial and Organisational Psychology* (pp. 820–823). Rogelberg, S.G., Ed. SAGE Publications.

Cahapay, M. (2021). Kirkpatrick Model: Its Limitations as Used in Higher Education Evaluation. *International Journal of Assessment Tools in Education*, *8*(1), 135–144. https://doi.org/10.21449/ijate.856143

Dewi, L. R., & Kartowagiran, B. (2018). An evaluation of internship program by using Kirkpatrick evaluation model. *REiD (Research and Evaluation in Education)*, *4*(2), 155–163. https://doi.org/10.21831/reid.v4i2.22495

Dixson, M. D. (2015). Measuring Student Engagement in the Online Course: The Online Student Engagement Scale (OSE). *Online Learning*, *19*(4). https://doi.org/10.24059/olj.v19i4.561

Embi, Z., Neo, T.-K., & Neo, M. (2017). Using Kirkpatrick's Evaluation Model in a Multimedia-based Blended Learning Environment. *Journal of*

Multimedia Information System, *4*(3), 115–122. https://doi.org/10.9717/JMIS.2017.4.3.115

Galvis, Á. H. (2018). Supporting decision-making processes on blended learning in higher education: literature and good practices review. *International Journal of Educational Technology in Higher Education*, *15*(1), 25. https://doi.org/10.1186/s41239-018-0106-1

Garrison, D. R., & Vaughan, N. D. (2013). Institutional change and leadership associated with blended learning innovation: Two case studies. *The Internet and Higher Education*, *18*, 24–28. https://doi.org/10.1016/j.iheduc.2012.09.001

Graham, C. R., Woodfield, W., & Harrison, J. B. (2013). A framework for institutional adoption and implementation of blended learning in higher education. *The Internet and Higher Education*, *18*, 4–14. https://doi.org/10.1016/j.iheduc.2012.09.003

Hamid, J. A., & Pihie, Z. L. (2004). Students' perception of the quality of teaching and learning in business studies programs. *Pertanika J. Soc. Sci & Hum.*, *12*(1), 71–86.

Kaur, M. (2013). Blended Learning - Its Challenges and Future. *Procedia - Social and Behavioural Sciences*, *93*, 612–617. https://doi.org/10.1016/j.sbspro.2013.09.248

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The Four Levels* (3rd ed.). Berrett-Koehler Publication.

Masood, Q. R., & Usmani, M. A. W. (2015). Evaluation of medical teacher's training program through Kirkpatrick's model. *KMUJ 2015*, *7*(2), 76–80.

Pineda, P. (2010). Evaluation of training in organisations: a proposal for an integrated model. *Journal of European Industrial Training*, *34*(7), 673–693. https://doi.org/https://doi.org/10.1108/03090591011070789

Porter, W. W., Graham, C. R., Bodily, R. G., & Sandberg, D. S. (2016). A qualitative analysis of institutional drivers and barriers to blended learning adoption in higher education. *The Internet and Higher Education*, *28*, 17–27. https://doi.org/10.1016/j.iheduc.2015.08.003

Porter, W. W., Graham, C. R., Spring, K. A., & Welch, K. R. (2014). Blended learning in higher education: Institutional adoption and implementation. *Computers & Education*, *75*, 185–195. https://doi.org/10.1016/j.compedu.2014.02.011

QM Higher Education Rubric, Quality Matters. Used under licence. All rights reserved. (2020). https://www.qualitymatters.org/sites/default/files/PDFs/StandardsfromtheQMHigherEducationRubric.pdf

Rienties, B., Brouwer, N. Z., & Lygo-Baker, S. (2013). The effects of online professional development on higher education teachers' beliefs and intentions towards learning facilitation and technology. *Teaching and Teacher Education*, *29*(1), 122–133. https://doi.org/10.1016/j.tate.2012.09.002

Rosmah Mohamed, & Arni Ariyani Sarlis Alias. (2012, March 12). *Evaluating the Effectiveness of a Training Program Using the Four Level Kirkpatrick Model in the Banking Sector in Malaysia* [Conference]. 3rd International Conference on Business and Economic Research, Bandung, Indonesia. http://library.oum.edu.my/repository/717/

Ruiz, J., & Snoeck, M. (2018). Adapting Kirkpatrick's evaluation model to technology enhanced learning. *Proceedings of the 21st ACM/IEEE International Conference on Model Driven Engineering Languages and Systems: Companion Proceedings*, 135–142. https://doi.org/10.1145/3270112.3270114

Schreier, Margrit. (2012). *Qualitative Content Analysis in Practice*. Sage Publications.

Vaughan, N. (2007). Perspectives on Blended Learning in Higher Education. *International Journal on E-Learning*, *6*(1), 81–94.

Wu, B., Hu, Y., Gu, X., & Lim, C. P. (2016). Professional Development of New Higher Education Teachers with Information and Communication Technology in Shanghai. *Journal of Educational Computing Research*, *54*(4), 531–562. https://doi.org/10.1177/0735633115621922

Digital Maturity of Schools - Explanation, Literature Review and Analysis

Melita Milić, Blaženka Divjak
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, Varaždin, Croatia
{mzahtila, bdivjak}@foi.hr

Abstract. *In this paper we present a systematic literature review of studies related to digital maturity of schools. The focus is on the identification of the most significant papers from the field of digital maturity of schools, essential elements of digital maturity frameworks and their validation, as well as the connection with other sectors related to the school. The analysis shows that ICT infrastructure is an important but not sufficient element for digital transformation. Digital transformation is also influenced by digital competences of schools' employees, use of ICT in teaching and learning and administration, and the strategic approach to digital transformation.*

Keywords. digital maturity, digital maturity framework, schools, systematic literature review

1 Introduction

Digital maturity of schools is an important concept which shows the extent of schools' willingness for a digital transformation and/or the level of implementation of digitalization. Digital transformation is a systematic and planned change which schools must undergo in order to prepare their students for the use of digital technology in an efficient and safe manner.

The term "digital maturity" is often interchangeably used with "digital transformation" to describe what a company or an institution has already accomplished in terms of conducting transformational efforts and in what ways they are systematically preparing for adapting to an increasingly digital environment in order to remain competitive (Begicevic Redjep et al., 2021). Further, Chanias & Hess (2016) defined digital maturity as "the status of a company's digital transformation" – it describes "what a company has already achieved with regard to transformation efforts". Therefore, digital maturity models „on the one hand aim at measuring the current level of a company's digitalization, on the other hand at providing a model path to digital maturity" (Henderson et al., 2021). The development of digital competences is necessary to support people in strengthening themselves for life and work in a digital environment during the time of the 4th industrial revolution.

Schools have an important role in transferring knowledge and developing digital competences. Therefore, the purposeful use of Information and Communications Technology (ICT) in schools must be systematically planned and implemented. The use of ICT in schools is not a matter of individuals' enthusiasm anymore, but rather of a systematic approach planned and implemented at a school level, in accordance with local and state policies (Zugec et al., 2018). A European Union (EU) survey showed that 44% of EU citizens and 47% of Croatian citizens between the ages of 16 and 74 do not have basic digital skills (DESI - Compare Countries Progress — Digital Scoreboard - Data & Indicators, n.d.), while retraining and acquiring new skills should be the key element of national recovery plans related to the COVID-19 pandemic. Educational authorities and teachers must put in additional effort to improve digital competences of teachers to be used in the regular educational process, as well as in extraordinary contexts and challenges such as those brought by the COVID-19 pandemic (Svetec & Divjak, 2021).

Systematic application of digital technologies in educational and administrative processes in schools represents the basis of digital maturity in schools. Due to the growing importance of technology, the concept of digital maturity of schools is becoming more significant within contemporary educational systems. Continuous monitoring of schools encourages the improvement of digital skills of the 21st century citizens in order to support them in the safe use of digital technology for improving social engagement (Fozo Attila & Racsko, 2020).

The concept of digital maturity has been mostly used in industry. There is a systematic literature review (SLR) of digital maturity models (DMM) for companies that evaluates existing DMMs with regard to their conformity to quality standards and theoretical foundation. (Thordsen et al., 2020)

Literature from the field of digital maturity of schools is scarce and points to inconsistencies, in defining the model of maturity, as well as in the terminology. Pata et al. (2021) noticed that many articles cover digital competences of teachers and/or

students rather than considering a holistic approach which would also include other domains.

The main purpose of this research emerges from the previous statement, and the goal of this review is to provide an in-depth and systematic overview of the literature that refers to models of digital maturity of schools.

Furthermore, for maturity models, it is necessary to check their validity. According to Thordsen et al. (2020), the validity of measurement is evaluated through the underlying assertions, building a complex net of arguments to back up the findings.

This paper presents a literature review and an analysis of identified models of digital maturity of schools. Based on the literature review, good practices and shortcomings in former studies are identified and recommendations for further research are offered.

This paper is structured as follows. The introductory part which explains the term of digital maturity in schools, the actuality and relevance of the topic, goals, and contributions of the paper. In the second section, the methodology of the research is introduced, and research questions are given, whilst in the third section the findings on digital maturity frameworks for schools are presented, including analysis of former research and recommendations for further research. The conclusion offers a summary of the paper.

2 Research questions and search methodology

At the beginning of this paper the most important research questions were highlighted, followed by the description of the research methodology. The goal of this paper was to implement an SLR and then to answer four research questions related to the analyzed papers:

RQ 1. Which are the most significant papers in the field of digital maturity of schools?

RQ 2. Is a digital maturity framework described in the paper?

RQ 3. Does the paper cover the validation of the framework?

RQ 4. Is the digital maturity of schools connected to other sectors, organizations, and public bodies?

The process of acquiring answers to the posed research questions was initiated with a systematic overview of the literature composed of published scientific papers which were searched for in relevant databases: Web of Science (WoS) and Scopus, accessed on the 1st of December 2021 and 27th of April 2022. Similar to Divjak et al. (2022), the process of acquiring the answers was conducted in three phases, as shown in Figure 1.

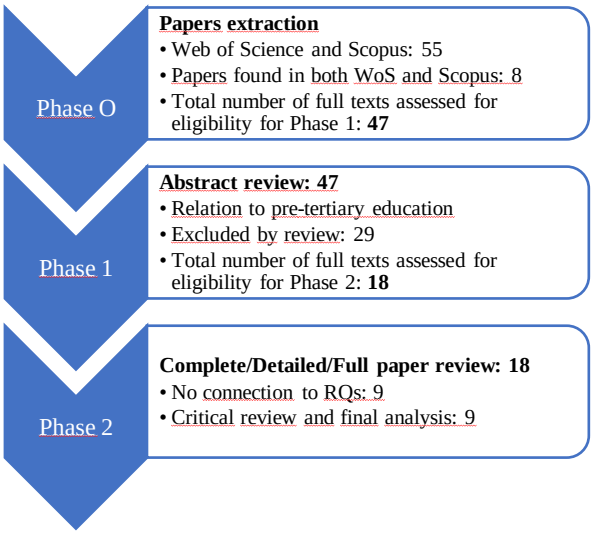


Figure 1. Phases of literature review

WoS and Scopus were searched based on the following keywords: "digital maturity" AND school.

The searching structure and the number of chosen papers from individual sources is shown in Table 1.

A total of 55 research papers were found in both databases. Eight identical papers were identified in both databases, thus making 47 the total number of research papers extracted in Phase 0. In Phase 1, the papers' summaries were reviewed in order to identify and analyze research related to pre-tertiary education (schools).

Finally, we got 18 relevant papers shown in Table 1., sorted alphabetically. Aside from basic data about the paper, the table displays answers to the second, third and fourth research question.

Table 1. Review of RQs

Reference	RQ 2.	RQ 3.	RQ 4.
(Balaban et al., 2018)	yes	yes	yes
(Begicevic Redjep, n.d.)	yes	yes	yes
(Dejanovic & Jugo, 2019)	no	no	no
(Gaftandzhieva et al., 2022.)	yes	no	yes
(Ifenthaler & Egloffstein, 2020)	yes	yes	no
(Jeladze & Pata, 2018b)	no	no	yes
(Jeladze & Pata, 2018a)	yes	yes	no
(Jugo et al., 2017)	yes	no	yes
(Klacmer Calopa et al., 2018)	no	no	yes
(Kolic- Vehovec et al., 2015)	no	no	no
(Fozo Attila & Racsko, 2020)	no	no	yes

(Leonidova et al., 2020)	no	no	no
(Mekovec et al., 2019)	no	no	no
(Pata et al., 2021)	no	no	yes
(Sotiriou et al., 2016)	yes	no	yes
(Towndrow & Fareed, 2015)	no	no	no
(Vuk et al., 2020)	no	no	no
(Zugec et al., 2018)	yes	no	yes

3 Findings and discussion

Below we describe the results and certain limitations of our study.

Among the 18 chosen research papers, Croatia is the most represented country based on the number of published papers (nine), followed by United Kingdom (seven). Other countries are represented with a maximum number of three papers. Considering the research papers which were chosen due to their relevance for pre-tertiary education, Croatia once again stands out with nine published papers, which makes up a half of the chosen research papers. The explanation behind such a great number of Croatian papers is the E-Škole (e-Schools) project ('E-Škole', n.d.) during which broad research about the digital maturity of schools was conducted.

The full name of the entire program is "e-Schools: a comprehensive informatization of school operation processes and teaching processes aimed at the development of digitally mature schools for the 21st century". The program consisted of two phases. The first pilot phase of the project started in 2015 and lasted until 2018, and the second phase lasted from 2018 until 2022. The goal of the e-Schools program was to establish a system for digitally mature schools through piloting and evaluating the application of ICT in educational and administration processes in 10% of schools in Croatia. The general goal of the e-Schools program is to support the strengthening of primary and secondary education systems' capacity to prepare students for the labor market, further education and lifelong learning ('E-Škole', n.d.).

3.1 Quote analysis

In the analysis of the 18 research papers which were identified as relevant for pre-tertiary education, four papers stand out. First, Ifenthaler & Egloffstein (2020), with 10 citations in the Scopus database and five citations in the WoS database, is the highest cited. Moreover, there are three papers (Balaban et al., 2018; Sotiriou et al., 2016; Jeladze & Pata, 2018b) with nine citations each in the Scopus database, and fewer quotes in the WoS database.

The oldest research paper from the chosen set of papers related to pre-tertiary education is a paper published in 2015 (Kolic-Vehovec et al., 2015).

In order to gain insight into the interconnections of the selected papers, it was necessary to search for all the references listed in the 18 papers. The Gephi tool was used for the analysis. The citation analysis was conducted on the chosen papers, all 18 of them. Those 18 papers were cited by 47 unique papers, which makes 65 nodes in total. The analysis found 53 connections between them (edges in the graph). Visualization was created by using Force Atlas 2, which is appropriate for small and medium graphs and adapted for a qualitative interpretation of graphs. The option of non-overlapping was enabled in order to see the formed clusters more clearly.

By extracting the biggest nodes in the Gephi tool, the key research papers were recognized, which answers RQ1: Which are the most significant papers that cover the field of digital maturity of schools?

Figure 2 shows the visualization containing all the 18 research papers and three more important papers that appeared in the citation analysis.

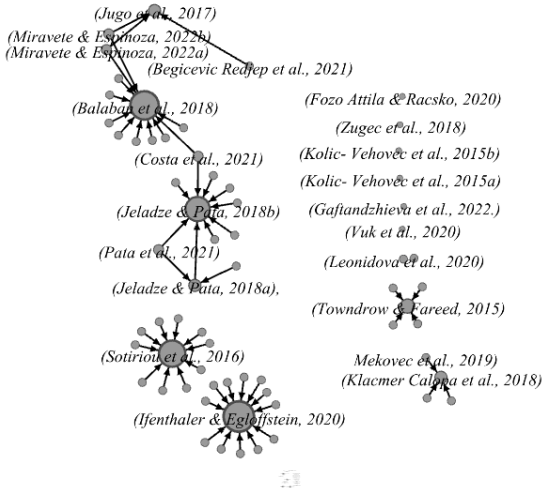


Figure 2. Correlation of quoted papers

The right side of Figure 2 shows papers with less citations. Besides Ifenthaler & Egloffstein (2020) and Sotiriou et al. (2016) on the left side, there is one larger group which is presented by Balaban et al. (2018) and one by Jeladze & Pata (2018b). Those two papers are citated by Costa et al. (2021), presenting the bridge between the two groups. Costa et al. (2021) and both papers from authors Miravete & Espinosa were not found neither in the WoS nor in the Scopus database during the search phase, and they should be included in further consideration. Balaban et al. (2018) is the earliest paper, with the highest number of citations, therefore it is the most significant in third group.

3.2 Digital maturity of schools

The coordination of all stakeholders and elements of the change is essential for the digital transformation of

the schools and project sustainability. In order to support meaningful and sustainable digital transformation, a unique and fully comprehensive concept of digital maturity of schools is needed. One such concept is described in Begicevic Redjep et al. (2018), but we will show other relevant concepts as well.

The analysis regarding the research questions is shown in Table 1. The place foreseen for the answer to RQ 2 is marked affirmatively even though the paper only includes some elements of the framework.

The summative overview of the answers to the research questions is shown in Table 2.

Table 2. Summative overview to the posed RQs

	RQ 2	RQ 3	RQ 4
Yes	8	4	10
No	10	14	8

Only eight studies refer to the digital maturity framework. From out of these eight research papers, only four had undergone framework validation. Most of the papers, ten of them, connect the digital maturity of schools with school investments and other sectors. The investment sponsors are the European Union (European Social Fund, European Regional Development Fund), and national authorities, such as education ministries. Other bodies and institutions, such as the Institute for Prospective Technological Studies, Joint Research Centre, provide support in the development of digital maturity models. There are also different initiatives in countries, such as strategies for digital education (e.g., Hungary, Croatia), which also support the development of digital maturity models. From the 18 analyzed papers, six do not provide elements which could offer positive answers to posed research questions RQ 2. - RQ 4., thus they are omitted from Figure 3.

The answers to all three research questions (RQ 2 – RQ 4) were affirmative only in two papers (Balaban et al., 2018; Begicevic Redjep et al., 2021) that discuss the Framework for Digital Maturity (FDMS) designed for the needs of Croatian schools within the e-Schools project.

The FDMS is described in four of the eight papers, and it influenced the model named DigBGSchool of digital maturity of schools in Bulgaria. Self-reflection in the Effective Learning by Fostering the Use of Innovative Educational Technologies (SELFIE) tool was taken as a basis for FDMS. SELFIE originated within EU's educational action as a free tool designed to help schools embed digital technologies into teaching, learning and assessment. (About SELFIE | European Education Area, n.d.).

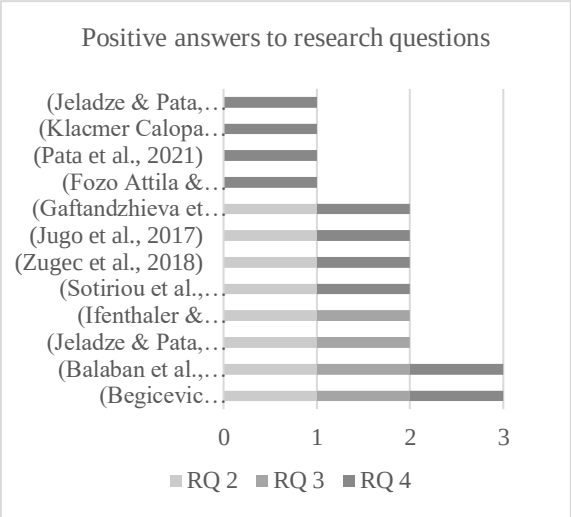


Figure 3. Display of research papers that positively answered the individual research questions

An Irish model, the Digital Schools of Distinction Program (DSoDP) (Coffey, n.d.), was developed before the FDMS, while the Maturity Model for Educational Organizations (MMEO) (Ifenthaler & Egloffstein, 2020) came to be after the FDMS. The MMEO was developed based on six models: Digital Maturity & Transformation Report, Digitalization Baromenter, Maturity Index Industry 4.0, Digital Transformation Report, Business Index DIGITAL and Model of digital maturity (Ifenthaler & Egloffstein, 2020). All of the models are recent and the most recent one is the BigBGSchool, developed in 2021 for the needs of schools in Bulgaria (Gaftandzhieva et al., 2022).

In all the frameworks (models) the ICT infrastructure element is recognized as a cornerstone element without which achieving the digital maturity of schools is impossible. We noticed that all models have also the element of planning, managing and strategic leadership in some form, as well as the ICT culture element. Most differences can be found in relation to the ICT learning and teaching domain. This domain is not explicitly mentioned in the MMEO. Furthermore, the domain related to developing digital competences is not emphasized as a separate domain in the DSoDP.

Table 3. can contribute to answering the third research question related to the validation of the framework, if any.

Table 3. Validation of the framework

Reference	RQ 3.
(Balaban et al., 2018)	yes *p*-value Spearman correlation coefficient
(Begicevic Redjep, n.d.)	yes Fornell–Larcker
(Ifenthaler & Egloffstein, 2020)	yes *p*-value

	Spearman correlation coefficient
(Jeladze & Pata, 2018a)	yes regression analysis

Four research papers which contain validation of the framework are marked with the medium color in Figure 3. Two of the papers refer to regression analysis (Sotiriou et al., 2016; Jeladze & Pata, 2018a). Data collected through the external evaluation of the e-Schools pilot project in Croatia were used in corresponding papers (Balaban et al., 2018; Begicevic Redjep et al., 2021) for validation purposes. The paper by Balaban et al. (2018) describes 38 elements and validation by the Spearman´s coefficient. At the end of the pilot project, Begicevic Redjep et al. (2021) measured the validity and reliability with SmartPLS and Cronbach alpha, and 24 indicators were retained as relevant. Ifenthaler & Egloffstein (2020) also used Spearman´s coefficient in the validation process. To conclude, only four papers (Balaban et al., 2018; Begicevic Redjep, 2021; Ifenthaler & Egloffstein, 2020; Jeladze & Pata, 2018a) conducted the maturity framework validation.

As many as ten research papers connect digital maturity frameworks with other sectors, institutions or bodies that are related to schools, as shown in Table 4.

Table 4. Connection with other sectors

Reference	RQ 4.
(Balaban et al., 2018)	yes investing in education - European Social Fund (ESF); European Regional Development Fund (ERDF)
(Begicevic Redjep, n.d.)	yes investing in education - ESF; ERDF
(Gaftandzhieva et al., 2022.)	yes investing in education - Ministry of Science and Education of Bulgaria
(Jeladze & Pata, 2018b)	yes development of society as a whole
(Jugo et al., 2017)	yes Investing in Education - EU Commission, Institute for Prospective Technological Studies, Joint Research Centre
(Klacmer Calopa et al., 2018)	yes Investing in Education - ESF; ERDF
(Fozo Attila & Racsko, 2020)	yes Investing in Education - Hungary's Strategy for Digital Education
(Pata et al., 2021)	yes Investing in Education - Estonian Ministry of Education and Research
(Sotiriou et al., 2016)	yes Investing in Education - Open Discovery Space
(Zugec et al., 2018)	yes Investing in Education - ESF; ERDF

In the paper by Jeladze & Pata (2018b) it is described that the digital maturity of schools affects the development of society in its entirety, while other papers (Klacmer Calopa et al., 2018; Pata et al., 2021; Fozo Attila & Racsko, 2020; Gaftandzhieva et al., 2022; Jugo et al., 2017; Zugec et al., 2018; Sotiriou et al., 2016; Balaban et al., 2018; Begicevic Redjep et al., 2021) link it to investment in education on a national level, as well as EU level, using e.g. the European Social Fund and the European Regional Development Fund, which shows that the importance of investing in skills of digital education has been recognized Ifenthaler & Egloffstein (2020) and Jeladze & Pata (2018) do not showcase any clear connections to other sectors.

When taking in consideration the entirety of the analyzed research on the digital maturity of schools, there is inconsistency in defining a maturity model and the use of terminology. Pata et al. (2021) noticed that there are many articles regarding digital competences of teachers and/or students and just a few taking a broader perspective on digital transformation of schools. In comparison with Thordsen et al. (2020), it can be determined that a series of articles about the FDMS formed comprehensive research in this context, including definitions and measurement procedures, the theoretical framework and empirical aspects.

In order to take an even broader perspective it would be necessary to link digital maturity of schools to other elements that are not part of the existing frameworks' scope or impact evaluation, for example monitoring student success in the next stage of education, as well as their careers after finishing school. The COVID-19 pandemic and isolation period discovered the fragility of digitally immature organizations (Fletcher & Griffiths, 2020; Svetec & Divjak, 2021) during external threats. Were digitally mature schools better at managing (themselves) during the pandemic by using all available resources (digital books, open content, …) and/or did this content contribute to the digital maturity of schools? This research question is still awaiting an answer. The limitations of this research are related to the selection of keywords, databases and the language in which the analyzed papers were written.

4 Conclusion

The goal of this paper was present a systematic review of literature that refers to identification of digital maturity frameworks for schools, elements of the frameworks, their validation, and connection to other sectors. The fact that only 18 research papers from the field of digital maturity of schools were identified shows that literature in this territory is scarce. Thanks to the e-Schools project, Croatia is the most represented country in literature that covers the field of digital maturity of schools.

Only four of 18 papers reported on some sort of framework validation.

The element of ICT infrastructure was recognized in all frameworks, while differences can be found in the ICT learning and teaching domain and the developing digital competences domain, which are not explicitly recognized all the frameworks as separate domains.

This SLR demonstrates the need for a holistic and comprehensive approach to defining and measuring the digital maturity of schools. Furthermore, in order to analyze the digital maturity of schools as part of broader educational and societal ecosystem, it would be necessary to link the digital maturity of schools to other elements, for example monitoring student success in the next stage of education, as well as their careers after finishing school.

Finally, the unprecedented situation with the COVID-19 pandemic revealed the fragility of less digitally mature organizations to act efficiently in times of crisis. Therefore, it is of utmost importance to further develop the concept and instruments related to the digital maturity of schools.

References

About SELFIE | European Education Area. (n.d.). Retrieved 17 May 2022, from https://education.ec.europa.eu/node/1544

Balaban, I., Begicevic Redjep, N., & Klacmer Calopa, M. (2018). The Analysis of Digital Maturity of Schools in Croatia. *International Journal of Emerging Technologies in Learning (IJET)*, *13*(06), 4. https://doi.org/10.3991/ijet.v13i06.7844

Begicevic Redjep, N. (n.d.). *Assessing digital maturity of schools: Framework and instrument.*

Begicevic Redjep, N., Balaban, I., & Zugec, B. (2021). Assessing digital maturity of schools: *Framework* and instrument. *Technology, Pedagogy and Education*, *30*(5), 643–658. https://doi.org/10.1080/1475939X.2021.1944291

Chanias, S., & Hess, T. (2016). How digital are we? Maturity models for the assessment of a *company's* status in the digital transformation. *Rep./Institut Für Wirtschaftsinformatik Und Neue Medien*, *2*, 1–14.

Croatia—National Development Strategy 2030 | Digital Skills and Jobs Platform. (n.d.). Retrieved 16 May 2022, from https://digital-skills-jobs.europa.eu/en/actions/national-initiatives/national-strategies/croatia-national-development-strategy-2030

Dejanovic, L., & Jugo, G. (2019). Managing Integration Of Digital Technologies In Schools—Principal Role And Competences. *12th International Conference Of Education, Research And Innovation (Iceri2019).*

DESI - Compare countries progress—Digital Scoreboard—Data & Indicators. (n.d.). Retrieved 16 May 2022, from https://digital-agenda-data.eu/charts/desi-see-the-evolution-of-two-indicators-and-compare-countries#chart={%22indicator%22:%22desi_hc_bds%22,%22breakdown%22:%22ind_total%22,%22unit-measure%22:%22pc_ind%22,%22ref-area%22:[%22HR%22,%22EU%22]}

Fletcher, G., & Griffiths, M. (2020). Digital *transformation* during a lockdown. *International Journal of Information Management*, *55*, 102185. https://doi.org/10.1016/j.ijinfomgt.2020.102185

Fozo Attila, L., & Racsko, R. (2020). The possibilities of assessment of the schools' digital maturity. *Civil Szemle*, *17*(3), 93-+.

Gaftandzhieva, S., Doneva, R., & Docheva, M. (2022). *Digital Maturity Level of Bulgarian Primary and Secondary Schools*. 14.

Henderson, D., Munday, M., & Roberts, A. (2021). How to Measure Digitalization? A Critical *Evaluation* of Digital Maturity Models. *National Institute Economic Review*, *255*, 42–55. https://doi.org/10.1017/nie.2020.48

Ifenthaler, D., & Egloffstein, M. (2020). Development and Implementation of a Maturity Model of Digital Transformation. *TechTrends*, *64*(2), 302–309. http://link.springer.com/10.1007/s11528-019-00457-4

Jeladze, E., & Pata, K. (2018a). *The model of self-organization in digitally enhanced schools*. 17.

Jeladze, E., & Pata, K. (2018b). Smart, Digitally Enhanced Learning Ecosystems: Bottlenecks to *Sustainability* in Georgia. *Sustainability*, *10*(8), 2672. https://doi.org/10.3390/su10082672

Jugo, G., Balaban, I., Pezelj, M., & Begicevic Redjep, N. (2017). Development of a Model to *Assess* the Digitally Mature Schools in Croatia. In A. Tatnall & M. Webb (Eds.), *Tomorrow's Learning: Involving Everyone. Learning with and about Technologies and Computing* (Vol. 515, pp. 169–

178). Springer International Publishing. https://doi.org/10.1007/978-3-319-74310-3_19

Klacmer Calopa, M., Tomicic Pupek, K., & Begicevic Redep, N. (2018). *Methodology For Strategic Planning Of Ict Integration In Educational Institutions: Evidence From Primary And Secondary Schools In Croatia.* 5478–5484. https://doi.org/10.21125/edulearn.2018.1322

Kolic- Vehovec, S., Miletic, I., K. Maglica, B., & M. Dorcic, T. (2015). Information And *Communication* Technologies (Ict) In Education: A Case Study Of Two Elementary Schools. *7th International Conference on Education and New Learning Technologies*, 4949–4957.

Leonidova, G., Valiakhmetov, R., Baimurzina, G., & Babich, L. (2020). Problems and Prospects of Distance Learning in the Estimates Provided by Teachers and Schoolchildren's *Parents*. *Economic and Social Changes: Facts, Trends, Forecast / Экономические и Социальные Перемены: Факты, Тенденции, Прогноз*, *13*(4 (70)), 202–2019. https://doi.org/10.15838/esc.2020.4.70.12

Mekovec, R., Žajdela Hrustek, N., Kirinić, V., *Hutinski*, Ž., & Klačmer Čalopa, M. (2019). *Defining Competence Description Elements For School Principals' Digital Competence Framework: Croatian Example*. 5074–5081. https://doi.org/Managing integration of digital technologies in schools - principal role and competences

Pata, K., Tammets, K., Väljataga, T., Kori, K., Laanpere, M., & Rõbtsenkov, R. (2021). The Patterns of School Improvement in Digitally *Innovative* Schools. *Technology, Knowledge and Learning*. https://doi.org/10.1007/s10758-021-09514-5

Sotiriou, S., Riviou, K., Cherouvis, S., Chelioti, E., & Bogner, F. X. (2016). Introducing Large-Scale Innovation in Schools. *Journal of Science Education and Technology*, *25*(4), 541–549. https://doi.org/10.1007/s10956-016-9611-y

Svetec, B., & Divjak, B. (2021). Emergency Responses to the COVID-19 Crisis in Education: A Shift from Chaos to *Complexity*. *EDEN Conference Proceedings*, *1*, 513–523. https://doi.org/10.38069/edenconf-2021-ac0051

Thordsen, T., Murawski, M., & Bick, M. (2020). How to Measure Digitalization? A Critical Evaluation of Digital Maturity Models. In M. Hattingh, M. Matthee, H. Smuts, I. Pappas, Y. K. Dwivedi, & M. Mäntymäki (Eds.), *Responsible Design, Implementation and Use of Information and Communication Technology* (Vol. 12066, pp. 358–369). Springer International Publishing. https://doi.org/10.1007/978-3-030-44999-5_30

Towndrow, P., & Fareed, W. (2015). Growing in digital maturity: Students and their computers in an academic laptop programme in *Singapore*. *Asia Pacific Journal of Education*, *35*(4), 438–452. https://doi.org/Introducing Large-Scale Innovation in Schools

Vuk, B., Mornar, V., & Boticki, I. (2020). ICT in School Education in Croatia. In C.-K. Looi, H. Zhang, Y. Gao, & L. Wu (Eds.), *ICT in Education and Implications for the Belt and Road Initiative* (pp. 37–52). Springer Singapore. https://doi.org/10.1007/978-981-15-6157-3_3

Zugec, B., Balaban, I., & Divjak, B. (2018). *The Development Of An Instrument For Assessing Digital Maturity Of Schools.* 8557–8565. https://doi.org/10.21125/edulearn.2018.1990

Exploring the Usage of an Adaptive Learning System for Elementary School Math Classes

Gordana Borotić, Gordan Gledec, Tomislav Jagušt
Faculty of Electrical Engineering and Computing
University of Zagreb
Unska 3, 10000 Zagreb
{gordana.borotic, gordan.gledec, tomislav.jagust}@fer.hr

Abstract. *In this paper, we present the preliminary results of testing an adaptive learning system on tablet computers developed for mathematics classes in the lower grades of primary school.*
The developed system consists of two parts: the mobile application for repetition and revision of math exercises, and an administrative interface, through which teachers can create new tasks, monitor students' progress, and analyse the results of each student.
The adaptivity algorithm, based on various available information, assesses the current knowledge level of each student and populates a list of next tasks with ones of appropriate difficulty.
The preliminary study was conducted with a small class of third-grade students, and the results show that the adaptivity algorithm generally successfully assessed the students' knowledge level (the level which coincides with the teacher's assessment). Exceptions and border case issues are analysed in detail, and improvements to the algorithm are suggested.

Keywords. adaptive learning, adaptive systems, technology-enhanced learning, primary education

1 Introduction

Adaptive learning systems (ALSs) are software applications that monitor student behaviour and learning characteristics continuously dynamically adjusting the process in order to improve the learning experience of each student.
The adaptive educational systems were rapidly evolving in recent years since they greatly rely on computer science, programming, and technology advancement, which are the fields that have experienced fast and immense progress in the last 30 years (V. Shute et al., 2021; V. J. Shute & D, 2012; Van Schoors et al., 2021).
As stated in a paper by Knutov et al. (Knutov et al., 2009), there are three basic forms of adaptation: content adaptation, adaptive navigation, and adaptive presentation, which are supported by a number of different technologies and approaches, like data mining, context awareness or grouping. Park and Lee recognized and described several adaptive instructional models, with macro and micro-adaptive instructional models as the two most common types of learning adaptation today (Park & Lee, 2003). Similarly, Alshammari et. Al. (Alshammari et al., 2014) analyse three main perspectives (learner, domain and adaptation), while Paramythis and S. Loidl-Reisinger (Paramythis & Loidl-reisinger, 2004) look at the adaptive environments from 4 different perspectives: domain, learner, group, and adaptive model perspective, with each system implementing one or more of these models creating different levels of personalized experiences. On the basis of Paramythis and S. Loidl-Reisinger's work, Jianu and Vasilateanu proposed their adaptive gamified e-learning system, using adaptive questions and rewards in the form of levels and points (Jianu & Vasilateanu, 2017).
One of the learning activities often being adapted are task repetitions and practice, especially in subjects like Mathematics or Chemistry, where solving a large number of numerical tasks is a common way to gain knowledge and experience.
Solving tasks and practicing with the use of technology, e.g. with mobile applications could be more interesting to students than solving tasks with a pen and paper or in a workbook (Carruthers, 2015; Mendez et al., 2018). The students could take it more like a game and fun than a school obligation. On the other hand, solving tasks with the help of the software application offers many benefits. The application could, and indeed often does automatically check for validity and accuracy of the answer, and students get immediate feedback. If the task was solved incorrectly, the correct answer or a hint could be displayed. By showing the tasks one by one, students could keep a better focus, fixating only on the current task instead of a whole paper or workbook page with tens of similar tasks. Finally, the tasks in the workbooks are pre-written and cannot be adapted to different students and their needs. Some students need more time and practice to master the material, and others less. For students who know the material better, solving easy tasks could become boring and monotonous, while those with less knowledge could be discouraged by difficult tasks. For

this reason, we designed and developed an application that uses an adaptivity algorithm to estimate the current knowledge level of each student and adapt new set of tasks accordingly, so that the tasks are always challenging enough for all the students in the classroom.

2 Related Work

The field of Mathematics, especially the repetitions and task practicing, is often a target for the implementation of different adaptive learning systems. Recently, several authors have experimented in the area, designing and developing different adaptivity algorithms and computer systems that could have a potential to improve learning. Very recently Bang, Li, and Flynn (Bang et al., 2022) conducted a study to measure learning outcomes and engagement improvement of students using the adaptive learning application compared to students who did not use it. They noticed significant learning gains and skill improvement in students using their application, while teachers praised the application as a helpful learning resource and supplement to existing lessons and curricula. Rosen et al. (Rosen et al., 2018) implemented an adaptive learning system inside an existing MOOC. and found out that the adaptivity increased the learning gains, especially if the student was offered the topics in which the lowest level of mastery was shown. Jagušt et al. (Jagušt et al., 2018) conducted a study with different gamified learning activities, including an adaptive one. They measured improved performance levels in students using the adaptive application, which could lead to better learning outcomes, especially combined with improved interest and motivation, which has been achieved with added gamification elements. Zlatović, Balaban and Hutinski (Zlatović et al., 2022) created an online knowledge assessment system for university students, with adaptivity and personalization elements. The system proved to be efficient for achieving individual learning results.

Also, some commercially available applications offer different levels and types of adaptivity. For example, Knowji[1], a language learning platform, monitors the progress in word learning, and shows more often the words that are not yet mastered. It also tries to predict when a particular word will be forgotten and invites the user to repeat it. Drillster[2] learns about users from the answers and generates questions according to the area in which it considers the user to be weaker, encouraging them to practice weaker tasks more often.

On the other hand, some researchers reported negative or neutral results. For example, Jansen et al. (Jansen et al., 2016) researched the possibilities of task difficulty adaptation in math practice. They allowed students to set their estimated success rate and based on that provided tasks of different difficulty. The approach didn't show any benefit in the terms of skills improvement or self-belief, possibly because the whole process was not automatic, and students could be giving the wrong estimation on purpose, to gain more points later in the gamified application. Similarly, Papoušek and Pelánek (Papoušek & Pelánek, 2017) developed a system where users could adapt the question difficulty, and found only a small effect of student self-adjustments in their geography class experiment. In the end, they suggest and give some advice on how to automatically estimate student knowledge levels.

3 Methodology

In this case study, an adaptive system for the practice of mathematics tasks was designed, developed, and tested with a small group (N=12, 6 boys and 6 girls) of 3rd-grade lower primary school students.

3.1 The application

The developed learning system consists of a mobile application and a corresponding server that contains pre-prepared lessons. Each lesson contains one or more task groups, which in turn contain tasks the students have to solve. Task groups have the function of grouping tasks by certain similarities. There were three types of tasks - multiple-choice including four possible answers, true-false questions, and tasks that required input of the numerical solution. Each task has an estimated difficulty level, which can range from 1 to 3 (1 - easy, 2 - medium, 3 - hard). The difficulty levels were estimated by teachers (Figure 1).

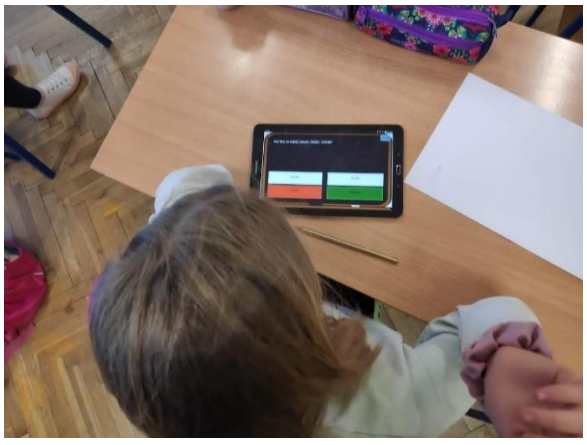


Figure 1. Student solving an ABC question on a tablet computer

When a student starts an application, starting screen of the app interface offers a list of existing lessons.

[1] https://www.knowji.com/

[2] https://drillster.com/

After the student selects a lesson, the adaptivity algorithm selects a set of 15 tasks for solving. Tasks are displayed one by one on the screen. After the student enters a numeric answer or selects one of the offered answers, the answer is logged on the server, the correctness is checked, and the feedback is displayed to the student. If the student answered incorrectly, the correct answer is displayed.

After the student has solved all 15 tasks, the success of solving the lesson is evaluated. Users receive feedback in a form of a simple 1-5 grade, calculated as a number of correct answers multiplied by the difficulty level of each task, then divided by a maximum possible score, and finally scaled to the 1-5 interval.

3.2. The adaptivity algorithm

Based on the existing data from the system log, and similar to studies described in (Balaban, 2015; Zlatović et al., 2022), to the built-in adaptivity algorithm estimates the student's knowledge of the selected lesson, to make it easier for students to master and permanently adopt the learning material. Based on the estimated knowledge level, the algorithm selects future tasks (in sets of 15 tasks) for each student. The more the students answer correctly, the adaptivity algorithm will generate more and more difficult tasks, and, in this way, students will gradually improve their knowledge. The algorithm will more often choose tasks that student had problems with solving in the past (or similar ones) for the student to eventually master. Also, over the time people tend to forget the learned material, and if they don't refresh their knowledge regularly, the level of knowledge may fall. In this case, after the algorithm detects a knowledge drop, a set of easier tasks is generated for student, to refresh the material as soon as possible.

In the beginning, while there is no recorded log data, the algorithm does not have information about the student's knowledge. In that case, it selects a set of tasks consisting of 5 tasks of each difficulty level. It selects tasks from different task groups to make the generated set of tasks as diverse as possible. After the student solved the lesson for the first time, the algorithm has some data to work on and can assess the knowledge level of that student.

If the student previously solved the selected lesson, the recorded information is used when the algorithm generates a set of new tasks, with special emphasis on the tasks from the last attempt. Based on the result (scaled to 0 – 3 interval) of the last attempt, the number of future tasks with each difficulty level is calculated, as shown in table 1. The generated set always contains the tasks of all difficulty levels, to minimise the possibility of wrong student knowledge level assessment by the algorithm. With each new result saved in the data log, the algorithm adjusts students' score, which in turn, changes a list of future tasks (Table 1).

Table 1. Number of selected tasks with each difficulty level based on the calculated student knowledge level

Knowledge level	Easy (1)	Medium (2)	Hard (3)
[0,1]	8	5	2
<1,2]	4	7	4
<2,3]	2	5	8

After determining the number of tasks of each difficulty level, the algorithm selects them randomly from appropriate task groups, giving priority to the tasks that the student didn't try to solve yet, or didn't solve correctly. For each task that was solved incorrectly in the previous attempt, an equivalent is found by taking a task of the same difficulty from the same group of tasks (similar tasks are in the same group of tasks).

4 Results

To test the algorithm and detect possible issues or anomalies, the experiment with a group of 3rd-grade students was conducted, and application usage data was collected in the server log, and analysed.

The future tasks for each student are selected based on several criteria like the accuracy of task solving, assumed level of knowledge, task characteristics etc. When selecting tasks with certain characteristics (task difficulty and group of tasks), priority is given to tasks that the student has not yet solved, and then the tasks that the student solved incorrectly. The described behaviour of choosing tasks that have less accuracy of solving can be seen in Figure 2. The results are presented for a specific student and the tasks from one lesson - the algorithm more frequently selects tasks that the student has not yet mastered.

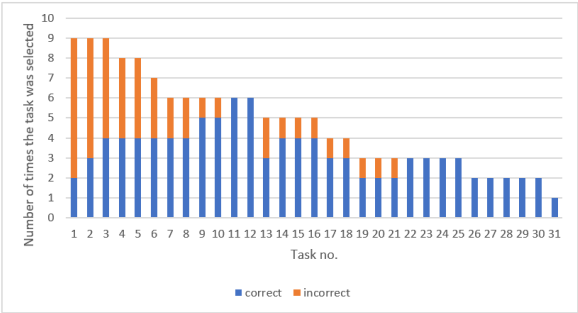


Figure 2. The frequency of selection of individual tasks in relation to the correctness

When observing the adaptability of the algorithm, we will first observe one student and his attempts to solve a specific lesson because the adaptivity algorithm evaluates students' knowledge for each lesson separately and generates sets of tasks based on the assessed knowledge. We observed the same student

and the same lesson as in the previous example, but this behaviour is also manifested in other students and lessons (Figure 3).

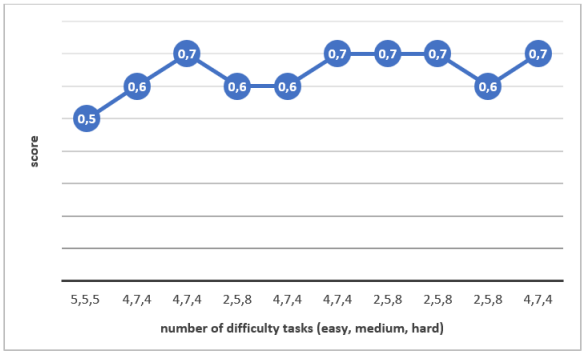


Figure 3. The change in the generated task list as a result of student knowledge level change

The x-axis shows a number of tasks of individual difficulty that have been selected for student at the beginning of each round. The first value on the x-axis represents the very first attempt of the student to solve the lesson. At that point, the algorithm has no information of the student's knowledge for the selected lesson, so it generates an equal number of tasks of each difficulty. The student solved the selected tasks with a score of 0.5. Taking this result into account, the next set of tasks is generated for him (4 easy, 7 medium, and 4 hard tasks). From the picture, one can notice that the algorithm is constantly adapting the list of future tasks. If the student has solved the previous set of tasks well, the algorithm picks more difficult tasks for the next round, and vice versa.

After observing the work of the algorithm individually on students regarding the accuracy of task solving, we wanted to check the global picture of knowledge assessment. We asked the class teacher to rank students according to their assessment of math knowledge. Using the arithmetic mean of the results of all solutions, we calculated how the algorithm evaluates the knowledge of each student individually.

In Table 2, students are ranked within the class according to the assessed knowledge that was calculated by the algorithm, while the last column shows students' rank in a class according to the teacher's assessment. In the second column, the assessed knowledge from algorithm is visible. We can notice that there is a small difference in the assessed knowledge between some students, therefore a small shift in the ranking list with regard to the teacher's assessment should not be a problem.

Generally, students who were rated as "above class average" in math by the teacher were also put in the upper part of the list by the algorithm and vice versa. Nevertheless, there were some deviations. In table 2 the student who was rated 4th by the algorithm was ranked last (12th) by the teacher, which is a "mismatch" of 8 places. A deeper analysis of the log data revealed the reason - that the student was solving only one lesson repeatedly, until solved it perfectly, and did not try to solve any other lesson.

Table 2. Students ordered by their knowledge level by the adaptivity algorithm and by their teacher

Student	knowledge level by the algorithm	Order by algorithm	Order by teacher	Diff.
s1	4.57	1	4	3
s2	4.55	2	2	0
s3	4.50	3	3	0
s4	4.38	4	12	8
s5	4.36	5	1	-4
s6	4.28	6	5	-1
s7	4.21	7	6	-1
s8	3.90	8	10	2
s9	3.71	9	9	0
s10	3.50	10	8	-2
s11	3.50	11	7	-4
s12	3.33	12	11	-1

For that reason, the algorithm, which calculates an arithmetic mean of all task-solving activities for each lesson, calculated the student's score as very high (although, in total it was not high enough for the first place). To improve the application performance in this and similar border cases, some modifications to the algorithm should be done, for example, it could calculate student score differently if student solved only one lesson, or repeatedly solved a small number of lessons in a short period of time. Another option is to take the re-solving of the same task into the calculation with different (smaller) weighting factor, especially if it was done shortly after the student tried to solve that task for the first time. The second biggest mismatch between students' knowledge level and teacher's list, and in the opposite direction, was with student s5, which is, by the teacher, "the best mathematician" in the class. Here the difference between algorithm and teacher is 4 places, and the knowledge level of students in front of him is not much higher than the knowledge level of s5. The log analysis revealed the fact that s5 solved the most tasks in the whole class, which, in turn, could be a reason for somewhat "lower" performance. The student was trying to solve as many tasks as possible in the given time and made more mistakes because of this self-imposed time pressure.

For student s11, although the mismatch was also 4, there is no discussion, because the teacher was not so confident in the order of students in the middle of the list, as she was with top students or the students whose knowledge level is lagging significantly behind the class average.

5 Conclusion and Future Work

In this paper we presented a mobile educational application for math task practice and repetition, that uses the adaptivity to continuously modify a list of tasks that are offered to the student, enhance the quality of learning, improve the amount of learned material and minimise the potential losses by resetting the student's forgetting curve. An experiment conducted on a group of twelve 3rd grade students showed promising results, the system managed to sort the students in a similar way as their teacher, with some anomalies which were discussed in more detail. Also, some improvements to the algorithm were proposed.

This case study adds to the existing corpus of knowledge in the area of adaptive learning systems, giving researchers in the field new implementation case, and teachers another practice example that has a potential to reduce teacher's workload, at the same time improving the personalization and individualization of the teaching process.

There are some noteworthy limitations to this study. Two main issues are a small sample size and only one experiment where the data was collected. Also, the starting task difficulty levels were estimated by teachers, which could introduce some difficulty bias (e.g. teachers could consider some tasks easy, but in reality students could have problems solving them). Although the algorithm included the option to adjust the task difficulty based on the previous student results, since this was the first time the application was used, there was not enough history data to use the difficulty adjustment feature.

In future, we plan to test the application on a much bigger number of students, for a longer period of time, and adjust the task difficulty level on the basis of previous results. This way more usage data will be collected, which would allow more precise student knowledge level assessment.

References

Alshammari, M., Anane, R., & Hendley, R. J. (2014). Adaptivity in E-learning systems. *Proceedings - 2014 8th International Conference on Complex, Intelligent and Software Intensive Systems, CISIS 2014*, 79–86. https://doi.org/10.1109/CISIS.2014.12

Balaban, I. (2015). Personalizing questions using adaptive online knowledge assessment. *ELearning 2015-6th International Conference on e-Learning, March*.

Bang, H. J., Li, L., & Flynn, K. (2022). Efficacy of an Adaptive Game-Based Math Learning App to Support Personalized Learning and Improve Early Elementary School Students' Learning. *Early Childhood Education Journal*, *0123456789*. https://doi.org/10.1007/s10643-022-01332-3

Carruthers, C. (2015). *An Interactive Learner-Centered Classroom: Successful Use of Tablet Computing and Dyknow Software in Foundational Mathematics*. 315–321. https://doi.org/10.1007/978-3-319-15594-4_33

Jagušt, T., Botički, I., & So, H.-J. (2018). Examining competitive, collaborative and adaptive gamification in young learners' math learning. *Computers & Education*, *125*(June), 444–457. https://doi.org/10.1016/j.compedu.2018.06.022

Jansen, B. R. J., Hofman, A. D., Savi, A., Visser, I., & van der Maas, H. L. J. (2016). Self-adapting the success rate when practicing math. *Learning and Individual Differences*, *51*(May), 1–10. https://doi.org/10.1016/j.lindif.2016.08.027

Jianu, E. M., & Vasilateanu, A. (2017). Designing of an e-learning system using adaptivity and gamification. *2017 IEEE International Symposium on Systems Engineering, ISSE 2017 - Proceedings*, 16–19. https://doi.org/10.1109/SysEng.2017.8088270

Knutov, E., De Bra, P., & Pechenizkiy, M. (2009). AH 12 years later: A comprehensive survey of adaptive hypermedia methods and techniques. *New Review of Hypermedia and Multimedia*, *15*(1), 5–38. https://doi.org/10.1080/13614560902801608

Mendez, D., Mendez, M., & Anguita, J. (2018). Motivation of 14 year-old students using tablets, compared to those using textbooks and workbooks. *International Journal of Interactive Mobile Technologies*, *12*(4), 86–96. https://doi.org/10.3991/ijim.v12i4.9203

Papoušek, J., & Pelánek, R. (2017). Evaluation of learners' adjustment of question difficulty in adaptive practice of facts. *UMAP 2017 - Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization*, 379–380. https://doi.org/10.1145/3079628.3079642

Paramythis, A., & Loidl-reisinger, S. (2004). Adaptive Learning Environments and e-Learning Standards. *Learning*, *2*(1), 181–194.

Park, O., & Lee, J. (2003). Adaptive instructional systems. *Handbook of Research for Educational Communications and Technology*, *1911*, 651–684. http://www.aect.org/edtech/ed1/25.pdf

Rosen, Y., Munson, L., Lopez, G., Rushkin, I., Ang, A., Tingley, D., Rubin, R., & Weber, G. (2018). The effects of adaptive learning in a massive open online course on learners' skill development. *Proceedings of the 5th Annual ACM Conference on Learning at Scale, L at S 2018*. https://doi.org/10.1145/3231644.3231651

Shute, V. J., & D, Z.-R. (2012). Adaptive educational

systems. *Adaptive Technologies for Training and Education*, *7*(27), 7–27.

Shute, V., Rahimi, S., Smith, G., Ke, F., Almond, R., Dai, C., Kuba, R., Liu, Z., Yang, X., & Sun, C. (2021). Maximizing learning without sacrificing the fun: Stealth assessment, adaptivity and learning supports in educational games. *Journal of Computer Assisted Learning*, *37*(1), 127–141. https://doi.org/10.1111/jcal.12473

Van Schoors, R., Elen, J., Raes, A., & Depaepe, F. (2021). An overview of 25 years of research on digital personalised learning in primary and secondary education: A systematic review of conceptual and methodological trends. *British Journal of Educational Technology*, *52*(5), 1798–1822. https://doi.org/10.1111/bjet.13148

Zlatović, M., Balaban, I., & Hutinski, Ž. (2022). A Model of the Continual Adaptive Online Knowledge Assessment System. In *E-Learning and Digital Education in the Twenty-First Century* (p. 13). IntechOpen. https://doi.org/10.5772/intechopen.95295

A Preliminary Investigation of the Effects of Communication Quality and Use of Communication Platforms on Students' Perception of Virtual Team Performance

Antonela Čižmešija, Goran Bubaš, Igor Balaban
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{acizmesi, gbubas, ibalaban}@foi.unizg.hr

Abstract. *The central topic of this research paper is virtual team performance in an academic context in which students use communication platforms to perform their tasks. The Task-Technology Fit model and various aspects of communication quality and usability of communication tools were used as a theoretical framework as they are considered to be important for the virtual team performance, but still understudied in that context. This paper presents a preliminary evaluation of a measurement instrument that was used to analyse processes in virtual teams in a university setting. Participants were university graduates in software engineering (N=64) who collaborated in virtual teams and extensively used various communication applications for their task related activities. The results confirmed the reliability and validity of the scales of our measurement instrument and revealed that communication quality was the most influential and statistically significant predictor of perceived performance in virtual teams.*

Keywords. virtual teams, task-technology fit theory, communication quality, communication tools, usability, team performance, instrument validation.

1 Introduction

A recent extensive review of literature on virtual teams (Garro et al., 2020) revealed that in the last 25 years most of the publications have been related to the topics (studied constructs) of *performance and communication* in virtual teams, including *communication technology* (equipment) that was used. A newer study (Muszyńska, 2021) focused on publications on virtual teams over the last two decades and found that the most frequent topic of the research papers was related to the *methods and tools for communication* in virtual teams. According to this study, other frequent topics were associated with the factors influencing team communication, effect of communication on team performance, as well as communication challenges, practices and patterns.

After the onset of the COVID-19 pandemic there was an increased pressure in organizations to shift into virtual teamwork, with a possible effect of gradual improvement of team action processes and better adaptation of teams to virtual work over time (Klonek et al., 2022). The adaptation of team work to the virtual environment and online collaboration technology during the COVID-19 pandemic can be analysed regarding *task* interactions, *process* interactions, and *relationship* interactions. However, those interaction domains are almost always reliant on the use of appropriate digital tools for synchronous and asynchronous communication (as reported in: Whillans, 2021). In addition, it is important to identify various human and organizational factors that affect the productivity of software teams during the COVID-19 pandemic like communication, collaboration, organization of teams, motivation and work environment (see for instance: Bezerra et al., 2020). In fact, abundant advice is offered in literature for the rapid deployment of virtual teams when lockdown and social distancing are used as a countermeasure for the waves of COVID-19 pandemic (for example see: Kilcullen et al., 2021).

2 Shift in communication in higher education

The COVID-19 pandemic imposed numerous challenges on teachers and students, including keeping the course related communication effective. A notable breakthrough in teaching and learning practice was achieved by an intensified use of web conferencing systems, for example, *BigBlueButton* (Čižmešija & Bubaš, 2020), as well as communication platforms like *Telegram* to support effective educational interactions

(Aladsani, 2021). On the other hand, the practices of using communication platforms among students for collaboration in teams for performing tasks, projects, and other academic activities were unquestionably influenced by the pandemic. Consequently, over the past two years, a significant shift has been made from the students' use of classic instant messaging tools, whose main purpose is the exchange of messages and photos, to more advanced collaboration tools. A recent study (Gonçalves, Sousa, & Pereira, 2020) revealed that students favoured communication platforms (e.g., Zoom or Google Hangouts) over e-mails, smartphone chat, or e-learning platforms as means of communication, with a preference for richer communication channels for facilitating the learning processes and group work.

One review of scientific literature on *virtual teams in higher education* (Jony & Serradell-López, 2019) indicated that some of the popular topics in related research papers are denoted by keywords like *team performance, success, effectiveness* and *satisfaction*. However, only several research papers were found that used the Task-Technology Fit (TTF) model to investigate virtual teamwork in higher education. For instance, a representative paper by Fuller & Dennis (2009) included students as participants in a study of TTF variables with a focus on prediction of team performance. Another study (Aiken et al., 2013) investigated how the task-technology fit mediates between knowledge sharing and team satisfaction with the use of technology in form of the WebCT learning management system.

In literature, perceived communication quality is the construct that comprehensively maps different interactions and actions in the academic context in which students working in virtual teams use communication platforms to perform various tasks. Common sub-dimensions of communication quality include information elaboration, knowledge sharing, openness of communication and general information sharing (Marlow et al., 2018). Since communication quality has a greater influence on perceived virtual team performance than communication frequency, and given that its sub-dimensions are connected to various aspects of interactions among students collaborating in virtual teams, the authors of this research paper found it worthwhile to explore its role in the academic context in conjunction with the TTF model.

3 Theoretical framework

One approach to investigating virtual team performance both theoretically and empirically is to use the Task-Technology Fit (TTF) model (see: Goodhue & Thompson, 1995; Goodhue, 1998). This model assumes that information systems (ISs) can be instrumentally used for a task or a collection of tasks and that the correspondence between task requirements and technology/IS functionalities ("task-technology fit") influences performance. In other words, the characteristics of tasks "moderate" the association between characteristics of ISs and how the users evaluate them. Information technology can thus have a positive impact on performance if it is utilized and if its features are in correspondence with the requirements of the tasks. According to a recent review of literature on TTF (Spies et al., 2020), the last decade and a half has seen a growing interest in the ways to apply this model in practice, from the use of various technologies (mobile, ISs, software, communication etc.) to its application in diverse fields (from healthcare, education and finance, to software support and other areas). A newer literature review (Hidayat et al., 2021) on the research application of TTF confirmed its use in diverse domains (education, social media, business and management, finance and banking etc.), while also evidencing continued interest of researchers for TTF over the past several years. Another relevant finding of this literature review was that the most frequent antecedent variables in TTF related scientific research were *Task Characteristics* and *Technology Characteristics*, while the most numerous dependent variables were *Perception of Usefulness* of technology, *Performance*, *Satisfaction* and *Perception of Ease of Use*. Still, it must be noted that virtual teamwork has both advantages and potential problems, as documented in a comprehensive review paper by Morrison-Smith and Ruiz (2020).

In our paper the constructs of the Task-Technology Fit model are used for the evaluation and improvement of existing assessment scales developed by other authors for the measurement of TTF related variables, as well as for the construction of new scales for this purpose. This is performed in the empirical context of university students' virtual teams and their use of communication and collaboration technologies (e.g. *WhatsApp, Discord, Facebook Messenger, MS Teams, Slack, Zoom*). After evaluation of the scales that were designed to measure the TTF related variables in the context of virtual teams, the intercorrelation of those variables was investigated, as well as the predictors of the variable *Perceived Team Performance*.

4 Research problems and goals

The *first and main* goal of our research and pilot study was to evaluate the assessment scales for the measurement of task-technology fit (TTF) constructs in the context of students' virtual teams. The *second goal* was to examine the relationships between the following variables: (a) quality of communication, (b) characteristics of communication platforms, (c) perception of the task-technology fit and acceptance of technology in the context of virtual team tasks, as well as (d) perceived performance in a virtual team. The *third goal* was to examine the students' usage of communication platforms in the context of performing virtual team tasks. In other words, the aims of our study

were to investigate the potential influence of the aforementioned variables on the perceived students' performance in virtual teams and also to test the measurement instrument that was designed for that purpose, including testing of its reliability and usefulness so that it could be applied in further research of virtual teams in other contexts.

According to the goals of our study and the *Task-Technology-Fit* theoretical model, the following research questions were defined:

RQ1: Which platforms for communication are most frequently used by the respondents (virtual teams) in our study?

RQ2: What is the reliability of the assessment scales that measure the corresponding constructs of the *Task-Technology Fit* theoretical model?

RQ3: What are the predictors of the TTF dependent variable *Perceived Performance* in the context of virtual teams?

5 Methodology

5.1 Data collection instruments

Possible effects of communication quality, characteristics of communication platforms, task-technology fit and evaluations of virtual team tasks on students' performance in virtual teams have not yet been investigated in scientific literature. Therefore, the first step in our research was to develop a measurement instrument with assessment scales designed for the specific purpose of measuring variables associated with the TTF model as a theoretical background. The first part of our survey included demographic questions regarding age, gender and year of study. The remaining sections of our survey included assessment scales specially designed to measure the constructs related to the TTF model and to collect students' responses regarding their frequency of use of various communication platforms in virtual teams. For the responses to the items of the assessment scales a five-point Likert-type scale in the range from *"1 – Totally inaccurate"* to *"5 – Absolutely accurate"* was used.

As was mentioned earlier, the main research constructs in our study were taken from the *Task-Technology-Fit model* (Goodhue & Thompson, 1995), which was used as a starting point for the operationalization of variables. It is important to note that the items used for the measurement of the construct-related variables were adapted to the context of students' virtual teams and the academic environment in which the research was conducted. Furthermore, some of the items were adapted from the assessment scales used in related research by other authors. Each of the variables (constructs) in our study is explained in more detail in the continuation of this section of our paper and sample items for their measurement are provided for illustration purpose.

Communication Quality is the level at which the content of communication is distributed and shared in a virtual team (Daim et al., 2012). This variable was measured using the following three relevant dimensions of communication quality identified in literature: (1) openness of communication, (2) knowledge sharing, and (3) information elaboration (see: Marlow et al., 2018). *Communication openness* refers to the level at which individuals in a team receive ideas, values, opinions, and emotions from others and share them with each other (Carlson, Carlson, Hunter, Vaughn, & George, 2013). An example of an item for the measurement of communication openness in our study is: *"Virtual team members were able to easily ask for advice from any other team member".* In the context of communication quality in a virtual team, *knowledge sharing* refers to virtual team members' willingness to share their functional experience or knowledge acquired through education to others within a team. Accordingly, the statement *"Members of my virtual team shared their previous experiences and examples of good practice within the team in performing related tasks"* is an illustration of a questionnaire item for the sub-dimension knowledge sharing. *Information elaboration* is associated with the complexity and assimilation of information to process the team/individual ideas comprehensively. This is a representative item for its measurement in our study: *"While working on a task or project in my virtual team, members sought to gather and use all available information."*

Communication Application Characteristics are related to the features of tools that individuals use to execute their tasks (Cane & McCarthy, 2009). In our study this multi-faceted construct was measured by the following subscales: *teamwork efficiency, perceived ease of use* and *reliability of communication platforms*. An example item used for measuring the reliability of communication applications is *"During the work of my virtual team with communication technologies, technical difficulties and technology-related downtimes were very rare."*

The ***Task-Technology Fit & Acceptance*** construct can be described as actions that team members accomplish and how these actions rely on specific functionalities of information technology (Hauder, Fiedler, Matthes, & Wüst, 2013).*"For all important activities and tasks in my virtual team, we were able to find and use a suitable functionality in the selected communication platforms"* and *"Generally speaking, members of my virtual team accepted the selected communication platforms well in relation to our mutual activities and tasks"* were two of the items used in the assessment scale for the measurement of the correspondence of tasks with technology choice and acceptance of selected technology for the tasks. It should be noted that three items that refer to *acceptance of technology for tasks at hand* were included in this assessment scale.

The *dependent variable* in our study was ***Perceived Performance*** in a virtual team. In the TTF model, an individual's (e.g. student's) performance refers to the accomplishment of a portfolio of tasks by an individual (Goodhue & Thompson, 1995). Two examples of sample items in our study that were formulated according to the measures of team effectiveness identified by Hertel et al. (2004) were: *"The members of my virtual team were effective in achieving team goals."* and *"In my virtual team, members typically completed work tasks within agreed deadlines."*

5.2 Respondents and data collection

The initial convenience sample of respondents in our study were 70 graduate IT students at a higher education institution in Croatia. The survey was anonymous and administered in an online format with the *Google Forms* online questionnaire tool. The students' participation in the online survey was on a *voluntary* basis. Consequently, only 70 of the 113 students who were enrolled in the course filled out the online survey.

To check the respondents' overall determination and honesty, at the end of the survey the following item was included: *"How committed, focused, and honest were you in answering the questions in this survey?"*. In case of answers to this question which indicated an unserious approach to the survey, the respondents' data was excluded from further analysis. Also, the responses to survey items which indicated the *general frequency of communication platform use*, as well as *how much these platforms were used on a daily and weekly basis*, were carefully reviewed. Students who responded to the related questions with "*Never/I don't use it"* or *"Very rarely"* were omitted from the data set before statistical analyses because our study was focused on the evaluations and perceptions of students who had sufficient experience with interacting with team members using communication platforms like Facebook Messenger, WhatsApp, MS Teams, Slack, Zoom, Jira, Skype, and Google Hangouts while performing their tasks in virtual teams.

After data cleaning and removal of records with responses that were not found to be suitable to be included in statistical analysis, *the final convenience sample consisted of 64 respondents* (41 male and 23 female). Most of the subjects, 53.1% of them, were aged between 21-22 years, followed by 42.2% of them aged between 23-24, and the rest, who were 25 or more years old. Also, 60 subjects were in their first year of graduate study, and 4 subjects were in the second (final) year of their graduate education.

It must be noted that before the data collection for our study the approval of the *Ethical Committee* was obtained.

5.3 Respondents' virtual team activities

The subjects in our study were students enrolled in a course on *technologically mediated communication* in the winter semester of the 2021/2022 academic year and the study was conducted during January 2022, after the students had finished almost all course activities. Due to the COVID-19 related epidemiological measures at the time when the teaching in this course was performed, the students in the course were attending fully synchronous online video lectures from the beginning until the end of the semester, as well as laboratory exercises in various formats: (a) synchronous online sessions, (b) asynchronous online sessions, (c) mixed on-site sessions in the computer lab and simultaneously streamed online via a videoconferencing tool, as well as (d) fully on-site sessions in the computer laboratory. The use of the previously mentioned formats of laboratory exercises depended on the epidemiological situation during the semester. Because of the anti-COVID measures, in most cases and for the majority of students the laboratory sessions were performed with a significant degree of the online component (either as fully online synchronous exercises or as on-site exercises combined with simultaneous live stream exercises). Students were also able to choose to attend one fully online study group of laboratory exercises.

A part of students' practical assignments in this course was to work on a team project that included a search of scientific literature on a specific topic, defining the topic/theme of their project, performing a survey on the topic of their project, and oral presentation of project results to other students with online presentation software like *Prezi* or *Emaze*, as well as the design of an asynchronous multimedia presentation of the project using the Mahara e-portfolio system. Four weeks after the beginning of the semester the students were organized in teams with 2-3 members who worked together on the project topic they selected.

During the period of more than 10 weeks, the members of students' teams used various communication platforms of their choice in order to perform diverse tasks that were assigned to them at different project stages. The context of the COVID-19 pandemic and mixed format of their laboratory exercises forced most of the students' teams to collaborate "virtually" in an online communication environment. The fact that project assignments were performed as an out-of-class learning activity, with most of the students being physically away from lecture halls, classrooms and laboratories at their college caused the dispersion of team members, who were forced to predominantly collaborate virtually via communication platforms to execute and complete the project assignment. Therefore, it can be considered that the project tasks that demanded considerable time and effort to complete were in fact performed by virtual student teams. In other words, the virtual teams consisted of groups of students that collaborated remotely on a joint task for the duration of this task (a

course project). The teams were assembled on a voluntary basis and spontaneously, without any direction of the instructor regarding specific team membership and composition.

6 Results

6.1 Reliability analysis

The verification of internal consistency or *Cronbach's alpha reliability* is an important phase in the development of an instrument that consists of self-assessment scales designed to measure various constructs or research variables. Cronbach's alpha is an indicator to which extent a proposed set of items of an assessment scale measure a single construct or variable (Ursachi, Horodnic, & Zait, 2015). In *Table 1*, values of Cronbach's alpha reliability coefficients for the assessment scales that were used in our study are displayed, as well as the number of items in each of the scales. The results of this analysis indicated that the original sets of items in almost all of the assessment scales that measured the corresponding constructs were of appropriate reliability because the Cronbach's alpha values were above the 0.7 (this is considered as acceptable in the educational research context; see: Taber, 2018). Only in the case of the assessment scale that measures the construct ***Task-Technology Fit & Acceptance*** one item was omitted to achieve a higher value of Cronbach's alpha coefficient.

Table 1. Cronbach's alpha values for constructs; labels for (total) scales are written in boldface and for their subscales in italics (N=64)

Variable name	Number of items in scale	Cronbach's alpha coefficient
Perceived Performance	**5**	**0.743**
Communication Quality	**15**	**0.851**
- Subscale *Openness of Communication*	6	0.748
- Subscale *Knowledge Sharing*	5	0.743
- Subscale *Information Elaboration*	4	0.712
Task-Technology Fit & Acceptance	**6**	**0.759**
Communication Application Characteristics	**9**	**0.778**
- Subscale *Teamwork Efficiency*	3	0.716
- Subscale *Perceived Ease of Use*	3	0.781
- Subscale *Reliability*	3	0.783

6.2 Students' use of communication platforms for virtual team activities

In the context of our study it was important to explore students' general usage of communication platforms for their activities in virtual teams. According to the analysed data, 29 participants in our study (45.3%) stated that they used communication platforms for virtual team activities *"very often"*, followed by 22 of them (34.4%) who responded with *"often"* and 13 (20.3%) who responded with *"neither often, nor rarely"*. To better understand the frequency of use of communication applications for team tasks by students, the responses related to the time that they spend daily on using them were also analyzed. In total, 62.6% of the subjects in our study reported that their use of communication applications that can be utilized for virtual team activities was approximately 2 or 3 hours per day, 20.3% of them reported that they allocated about 1 hour a day to such applications, and the remaining 17.1% reported using them 4 or 5 hours per day. Besides the frequency of usage, the use of specific communication applications by students was also investigated. Their most frequent choice was *WhatsApp* (43.75%), followed by *Discord* (31.25%). *Facebook Messenger* was the first option for 9.38% and Zoom for 4.69% of the students in our sample, while other communication tools were given as the preferred choice by 1.56% to 3.13% of respondents.

It is important to emphasize, however, that most of the students reported using at least three or four different tools for communication purposes. In further analysis of the multiple-choice answers regarding the use of communication platforms (checkboxes question type), it was found that 25 respondents (39.06%) stated that they also use *Jira*, which is a common communication tool in software engineering.

In their responses to an open-ended question regarding their *most common conversation topics* while using communication platforms for work in virtual teams, the students pointed out scheduling and planning activities and tasks for the virtual team, monitoring the workflow and progress of team members, coordination of activities and deadlines, discussing how to solve tasks and sending additional materials such as pictures (screenshots) and videos that were useful for learning that was needed for task accomplishment. Apart from using communication platforms for virtual team activities in our study, students also used communication tools for other assignments in different courses that were related to software development, like writing a project plan, or preparing seminar papers and presentations.

To explore whether the main constructs and corresponding survey questions in our study were related to their real-life academic activities in virtual teams that students performed to meet course requirements, in the last section of the survey the question *"How many questions in this survey concerned you and your virtual team?"* was included. This was done to check if students had adequate experience for working in a virtual team to properly respond to the content of the survey questions. For the majority of students from the research sample in our study, the questions in the survey reflected a lot of their virtual team experiences (47 respondents, or 73.44%)

or were at least partly related to what they experienced in virtual teamwork (16 respondents, or 25%).

6.3 Correlation and regression analysis

6.3.1 Correlation analysis

To examine the relationships between the variables in our study in the context of virtual team performance and the use of communication platforms for virtual teamwork, we calculated the Spearman's rank correlation coefficient (r_s) between the total scores of assessment scales used to measure different TTF constructs. The Spearman's rank correlation coefficient (Spearman's rho) is suitable for nonparametric data distribution and measures the strength of the monotone association between two variables. This coefficient is not a measure of the linear relationship between two variables since the data are transformed into ranks before calculating the coefficient (Hauke & Kossowski, 2011). It can be interpreted as follows: a perfect negative or positive monotonic relationship between variables occurs in the case when $\rho = -1$ or $+1$. On the other hand, if $\rho = 0$, there is no association between variables.

In *Table 2* the values of Spearman's rho coefficients between the following variables are displayed: *Perceived Performance*, *Communication Quality*, *Task-Technology Fit & Acceptance,* and *Communication Application Characteristics*. It must be noted that for the purpose of correlation analysis the data collected with the *Communication Quality* scale was used as a summative score of its three subscales *Openness of Communication, Knowledge Sharing* and *Information Elaboration.*

Table 2. Spearman's rank correlation coefficient values for key TTF variables (N=64)

	Perceived Performance	Communication Quality	Task-Technology Fit & Acceptance
Communication Quality	.632**	-	-
Task-Technology Fit & Acceptance	.291**	.345**	-
Communication Application Characteristics	.421**	.364**	.605**

**Correlation is significant at the 0.01 level (2-tailed).

The results of the correlation analysis revealed that all of the key variables have positive intercorrelations that are statistically significant at the 0.01 level. The strongest correlations were identified in the case of the following variable pairs:

- There was a positive and strong correlation between *Perceived Performance* and *Communication Quality* ($r_s = .632$, $p < .001$).
- There was a positive and strong correlation between *Task-Technology Fit & Acceptance* and *Communication Application Characteristics* ($r_s = .605$, $p < .001$).
- There was a positive and moderate correlation between *Perceived Performance* and *Communication Application Characteristics* ($r_s = .421$, $p < .001$).

In the case of other variables, the correlations presented in *Table 2* can be qualified as positive and statistically significant but weak since their values were in the range between 0.20 and 0.39. The lowest positive (only weak) correlation was revealed between *Perceived Performance*, on the one side, and *Task-Technology Fit & Acceptance* ($r_s = .291$, $p < .001$).

6.3.2 Regression analysis

To determine which of the independent variables *Communication Quality, Task-Technology Fit & Acceptance* and *Communication Application Characteristics* had greatest association with the students' *Perceived Performance* in a virtual team, as a dependent variable, regression analysis was performed. The results of the regression analysis that are presented in *Table 3* reveal that the value of R^2=0.390 (coefficient of determination), i.e. that 39% of the variation in the *dependent variable* is shared or can be predicted from *independent variables* included in the regression analysis.

Table 3. Model summary of regression analysis – dependent variable *Perceived Performance* (N=64)

Model	R	R square	Adjusted R square	Std. error of the estimate
1	.625[a]	.390	.360	.391

a. Predictors: (Constant), *Communication Quality, Task-Technology Fit & Acceptance, Communication Applications Characteristics*

The next step in the regression analysis was to explore which of the included variables can be considered as the greatest statistically significant predictor of the criterion (dependent) variable *Perceived Performance*. According to the data displayed in *Table 4*, only in the case of variable *Communication Quality* the p-value was < 0.01. Therefore, *Communication Quality* could be taken into consideration as the most influential statistically significant independent predictor of *Perceived Performance* in a virtual team. For variables *Task-Technology Fit & Acceptance* and *Communication Application Characteristics,* the p-value was above 0.05, so they were not confirmed as statistically significant predictors in the regression model for the observed dependent variable *Perceived Performance* that is displayed in *Table 4*.

Table 4. Statistical significance of independent variables as predictors of *Perceived Performance* (N=64)

Model	Unstandardized coefficients		Standardized coefficients	t	Sig.
	B	Std. error	Beta		
(Constant)	.551	.626		.881	.382
Communication Quality	**.674**	**.134**	**.551**	**5.022**	**.000**
Task-Technology Fit & Acceptance	-.087	.144	-.083	-.603	.549
Communication Applications Characteristics	.243	.153	.219	1.582	.119
a. Dependent variable: *Perceived performance*					

7 Discussion

The participants in our study relied on their interaction in virtual teams to perform the task that was assigned to them in form of a course project. Most of the subjects (79.7%) stated that they used communication platforms *"very often"* or *"often"*. Regarding the first research question (RQ1) in our study, it must be noted that their most frequent choices of a communications platform were *WhatsApp* (43.75%), *Discord* (31.25%), *Facebook Messenger* (9.38%), *Slack* (3.13%) and *MS Teams* (3.13%). Most of the students in our study use multiple communication tools for performing educational activities on daily basis, which is in the line with the results of research on using Slack (see: Menzies & Zarb, 2020). The mentioned research also established that Facebook Messenger, WhatsApp, Snapchat and Discord were the most popular applications for communication.

Having in mind the previously presented results of data analysis and considering the second research question in our study (RQ2), we confirmed the satisfactory internal consistency (measured by the Cronbach's alpha coefficient) of all scales and subscales that were designed or adapted for our study. As can be seen in *Table 1*, all of the Cronbach's alpha coefficients had value above .70, which can be considered as satisfactory for preliminary research.

The results of regression analysis in our study (see *Table 2*) revealed that *Communication Quality* was the most influential predictor of the criterion variable *Perceived Performance* in virtual teams. This answers the third research question (RQ3) and is in correspondence with the findings of Garro-Abarca et al. (2020), who also used regression analysis for prediction of *virtual team performance* (as a dependent variable) and found that its variance is mostly determined by the predictors labeled *"Trust"* and *"Virtual Team Communication"* in their study.

8 Conclusion

Our pilot study used the *Task-Technology Fit* (TTF) model as a theoretical framework to design a measurement instrument for assessing the performance of virtual teams in higher education. All of the measurement scales for TTF related variables in our pilot study had acceptable internal consistency. Our research also revealed that the two most frequently used communication tools for collaboration in students' virtual teams were *WhatsApp* and *Discord*, which is in line with research results of other authors. Furthermore, it was revealed that the quality of communication is a most influential and statistically significant predictor of the perceived performance in virtual teams. This finding is in line with the claim in literature that the quality of communication is more important than its frequency. It also highlights the importance of accurate and understandable communication among students' virtual team members. The implication for researchers of students' virtual teams would be to focus on evaluating the quality attributes of students' communication rather than the number of posts or messages they exchange or publish.

In the research presented in our paper the sample was rather small since it involved students from only one course. However, such research context was suitable for examining the students working in predominantly virtual teams on project tasks due to the COVID-19 situation. Therefore, in our future research we aim to further test the measurement instrument by including subjects from several other university courses. Next, the sample was rather homogeneous, which means that their prior knowledge of their field of study, common ways of communication and task completion routine may have influenced high values of Cronbach alpha.

In the future, we plan to investigate the relationship between the TTF constructs in a context outside higher education, especially in IT companies, where employees extensively use communication platforms for collaboration in virtual teams.

In addition to testing the measurement instruments in the real context of virtual teams, the instrument can also be tested in an international academic environment, for example in courses where students collaborate with colleagues from other universities on virtual team tasks. In the latter case, the impact of cultural factors on the performance perception of team members should also be investigated. On the other hand, in the research on virtual teams in a real working environment more detailed task characteristics and their possible influence on the perceived performance of virtual team members should be taken into account.

References

Aiken, M., Gu, L., & Wang, J. (2013). Task knowledge and task-technology fit in a virtual team. *International Journal of Management, 30*(1), 3-11.

Aladsani, H. K. (2021). University Students' Use and Perceptions of Telegram to Promote Effective Educational Interactions: A Qualitative Study. *International Journal of Emerging Technologies in Learning*, *16*(9), 182-197. https://doi.org/10.3991/ijet.v16i09.19281

Bezerra, C.I.M., de Souza Filho, J.C., Coutinho, E.F., Gama, A., Ferreira, A.L., de Andrade, G.L., & Feitosa, C.E. (2020). How human and organizational factors influence software teams productivity in COVID-19 pandemic: A Brazilian survey. *SBES '20: Proceedings of the 34th Brazilian Symposium on Software Engineering*, ACM, New York, NY, USA, 606-615. https://doi.org/10.1145/3422392.3422417

Cane, S., & McCarthy, R. (2009). Analyzing the factors that affect information systems use: A task-technology fit meta-analysis. *Journal of Computer Information Systems*, *50*(1), 108-123. https://doi.org/10.1080/08874417.2009.11645368

Carlson, J. R., Carlson, D. S., Hunter, E. M., Vaughn, R. L., & George, J. F. (2013). Virtual team effectiveness: Investigating the moderating role of experience with computer-mediated communication on the impact of team cohesion and openness. *Journal of Organizational and End User Computing*, *25*(2), 1-18. https://doi.org/10.4018/joeuc.2013040101

Čižmešija, A., & Bubaš, G. (2020). An Instrument for Evaluation of the Use of the Web Conferencing System BigBlueButton in e-Learning. *Central European Conference on Information & Intelligent Systems*, (October), 63-71.

Daim, T. U., Ha, A., Reutiman, S., Hughes, B., Pathak, U., Bynum, W., & Bhatla, A. (2012). Exploring the communication breakdown in global virtual teams. *International Journal of Project Management*, *30*(2), 199-212. https://doi.org/10.1016/j.ijproman.2011.06.004

Fuller, R., & Dennis, A. (2009). Does Fit Matter? The Impact of Task-Technology Fit and Appropriation on Team Performance in Repeated Tasks. Information Systems Research, 20(1), 2-17. https://doi.org/10.1287/isre.1070.0167

Garro Abarca, V.M, Palos-Sanchez, P.R. & Rus-Arias, E. (2020). Working in virtual teams: A systematic literature review and a bibliometric analysis. *IEEE Access*, *8*, 168923-168940. https://doi.org/10.1109/ACCESS.2020.3023546

Gonçalves, S. P., Sousa, M. J., & Pereira, F. S. (2020). Distance learning perceptions from higher education students—the case of Portugal. *Education Sciences*, *10*(12), 1-15. https://doi.org/10.3390/educsci10120374

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly: Management Information Systems*, *19*(2), 213–233. https://doi.org/10.2307/249689

Goodhue, D.L. (1998). Development and measurement validity of a task-technology fit instrument for user evaluations of information system. Decision Sciences, 29(1), 105-138. https://doi.org/10.1111/j.1540-5915.1998.tb01346.x

Hauder, M., Fiedler, M., Matthes, F., & Wüst, B. (2013). Analyzing task and technology characteristics for enterprise architecture management tool support. Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC, 267-274. https://doi.org/10.1109/EDOCW.2013.36

Hauke, J., & Kossowski, T. (2011). Comparison of values of pearson's and spearman's correlation coefficients on the same sets of data. Quaestiones Geographicae, 30(2), 87-93. https://doi.org/10.2478/v10117-011-0021-1

Hertel, G., Konradt, U., & Orlikowski, B. (2004). Managing distance by interdependence: Goal setting, task interdependence, and team-based rewards in virtual teams. European Journal of Work and Organizational Psychology, 13(1), 1-28. https://doi.org/10.1080/13594320344000228

Hidayat, D., Pangaribuan, C. H., Putra, O. P. B., & Irawan, I. (2021). Contemporary studies of Task-Technology Fit: A review of the literature. 2021 International Conference on Information Management and Technology (ICIMTech), 309-313. http://doi: 10.1109/ICIMTech53080.2021.9535028.

Jony, A.I., & Serradell-López, E. (2019). Effective virtual teamwork development in higher education: A systematic literature review. EDULEARN19 Proceedings, 11th International Conference on Education and New Learning Technologies (pp. 873-882), Palma, Mallorca, Spain. http://doi: 10.21125/edulearn.2019.0285

Kilcullen, M., Feitosa, J., & Salas, E. (2021). Insights From the Virtual Team Science: Rapid Deployment During COVID-19. Human Factors, February 2021 (Online first). https://doi:10.1177/0018720821991678

Klonek, F.E., Kanse, L., Wee, S., Runneboom, C., & Parker, S.K. (2022). Did the COVID-19 lock-down make us better at working in virtual teams? Small Group Research, 53(2), 185-206. https://doi:10.1177/10464964211008991

Marlow, S. L., Lacerenza, C. N., Paoletti, J., Burke, C. S., & Salas, E. (2018). Does team communication represent a one-size-fits-all approach?: A meta-analysis of team communication and performance. *Organizational Behavior and Human Decision Processes*, *144*(February 2016), 145-170. https://doi.org/10.1016/j.obhdp.2017.08.001

Menzies, R., & Zarb, M. (2020, October). Professional communication tools in higher education: A case study in implementing Slack in the curriculum. In 2020 IEEE Frontiers in Education Conference (FIE) (pp. 1-8). IEEE.

Morrison-Smith, S., & Ruiz, J. (2020). Challenges and barriers in virtual teams: A literature review. *SN Applied Sciences*, *2*, 1096. https://doi.org/10.1007/s42452-020-2801-5

Muszyńska, K. (2021). A bibliometric review of research on communication in virtual project teams. *Procedia Computer Science*, *192*, 4770-4779. https://doi.org/10.1016/j.procs.2021.09.255.

Spies, R., Grobbelaar, S., & Botha, A. (2020). A scoping review of the application of the Task-Technology Fit Theory. In: Hattingh, M., Matthee, M., Smuts, H., Pappas, I., Dwivedi, Y., Mäntymäki, M. (eds) *Responsible Design, Implementation and Use of Information and Communication Technology. I3E 2020. Lecture Notes in Computer Scienc*e, 12066. Springer, Cham. https://doi.org/10.1007/978-3-030-44999-5_33

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, *48*(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Ursachi, G., Horodnic, I. A., & Zait, A. (2015). How Reliable are Measurement Scales? External Factors with Indirect Influence on Reliability Estimators. *Procedia Economics and Finance*, *20*(15), 679-686. https://doi.org/10.1016/s2212-5671(15)00123-9

Whillans, A., Perlow, L., & Turek, A. (2021). Experimenting during the shift to virtual team work: Learnings from how teams adapted their activities during the COVID-19 pandemic. Information and Organization, 31(1), 100343. https://doi.org/10.1016/j.infoandorg.2021.100343

Implementing Educational Humanoid Robots in Croatian Secondary Schools: Potential Challenges and Benefits

Snježana Babić, Darko Etinger
Faculty of Informatics
Juraj Dobrila University of Pula
Rovinjska 14, Pula, Croatia
(snjezana.babic,darko.etinger}@unipu.hr

Monika Lenić
Vrbani Elementary School
Listopadska ulica 8, Zagreb, Croatia
monika.lenic1@gmail.com

Abstract. *In this study, we have examined secondary school teachers' perceptions about the benefits and challenges of using educational humanoid robots in their teaching processes and their connection with the subject fields. In this regard, a survey was conducted among 154 Croatian secondary school teachers, almost none of whom were users of educational humanoid robots in the teaching process. The results of the research presented in this paper can contribute to understanding the factors of adopting educational humanoid robots by secondary school teachers and developing strategies for implementing this innovative educational technology.*

Keywords. educational humanoid robots, adoption, benefits, challenges, secondary school teachers, Croatian

1 Introduction

In general, humanoid robots can perform certain pre-programmed activities, communicate with humans, and receive commands from their users. From a technical point of view, this type of robot is equipped with sensors, speakers, cameras, and actuators and very often has the shape of a human body. Several categories of these robots on the market today are designed for specific areas of activity; for example, humanoid healthcare robots, social humanoid robots, and educational humanoid robots. (Choudhury et al., 2018). Nowadays, humanoid robots with social skills have been used in different fields of education, such as foreign language education, science education, and special education (Sisman et al., 2019). As an example, Pepper and NAO are humanoid robots designed to interact with humans used in education, healthcare, and research and can perform multiple actions. Compared to other educational technologies, the key advantage of humanoid robots in teaching is their social and physical presence and individual teaching capabilities, which can improve learning outcomes (Belpaeme and Tanaka, 2021). Implementing robotics-based activities in the teaching process can enable teachers to apply constructivism, constructionism, and inquiry-based learning that include collaboration, creative thinking and problem solving, and active learning to improve teaching and learning (Chalmers et al., 2021).

Considering the complexity of implementing educational humanoid robots in classes, Mishra et al. (2021) have proposed a multidisciplinary framework for the use of humanoid robots in an educational environment that integrates the following four perspectives: (1) technological (human-robot interaction); (2) pedagogical/didactic, (3) effectiveness of humanoid robots (psycho-social); (4) ethical implications of using humanoid robots.

Researchers in this field have identified numerous benefits and challenges of using humanoid robots as teaching and learning assistants. Since the implementation of humanoid robots in the teaching process depends on teachers' acceptance of this innovative technology, there is a need for research in this area in different educational contexts. This paper examines the teachers' perception of the benefits and challenges of using educational humanoid robots in Croatian secondary schools.

2 Literature Review

The results of qualitative research conducted by Ahmad et al. (2016) show that language teachers in primary and secondary schools, after using a NAO robot, perceived that a NAO robot could be helpful in language teaching due to its ability to answer frequently asked questions in class, the possibility of monitoring the child's memory, the ability to adapt to the child's personality and emotions in real-time, and the possibility of adapting to different cultures. The authors also have concluded that a user interface design that is easy to update with new lessons would be a factor that would make it easier for teachers to use the NAO robot in the classroom. After testing the robot Pepper as a storyteller in a real context in two modalities, De Carolis et al. (2021) have concluded

that the synthesized voice of the robot, in combination with non-verbal behavior, has a more positive effect on the emotional experience of children (from 8 to 9 years of age) by creating the impression of a more accessible learning compared to the first modality, when a human voice narrated the story. Pepper played "only the role of a device", i.e. when the robot was a kind of audiobook. In the example of the use of a NAO robot as a teaching assistant in a primary school, the results of Mubin et al. (2019) have found that the children have shown more activity in class, but with no significant difference in test scores. Furthermore, Karakosta et al. (2019) have found that the robot Kaspar has had a positive impact on children diagnosed with autism spectrum disorder behaviors in specific domains, such as communication and interaction, prompted speech, unprompted imitation, and focus/attention. According to Sisman et al. (2019), students' attitudes toward using humanoid robots in an educational environment are influenced by their perceptions of enjoyment, anxiety, and engagement. Alhashmi et al. (2021) have found that teachers have expressed concern and desire for ease of use of robots in class, while students have expressed satisfaction with using humanoid robots as co-teachers in class.

The results of the research conducted by Woo et al. (2021) in a sample of studies between 2000 and 2020, which examined social robots in classrooms, have shown that there is no solid evidence that social robots are more operative than human teachers. Prior to the aforementioned research, there had been very little research on the ethical and safety issues. Furthermore, Fridin and Belokopytov (2014) have found that teachers' intention to use NAO robots in teaching is strongly influenced by their perceived usefulness.

Therefore, as previously mentioned in this paper, the results of numerous studies have shown that teachers perceive multiple benefits of using humanoid robots as assistants in learning and teaching; for example, it helps in the development of communication skills and the development of teamwork skills (Burbaite et al., 2013; Fridin and Belokopytov, 2014; Khanlari, 2015; Chevalier et al., 2016), students are more creative, it stimulates good emotions, the robot is easy to use (Fridin and Belokopytov, 2014; Reich-Stiebert and Eyssel, 2016), it develops logical and creative thinking (Chevalier et al., 2016), robots bring added value to education, increase student activity in class, increase student motivation (Burbaite et al., 2013; Reich-Stiebert and Eyssel, 2016), encourage innovative pedagogical practices and can facilitate the education of students that encounter difficulties during the regular curriculum (Kradolfer et al., 2014).

Also, research has shown that teachers perceive significant technical, economic, and logistical challenges to introducing humanoid robots into teaching (Belpaeme and Tanaka, 2021). In this regard, as an example, teachers perceive a lack of time to start the robotic activity during the lesson, inadequate access to supporting materials, lack of technical and teaching support, and lack of technical and pedagogical knowledge (Burbaite et al., 2013; Almisis, 2013; Kradolfer et al., 2014). Based on the experience of implementing a NAO robot in an elementary school, Gaber (2021) has concluded that the main challenge for implementing a NAO robot in the classroom is its cost (e.g. a NAO robot costs between 8000 and 16000 USD) and the additional cost for the professional training of the user of the educational humanoid robot. Also, the author points out that there is a need to research challenges related to the implementation of educational humanoid robots in classrooms, which relate to problems such as safety and the need for professional training of teachers. Fox and Gambino (2021) emphasize that humanoid social robots in education are limited in the tasks for which they are programmed. They are thus limited in their ability to personalize; for example, most of the humanoid social robots have a small set of possible responses, cannot identify or differentiate between different users and have minimal conversational control. Based on the theoretical research, it is evident that there has been a justified need to conduct research in this context in different educational environments. In the continuation of this paper, the methodology of the conducted research and the key findings of the conducted study are described. Lastly, the most important conclusions of this work are given. It is essential to emphasize that almost all of the respondents were non-users of a humanoid robot in the teaching process.

3 Methodology

This paper focuses on the perceived benefits and challenges of using educational humanoid robots in the classroom by Croatian secondary school teachers to improve understanding of the adoption factors of this innovative technology. In this regard, the research questions are:

1. *What benefits and challenges of using educational humanoid robots in teaching do Croatian secondary school teachers see?*
2. *How do teachers see the benefits and challenges of using educational humanoid robots depending on the subject fields?*

Based on a voluntary and anonymous approach, the research was conducted from May to June 2021, for the master thesis of the co-author of this paper (Lenić, 2021). The questionnaire was created using an online Google Forms questionnaire forwarded to the Croatian Facebook group “School Staff Room”, WhatsApp and Viber. Some secondary school principals distributed the questionnaire via e-mail or posted it in virtual groups (e.g., Yammer). A new questionnaire containing 37 closed questions was created for this research. Answers to 13 questions related to the adoption factors of humanoid robots are modulated on

a five-point Likert scale (1 = completely false; 2 = mostly false; 3 = neither true nor false; 4 = mostly true; 5 = completely correct). Some items in this questionnaire have been adapted from the unified theory of acceptance and use of technology - UTAUT (Venkatesh et al., 2003), while some items are defined according to the researched literature (Alimisis, 2013; Burbaite et al., 2013; Fridin and Belokopytov, 2014; Kradolfer et al., 2014; Khanlari, 2015; Chevalier et al., 2016). In this research paper, only the selected results of the conducted study (Lenić, 2021) are presented.

4 Results and Discussion

4.1 Participants

A total of 154 respondents (Croatian secondary school teachers) participated in the research (75.97% women and 24.03% men). Most respondents were between 31 and 40 years old (29.9%), followed by those between 41 and 50 years old (26.6%), and those between 51 and 60 years old (24.7%), 5.8% of them were under 30 years old, while 4.5% of them were over 60 years old. Most respondents had higher education (75.3%), followed by higher vocational education (14.3%), and slightly fewer respondents (10.4%) had secondary vocational education. As for the subject area, most respondents were teachers from scientific-mathematical and professional fields (56.5%), followed by social-humanities (19.5%), language-communication (13.6%), IT-technical occupations (6.5%), art (1.9%) and other fields (1.9%). Most respondents had between 21 and 30 years of teaching experience (28%), followed by those between 4 and 10 years (27%), and between 11 and 20 years (25%), while 14% had less than three years of teaching experience, and 6% of them had 30 years or more of teaching experience. Most respondents were teachers working in vocational schools (96.02%), while 27.90% worked in grammar schools and 4.5% in other schools. The results of processing the collected data show that the most significant number of respondents have never used educational robots in class (92.2%), and 90.3% of them have not attended any form of education, nor have they educated themselves on the topic of using educational humanoid robots in the teaching process.

Most respondents have self-assessed their knowledge about using humanoid robots in the teaching process as insufficient (44.2%) and sufficient (39.6%). In comparison, 16.2% have stated that they know more about educational humanoid robots.

4.2 Findings

The results of the analysis of the data collected in this research on the factors of acceptance of educational humanoid robots by Croatian secondary school teachers are divided into benefits and challenges, which are shown in Figure 1 and Figure 2. A 5-point Likert scale was used to measure the items shown in the mentioned figures.

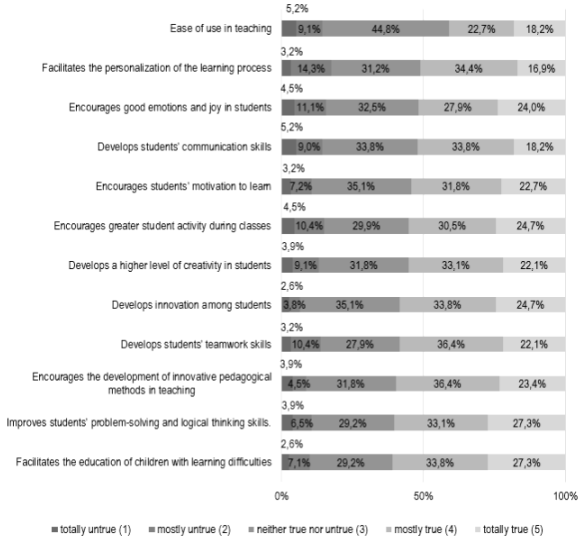


Figure 1: Croatian secondary school teachers' (N = 154) response percentage concerning their perceived benefits of educational humanoid robots (according to Lenić, 2021)

As can be seen in Figure 1, most participants (61.1%) in this study believe that the use of educational humanoid robots helps facilitate the education of children with learning disabilities, and 60.4% of them believe that this innovation in teaching can improve students' problem-solving and logical thinking skills, developing students' innovation (58.5%), and improve students' teamwork skills (58.5%). Furthermore, the results show (Figure 1) that 59.8% of the participants agree with the statement that the use of humanoid robots as teaching and learning assistants encourages the development of innovative pedagogical methods in the teaching process by secondary school teachers. Arithmetic mean (shown in Figure 1) indicates that 55.2% participants are of the opinion that robots can encourage greater student activity during classes and develop a higher level of creativity in students, and 54.5% of them believe that humanoid robots as teaching and learning assistants encourage students' motivation to learn. The results of the collected data (Figure 1) show that only half of the participants generally agrees that the use of a humanoid robot in class could develop students' communication skills (52%), that it facilitates the personalization of the learning process (51.3%), encourages good emotions and joy in students (51.09%). Most of them, 44.8%, are undecided in response to the statement about the ease of using an educational humanoid robot. This is confirmed by the result of the arithmetic mean (M=3.40, SD=1.08), which indicates that the participants, on average, are undecided. In addition, from the results shown in Figure 1, it is evident that a significant number of respondents do not agree that the use of humanoid robots as assistants in teaching and learning in the classroom facilitates the personalization

of the learning process (17.5%), they also do not agree that the humanoid robots encourage good emotions and joy in students (15 .5%), nor that they promote more significant student activity during class (14.9%).

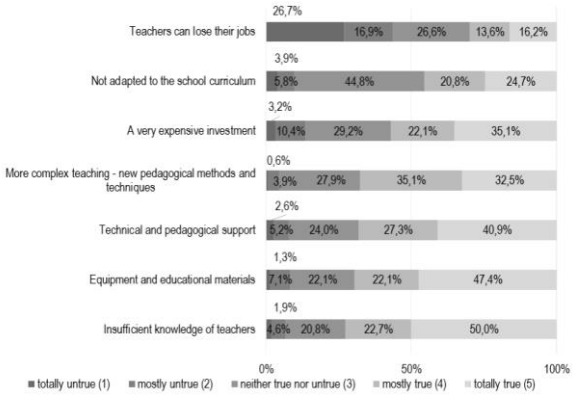


Figure 2. Croatian secondary school teachers' (N = 154) response percentage concerning their perceived challenges of educational humanoid robots (according to Lenić, 2021)

The results presented in Figure 2 show that the majority of the respondents (72.7%) believe that secondary school teachers do not have enough knowledge to use humanoid robots in class, 69.5% of them believe that they do not have adequate equipment and educational materials and that they do not have technical and pedagogical support for the use of humanoid robots in teaching (68.2%).

Also, most Croatian secondary school teachers (67.6%) believe that teaching with educational robots would be more complex, considering the use of new pedagogical methods and techniques. The arithmetic mean (M=3.75, SD=1.18) shows that, on average, the respondents agree that using humanoid robots in teaching is a costly investment (57.1%). Only slightly less than half of the respondents (45.5%) agree with the statement that the use of humanoid robots in classes "is not adapted to the school curriculum". In comparison, 44.8% of them remain undecided, and only 9.7% of the participants believe this item to be "mostly untrue" and/or "totally untrue". Furthermore, 43.5% of respondents believe they cannot lose their jobs if humanoid robots are used as assistants in teaching and learning. In comparison, almost a third of the respondents (29.8%) believe that implementing humanoid robots in education may result in losing a teaching job, and 26.6% of them are undecided (see Figure 2).

On average, those Croatian secondary school teachers that have participated in this research (M=3.54; SD=1.28) have stated that they intend to use educational humanoid robots in the near future if given the opportunity to do so. In this regard, 53.3% of them have agreed that in the near future, they intend to use a humanoid robot as an assistant in teaching and learning. In comparison, 27.9% are undecided, and only 18.8% have not declared their intention to use a humanoid robot in their teaching.

4.2.1 Differences in the perception of educational humanoid robots between secondary school teachers with no knowledge and those with partial knowledge of this technology

An independent samples t-test was conducted on a sample of 129 respondents, of whom 47.3% have reported partial knowledge and 52.7% reported no knowledge of educational humanoid robots to determine the differences in their perception of the benefits and challenges of the implementation of educational humanoid robots in their teaching.

According to the results obtained in Table 1, the results of the t-test for independent samples show that there is a statistically significant difference between the respondents in this study who have expressed partial knowledge and those who have expressed no knowledge at all about educational humanoid robots with regard to the perception of ease of use (t=2,69, p=,008), as well as facilitation of the personalization of the learning process (t=2.01, p=,047), stimulation of good emotions and joy in students (t=2,75, p=,007), development of students' communication skills (t=3,27, p=,001), stimulation of students' motivation for learning (t=3,04, p=,003), stimulation of greater student activity during classes (t=2,82, p=,006), development of a higher level of creativity in students (t=2,44, p=,016), development of innovation in students (t=3,02, p=,003), development of students' teamwork skills (t=3,53, p=,001), encouragement of the development of innovative pedagogical methods in teaching (t=2,79, p=,006), improvement the problem-solving skills and logical thinking of students (t=2,60, p=,010); with the group of teachers that have expressed partial knowledge about educational humanoid robots at a higher level perceiving the mentioned advantages compared to one teacher who has expressed that they had no knowledge at all about this innovative technology. Furthermore, the results of the t-test (t=1,97, p=,051) have shown that there are no statistically significant differences between the observed groups of subjects with regard to the perception of facilitating the education of children with learning difficulties (see Table 1).

However, the T-test for independent samples has confirmed a statistically significant ($p<0,05$) difference between the observed groups of respondents in this research regarding the opinion about insufficient knowledge about educational humanoid robots by teachers, where the respondents who have stated that they have insufficient knowledge about educational humanoid robots (t=-2.14, p=.034) believe more that teachers do not have enough knowledge about educational humanoid robots compared to the respondents who have stated that they have partial knowledge about this innovative technology (see Table 2).

Table 1. The results of the t-test analysis among respondents (N=129) about the perceived benefits of educational humanoid robots in teaching and teachers' knowledge about them

Variables/ benefits	Teachers' knowledge		T-test (p-value)
	sufficient (4)	insufficient (5)	
	M (SD)	M (SD)	
Ease of use in teaching	3,46 (0,91)	3,00 (1,02)	**2,69 (,008)**
Facilitates the personalization of the learning process	3,62 (1,10)	3,25 (1,01)	**2,01 (,047)**
Encourages good emotions and joy in students	3,79 (1,08)	3,26 (1,07)	**2,75 (,007)**
Develops students' communication skills	3,75 (1,01)	3,18 (0,99)	**3,27 (,001)**
Encourages students' motivation to learn	3,85 (0,95)	3,31 (1,07)	**3,04 (,003)**
Encourages greater student activity during classes	3,84 (1,04)	3,29 (1,13)	**2,82 (,006)**
Develops a higher level of creativity in students	3,75 (1,04)	3,31 (1,03)	**2,44 (,016)**
Develops innovation among students	3,92 (0,90)	3,41 (1,00)	**3,02 (,003)**
Develops students' teamwork skills	3,90 (1,01)	3,28 (0,99)	**3,53 (,001)**
Encourages the development of innovative pedagogical methods in teaching	3,85 (0,93)	3,37 (1,04)	**2,79 (,006)**
Improves students' problem-solving and logical thinking skills.	3,90 (1,01)	3,43 (1,06)	**2,60 (,010)**
Facilitates the education of children with learning difficulties	3,90 (0,93)	3,54 (1,11)	1,97 (,051)

Note: p<0.05

Table 2. The results of the t-test analysis among respondents (N=129) about the perceived challenges of educational humanoid robots in teaching and teachers' knowledge about them

Variables/ challenges	Teachers' knowledge		T-test (p-value)
	sufficient (4)	Insufficient (5)	
	M (SD)	M (SD)	
Teachers can lose their jobs	2,67 (1,33)	3,10 (1,42)	-1,78 (,078)
Not adapted to the school curriculum	3,61 (1,13)	3,65 (1,02)	-0,21 (,831)
A very expensive investment	3,80 (1,03)	3,87 (1,14)	-0,33 (,739)
More complex teaching - new pedagogical methods and techniques	4,02 (0,79)	3,96 (1,01)	0,38 (,708)
Technical and pedagogical support	4,11 (0,97)	3,99 (1,14)	0,69 (,491)
Equipment and educational materials	4,08 (1,02)	4,19 (1,04)	-0,60 (,549)
Insufficient knowledge of teachers	**3,69 (0,96)**	**4,07 (1,07)**	**-2,14 (,034)**

Note: p<0.05

4.2.2 The connection between perceived benefits and challenges of using humanoid robots in teaching and subject fields

One-way analysis of variance (ANOVA) and Tukey post hoc tests were used to determine the relationship between perceived benefits and challenges of using humanoid robots in teaching and the subject areas.

As previously stated, in this research, subject areas are divided into scientific-mathematical (28.6%) and professional fields (27,9%), social-humanities (19.5%), language-communication (13.6%), IT-technical occupations (6.5%), art (1.9%) and other fields (1.9%). Table 3 shows the results of the ANOVA analysis among respondents/Croatian secondary school teachers on the perceived benefits of educational humanoid robots in teaching and subject areas. According to the obtained results in Table 3, it is evident that regarding the subject field, there is a statistically significant difference in the perception of the following benefits of an educational humanoid robot in teaching by secondary school teachers: (it) facilitates the personalization of the learning process, encourages students' motivation to learn, encourages better student activity during classes, develops a higher

level of creativity in students, develops innovation among students, develops students' teamwork skills, improves students' problem-solving and logical thinking skills, facilitates the education of children with learning difficulties. By carrying out the Tukey post hoc test, it has been determined that secondary school teachers in the professional fields believe that using humanoid robots can facilitate the personalization of the learning process and develop innovation among students, develop students' teamwork skills, and improve students' problem-solving skills and logical thinking, in comparison to teachers from the social-humanities field. Furthermore, it has been determined that secondary school teachers of other fields believe that using humanoid robots can facilitate the education of children with learning disabilities and encourage better student activity during classes than teachers of scientific-mathematical and professional fields.

Table 3. The results of the ANOVA analysis among respondents (N=154) about the perceived benefits of educational humanoid robots in teaching and subject fields

Variables/Benefits	**Area of subject** ANOVA	
	F	p-value
Ease of use in teaching	1,14	,340
Facilitates the personalization of the learning process	**2,61**	**,019***
Encourages good emotions and joy in students	2,06	,062
Develops students' communication skills	1,90	,085
Encourages students' motivation to learn	**2,19**	**,047***
Encourages greater student activity during classes	**2,84**	**,012***
Develops a higher level of creativity in students	**2,28**	**,039***
Develops innovation among students	**3,03**	**,008***
Develops students' teamwork skills	**3,25**	**,005***
Encourages the development of innovative pedagogical methods in teaching	2,14	,052
Improves students' problem-solving and logical thinking skills.	**3,12**	**,007***

Note: p<0.05

Furthermore, as can be seen from the results of the ANOVA analysis in Table 4, regarding the subject area, there is a statistically significant difference in the perception of the following challenges of educational humanoid robots in the classroom: a costly investment, the need of technical and pedagogical support and insufficient knowledge of teachers.

Table 4. The results of the ANOVA analysis among respondents (N=154) about the perceived challenges of educational humanoid robots in teaching and subject fields

Variables/challenges	**Area of subject** ANOVA	
	F	p-value
Teachers can lose their jobs	2,12	,054
Not adapted to the school curriculum	1,92	,082
A very expensive investment	**2,57**	**,021***
More complex teaching - new pedagogical methods and techniques	0,67	,677
Technical and pedagogical support	**2,40**	**,031***
Equipment and educational materials	1,07	,384
Insufficient knowledge of teachers	**2,44**	**,028***

Note: p<0.05

The Tukey post hoc test has determined that secondary school teachers in other fields exhibit a significantly lower belief that implementing humanoid robots in their classes is a costly investment than secondary school teachers in the scientific-mathematical, social-humanities, and language-communication fields.

5 Limitations of the Study

The limitations of this study are reflected in the number of research participants, the number of adoption factors, and the participants' knowledge about the topic. According to the results of this research, 44.2% of the respondents have stated that they have insufficient knowledge, 39.6% of respondents have stated that they have sufficient knowledge, while only 16.2% of them have stated that they know more about educational humanoid robots. Regarding the abovementioned results, it should be noted that teachers do not have enough knowledge to be able to draw conclusions about the real benefits and challenges of this innovative technology. Accordingly, the interpretation of the results should be approached cautiously. But on the other hand, these results are considered to be of great importance because they confirm the difference in the perception of those secondary school teachers who have no knowledge and those who have partial knowledge about educational

humanoid robots. Consequently, selected factors can influence further acceptance of this technology by Croatian secondary school teachers, and therefore, they need to be additionally taken into account in further research in this context. For the sake of generalization, research based on our findings should be carried out in the whole country (national pilot study), as well as in an international context, to determine the difference in the perception of secondary school teachers about the benefits and challenges of this innovative educational technology in different educational settings, especially considering different educational areas. The number of adoption factors should be expanded, with other factors related to the cost of implementation in the educational context and secondary school teachers' competencies for using humanoid robots in teaching.

Despite the limitations, the findings offer an understanding of specific benefits, challenges, and recommendations for a better understanding of the determinants of acceptance of humanoid robots in secondary education.

6 Recommendations

Our findings can help researchers as a foundation for future research as well as all practitioners who wish to develop strategies for implementing humanoid robots in educational environments. Based on the results of this study, below are the requirements when deciding on the broader implementation of humanoid robots in Croatian secondary education:

- Pedagogical and technical training of secondary school teachers
- Technical and pedagogical support
- Defining the role of the humanoid robot in the teaching process so that teachers can be sure that their jobs are not threatened
- Defining the educational values of humanoid robots so that secondary school teachers can develop a positive attitude toward their use
- Extra funding would be required - educational humanoid robots, educational materials, professional training for all stakeholders
- Development of innovative pedagogical methods and techniques for implementing humanoid robots in teaching.

7 Conclusion

Humanoid robots as assistants in teaching and learning are innovative educational technologies in education. It takes time for them to be accepted by all stakeholders of the educational process; among them is the key role of the teacher.

In this paper, the adoption factors of educational humanoid robots in classes by Croatian secondary school teachers were specifically investigated, and 92.2% of them have stated that they are not users of this innovation. Furthermore, 83.8% of the respondents have said they do not know how to use educational humanoid robots in class.

Moreover, the collected data shows that Croatian secondary school teachers perceive the benefits and challenges of using humanoid robots in teaching. For example, 61.1% of participants believe that a humanoid robot can facilitate the education of children with learning difficulties.

This research has confirmed a statistically significant connection between Croatian secondary school teachers' perception of certain benefits and challenges of educational humanoid robots in classes and subject fields. This is consistent with the results of other studies mentioned in this paper.

A special contribution of this paper is in confirming the difference between those Croatian secondary school teachers who have stated that they do not have enough knowledge and those who do not have any knowledge about educational humanoid robots, with regard to the selected factors in this research. The aforementioned shows that the education of teachers about educational humanoid robots is very important for the implementation of humanoid robots in their teaching. Moreover, it confirms that the obtained factors can be used for further research in this context, as well as for defining the implementation strategy of this innovative technology in Croatian secondary schools.

The intention to use humanoid robots in courses in the near future has been expressed by slightly more than half of the respondents (53.3%), while 27.9% could not decide.

And finally, the research results have shown there is a need to develop strategies for training secondary school teachers to use educational humanoid robots.

References

Ahmad, M. I., Mubin, O., & Orlando, J. (2016). Understanding behaviours and roles for social and adaptive robots in education: teacher's perspective. *Proceedings of the fourth International Conference on Human Agent interaction*, 297-304. doi:10.1145/2974804.2974829

Alhashmi, M., Mubin, O., & Baroud, R. (2021). Examining the use of robots as teacher assistants in UAE classrooms: teacher and student perspectives. *Journal of Information Technology Education: Research*, 20, 245-261. doi: 10.28945/4749

Alimisis, D. (2013). Educational robotics: Open questions and new challenges. *Themes in Science and Technology Education*, 6(1), 63-71. Retrieved from https://files.eric.ed.gov/fulltext/EJ1130924.pdf

Belpaeme, T., & Tanaka, F. (2021). Social robots as educators. In OECD *Digital Education Outlook 2021: Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots*, 143-160. doi: 10.1787/1c3b1d56-en

Burbaite, R., Stuikys, V. & Damasevicius, R. (2013) Educational robots as collaborative learning objects for teaching Computer Science, *Proceedings of 2013 International Conference on System Science and Engineering* (ICSSE), 211-216. doi: 10.1109/ICSSE.2013.6614661.

Chalmers, C., Keane, T., Boden, M., & Williams, M. (2022). Humanoid robots go to school. *Education and* Information Technologies, 1-19. doi: 10.1007/s10639-022-10913-z

Chevalier, M., Riedo, F., & Mondada, F. (2016). Pedagogical uses of thymio II: How do teachers perceive educational robots in formal education?. In IEEE *Robotics & Automation Magazine*, 23(2), 16-23. doi: 10.1109/MRA.2016.2535080.

Choudhury, A., Li, H., Greene, C., & Perumalla, S. (2018). Humanoid robot-application and influence. Appl Sci Res Rev 5:17. doi: 10.48550/arXiv.1812.06090, Retrieved from https://www.primescholars.com/articles/humanoid-robot-application-and-influence-95412.html

De Carolis, B., D'Errico, F., & Rossano, V. (2021). Pepper as a Storyteller: Exploring the Effect of Human vs. Robot Voice on Children's Emotional Experience. In IFIP Conference *on Human-Computer Interaction* (471-480). Springer, Cham. doi: 10.1007/978-3-030-85616-8_27

Fridin, M., Belokopytov, M. (2014). Acceptance of socially assistive humanoid robot by preschool and elementary school teachers. *Computers in Human Behavior*, 33, 23-31. Retrieved from https://www.sciencedirect.com/science/article/abs/pii/S0747563213004652

Fox, J., & Gambino, A. (2021). Relationship development with humanoid social robots: Applying interpersonal theories to human–robot interaction. *Cyberpsychology, Behavior, and Social Networking*, 24(5), 294-299.

Gaber, M.A.M. (2021). Exploring the challenges and benefits of implementing humanoid robots in classrooms. *Athena International Journal of Education*, 01 (04), 6-11. Retrieved from https://athenaedujournal.com/uploads/issues/files/file_4078_6499_996.pdf

Karakosta, E., Dautenhahn, K., Syrdal, D. S., Wood, L. J., & Robins, B. (2019). Using the humanoid robot Kaspar in a Greek school environment to support children with Autism Spectrum Condition. Paladyn, *Journal of Behavioral Robotics*, 10(1), 298-317. doi: 10.1515/pjbr-2019-0021

Khanlari, A. (2015) Teachers' perceptions of the benefits and the challenges of integrating educational robots into primary/elementary curricula, *European Journal of Engineering Education*, 41(3), 320-330. doi: 10.1080/03043797.2015.1056106

Kradolfer, S., Dubois, S., Riedo, F., Mondada, F. i Fassa, F. (2014) Social Robotics. A Sociological Contribution to Understanding the Use of Robots in Schools: The Thymio Robot, (eds) Social Robotics. ICSR 2014. *Lecture Notes in Computer Science*, 8755, 217-228. Springer, Cham. doi: 10.1007/978-3-319-11973-1_22

Lenić, M. (2021). Educational robots: master thesis, Faculty of Informatics, Juraj Dobrila University of Pula, 62, Retrieved from https://repozitorij.unipu.hr/islandora/object/unipu:6055

Mishra, D., Parish, K., Lugo, R. G., & Wang, H. (2021). A framework for using humanoid robots in the school learning environment. *Electronics*, 10(6), 756. doi: 10.3390/electronics10060756

Mubin, O., Alhashmi, M., Baroud, R., & Alnajjar, F. S. (2019). Humanoid robots as teaching assistants in an Arab school. *In Proceedings of the 31st Australian Conference on Human-Computer-Interaction*, 462-466. doi: 10.1145/3369457.3369517

Reich-Stiebert, N., & Eyssel, F. (2016). Robots in the classroom: What teachers think about teaching and learning with education robots. *In International conference on social robotics* (671-680). Springer, Cham. doi: 10.1007/978-3-319-47437-3_66

Sisman, B., Gunay, D., & Kucuk, S. (2019). Development and validation of an educational robot attitude scale (ERAS) for secondary school students. *Interactive Learning Environments*, 27(3), 377-388. doi: 10.1080/10494820.2018.1474234

Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D. (2003). User acceptance of information technology: toward a unified view. *Journal MIS Quarterly: Management Information Systems*, 27(3), 425–447. doi: 10.2307/30036540

Woo, H., LeTendre, G. K., Pham-Shouse, T., & Xiong, Y. (2021). The use of social robots in classrooms: A review of field-based studies. *Educational Research Review*, 33, 100388. doi: 10.1016/j.edurev.2021.100388

Digitally Mature Education: The Role of Digital Maturity and Educational Leadership in Meaningful Digital Transformation of Schools

Dragana Kupres
Croatian Academic and Research Network - CARNET
Office of the CEO
Ul. Josipa Marohnića 5, 10000 Zagreb
dkupres@carnet.hr

Nati Cabrera Lanzo, Albert Sangrà Morer
Psychology and Education Sciences Department
Universitat Oberta de Catalunya
Av. del Tibidabo, 39, 08035 Barcelona, Spain
(ncabrera, asangra)@uoc.edu

Abstract. *This theoretical paper conceptualizes the digitally mature education. We propose that digitally mature education is to be co-created and co-lead with transformative and distributed practices while using digital maturity as an improvement and reflective framework for guiding meaningful digital transformation. In conceptualizing digitally mature education, we propose a transformative, inclusive, visionary, effective and accountable vision of education. We argue that the same perspective should be taken when thinking about digital transformation and the purpose of using digital technology in education. We claim that two leadership approaches should be combined to influence the meaningful digital transformation: transformational and distributed leadership.*

Keywords. Digitally mature education, educational leadership, school leadership, digital maturity, digital transformation.

1 Introduction

The last decade has witnessed extensive technological investments supporting national educational reforms, encouraged by covid-19 lockdown (among others European Commission, 2020b; Hakansson Lindquist et al., 2019; Kampylis et all, 2016; Toh and So, 2011). The calls for more in depth thinking on how to approach these initiatives to better harness the potentials of technology and avoid the risks of technological determinisms were present even before the transitioning to emergency remote teaching during pandemic, and are reinforced after it (European Commission, 2020; Hodges et al., 2020; European Schoolnet, University of Liege, 2013; OECD, 2015).

Additionally, recent studies have shown that, while more technology and computer equipment is entering schools as ever before and teachers are encouraged to use it, there is a lack of clear vision in schools themselves of why we need technology and which strategic direction to take (Balaban, 2018; CooperGibson Research, 2022). What kind of technology and how appropriate it is for the students and how do we afford it in the long run, bears tremendous sustainability issues on the shoulders of future policy makers, governments and school heads.

Digital Education Action Plan for 2021-27 states that "education and training system is increasingly part of the digital transformation" (EC, 2020a:p.2). In addition to that direction, we call to look upon the issue from the opposite perspective, in which digital transformation increasingly becomes a meaningful part of education and training system. For this reason, we introduce the term *digitally mature education* that is firstly and deeply rooted in the vision for education, and secondarily supported by meaningful digital transformation. In this article we explore what digitally mature education could encompass and how it can be conceptualized using terms of meaningful digital transformation and improvement/reflective framework of digital maturity.

This paper is a theoretical article that synthesises and adapts (Jakkola, 2020) educational leadership theories in the context of digitally mature education. It synthesises two major theories of educational leadership, transformational and distributed leadership, and it adapts them to the context of technology integration in education, relating them to the concept of digital maturity.

In conceptualizing digitally mature education, the article will propose three claims about how to conceptualize school leadership for meaningful digital transformation and support the claims with argumentations based on literature review, while answering three main research questions.

In the first section, based on literature review and in the context of this paper, we claim that education is a transformative, inclusive, visionary, effective and accountable process of cognitive and social development of a person and her community, aware of the wider global, environmental and digital changes (Anderson and Boyle, 2020; Daniels et al., 2019;

Freire, 1970; Peters and Jandrić, 2017; Townsend, 2011; Underwood and Dillon, 2004).

In the second section, we claim that digital transformation should serve this vision of education, and not take its separate course. Digital technology should support cognitive and social development of individuals and community in transformative, inclusive, effective and accountable ways. We propose to use digital maturity as an organisational improvement and reflective framework designed to guide this kind of meaningful digital transformation (Marshal, 2010, 2012; Teichert, 2019; Volungevičienė et al., 2021).

In the third section, we claim that two leadership approaches should be combined to influence the described meaningful digital transformation: transformational leadership and distributed leadership.

This paper is set to answer the following research questions:

1. What is the role of education when considering global, environmental and digital transformations our societies are experiencing and how can education be part of them? (addressed in section 1)
2. What kind of meaningful digital transformation can support educational processes to become more transformative, inclusive, visionary, effective and accountable and in the same time support cognitive and social development of a person and of her community? (addressed in section 2)
3. What kind of leadership schools need for continuous organisational improvements towards meaningful digital transformation, having in mind the transformative, inclusive, visionary, effective and accountable characteristics of education and related digital maturity? (addressed in section 3).

2 The Role of Education in Global, Environmental and Digital Transformations

The following section discusses the role of education in the present global, environmental and digital turmoil we all are experiencing as global citizens.

Individual growth and transformation are deeply interrelated with social growth and social transformation, and both should be considered as the goals of education (Anderson and Boyle, 2020; Brosio, 2006; Dewey, 1915/2016; Freire, 1970). In this sense the aim of education is to support the development of both cognitive and social abilities of learners (Underwood and Dillon, 2004:213). This is reflected in what Anderson and Boyle understand as 'good' education, "the growth and development of the whole person" (Anderson and Boyle, 2020:page 16). Individual growth is inextricably related to social growth, reflected primarily in the community in which we live. Education provides "opportunity to become contributing and content members of a healthy society". (Anderson and Boyle, 2020:page 16). Critical educational theory understood education as an altruistic and caring path to social and individual transformation, to democratic empowerment, socioeconomic justice, respect for diversity (Brosio, 2006). Dewey understood education as a vital force in forming democracy and civil society (Dewey, 1915/2016). Using the words of George Bernard Shaw, education is the process of "continual becoming" (St. John, 1931) in cognitive and social sphere. Learner is an active and dynamic participant in her education, a creator of her own future self. This thinking is in line with Freire's liberating potential of a person (Freire, 1970). The knowledge is the person who knows - in terms of Aristotle's "Knowledge [epistēmē], in its being-at-work, is the same as the thing it knows" (Aristoteles, 2001).

The role of education is changing and is influenced by the global, environmental and digital turn (Peters and Jandrić, 2017) of our societies. These shifts are not just negative or just positive but are complex and interrelated. Education is a living organism influenced by its environment, co-created by it and forms part of the wider "ecological system" (Harrison et al., 2014). Education has shifted towards a global market approach, including answering the accountability demands (among others Daniels, 2019; Gumus, 2018; Townsend, 2011). Education is a crucial part of sustainability, and both sustainable thinking and long-term visions should form part of changing the education to respond to the needs of environmental and societal transformations (Michelsen and Fischer, 2017). The digital shift in education is pushed more mainstream after pandemic of covid-19, sometimes with vested interest of commercial solutions in education (Teräs et al., 2020).

We cannot expect education to remain intact when the world is in flux. Education needs to change together with society and respond to the global, environmental and digital challenges we are all facing. In this response education should not be replaced by commodified, global enterprise, but needs to be:

- *liberative and transformative*, with strong mission to prepare students to be critical thinkers, global thinkers, active citizens (Anderson and Boyle, 2020; Freire, 1970; Townsend, 2011; Underwood and Dillon, 2004)
- *inclusive* of to the social needs of all citizens, both elites and marginalised groups, co-creating present and future democratic societies (Anderson and Boyle, 2020; Freire, 1970)
- *visionary*, aware of the challenges of the future job markets, climate change and social impact of digital technologies (Peters and Jandrić, 2017)

- *effective*, striving to achieve the educational outcomes (the school intends to achieve) (Daniels et al., 2019)
- *accountable* to the public that provides finances, by using public resources responsibly to provide the best quality service back to the citizens (Townsend, 2011).

Based on the presented arguments we claim that education should be understood as a transformative, inclusive, visionary, effective and accountable process of cognitive and social development of a person which contributes to both the achievement of the full potential of a person and of her community (Anderson and Boyle, 2020, Brosio, 2006, Dewey, 1915/2016, Freire; Underwood and Dillon, 2004). This process should be accountable to the public it serves, public resources it uses, and mindful of the wider global, environmental and digital transformations (Peters and Jandrić, 2017).

3 Digital Maturity as a Transformative, Inclusive, Visionary, Effective and Accountable Framework and a Reflective Tool for Digital Transformation of Education

This section reflects on the nature of digital transformation needed to support the transformative, inclusive, visionary, effective and accountable vision of education.

We acknowledge that digital turn (Peters and Jandrić, 2017) in education is becoming increasingly important with government-led national initiatives trying to use technology as a lever supporting various educational reforms (Hakansson Lindquist et al., 2019; Kampylis et all, 2016; Toh and So, 2011). However, there is a need for a "radical change" of how technology is exploited in teaching and learning (European Schoolnet, University of Liege, 2013), taking an approach that is "both gradually accepting and sceptical" (OECD, 2015).

We also acknowledge that the digital shift in education is pushed more mainstream after pandemic of covid-19. Closing educational institutions down during pandemic years in many ways transferred the responsibility for education to families of various social and educational backgrounds. This exposed latent social and educational inequalities and reinforced them even further, and brought about "economic, social, cultural, emotional and digital divides" (Tarabini, 2021:6). If not considered carefully, digital technology can be easily used to perpetuate and reinforce the same inequalities and can contribute to what Tarabini calls the crisis of meaning of the schools in the sense of "crisis of the very process of socialisation and cultural transmission" (Tarabini, 2021:10).

Furthermore, in exploring digital maturity of schools, recent studies have shown that the schools themselves often lack the vision and the strategic approach towards technology, including how to harness and where to lead digital infrastructure and skills already present to some extent in their schools. The research commissioned by the British government Department for Education (DfE) revealed that schools are better prepared in terms of infrastructure, equipment and skills but are missing the strategic direction, documentation, and discussion on how to strategically position technology to serve teaching and learning[1] (CooperGibson Research, 2022). Similar findings were found when analysing the digital maturity of schools based on the Framework for Digitally Mature Schools (FDMS) in Croatia. The results indicated that schools were not aware of the importance and use of overall strategies for technology adoption, while on the other hand having good developmental school ICT plan and programme proved to be an important missing element[2] (Balaban et al., 2018).

There is a lack of clear vision in schools of why schools need technology and which strategic direction to take, what kind of technologies and tools and how appropriate they are for students, how do we afford it in the long run, and consequently, what kind of continuous professional development teachers need, including both technical skills and pedagogical knowledge.

This reveals a need for comprehensive vision, followed by a strategic direction, including planning, management and leadership when using digital technology in school education and thinking about digital transformation in education. Therefore, to follow in the same path as education we want to witness, we claim that digital transformation should take the system-wide ecological perspective (Harrison et al., 2014) and also be transformative, inclusive, visionary, effective and accountable process, mindful of the social changes, having in mind the cognitive and social development education is aimed to achieve. This is especially relevant to school education, if the purpose of schools in the 21st century is still, if not even

1 „Schools' average scores in terms of their strategy (mean score of 0.27) was lower than their progress in terms of technology and capability (mean scores of 0.58 and 0.62 respectively)" (CooperGibson Research, 2022:p7).

2 ICT plan was found to be weekly or moderately correlated with almost all elements within the FDMS matrix: "Plan and programme of school development from ICT perspective (PML2) weakly correlates with the Management of the ICT integration into school's business activities (PML4) (r_s=.317, $p < 0.01$), the Learning analytics (PML5) (r_s =..391, $p < 0.01$), and with the Use of ICT in teaching students with special educational needs (PML7) (r_s =..293, $p < 0.01$). Moderate correlation is found with the Management of the ICT integration into learning and teaching (PML3) (r_s =.485, $p < 0.01$)." (Balaban et al., 2018:p9).

more, to ensure socialising and teaching of children and young people (Tarabini, 2021).

Digital maturity is a concept capable of having a holistic view on technology in education because it brings the "social aspects of technology acceptance" (Harrison et al., 2014: p. 1). Digital maturity supports and reflects upon digital transformation of educational institutions across a range of organisational, infrastructural, teaching and learning, competency and cultural issues. It is a "valuable proxy for indicating the extent of technology adoption across the whole eco-system of a school." (Harrison et al., 2014:page 346). It can and indeed should reflect and warn about unequal and insufficient resources schools have in using digital technologies for meaningful instructional and learning activities, both in terms of infrastructure and competencies, but also leadership capacities, to name only the most relevant for the topic of this article.

Digital maturity models serve as a roadmap, a guideline, support for institutional planning, organisational change and benchmarking, support continuous improvement and open space for reflection and understanding of the wider, system context (Marshal, 2010, 2012; Volungevičienė et al., 2021). Digital maturity models connect the complex variables and provide complex narratives of interventions over specific duration, by capturing complexity of evidencing the effects of ICT in educational settings including activities, processes, technologies and skills (Sanchez-Puchol et al., 2018; Underwood et al., 2004, 2010). It has been seen as a potent opportunity for a long-term assessment of digital transformation of institutions as well as large-scale national technology investments, seeking the evidence of the benefits of technological investments across educational systems. See for example Framework for Digitally Mature Schools in Croatia (Balaban et al., 2018) or Framework for Digitally Competent Organisations (Kampylis et al., 2016).

4 Educational Leadership for Digitally Mature Education

The last section proposes the leadership schools need for making meaningful digital transformation and for growing towards digitally mature educational organisations.

We start from a classical definition of leadership as "the process of influencing the activities of an organised group in its efforts toward goal setting and goal achievement" (Stogdill, 1950). We acknowledge that the very general meaning of leadership is the process of influence. If the goal we seek is digital transformation that support the school's vision of education, than what kind of leadership schools should practice?

We argue that (1) the distribution of leadership authority (Dexter, 2018; Gumus et al, 2016; Leithwood et al., 2008; Voogt et al., 2018) is one of the major prerequisites for the influence towards meaningful digital transformatiooon to happen in schools, motivated by (2) a transformational vision for self-actualisation and achievement (Daniëls et al, 2019; Leithwood and Jantzi, 2005) of the individuals, schools, community and the wider system. This claim is supported by Joke Voogt and a group of authors in stating that distributed leadership perspective, together with transformative and instructional leadership actions, are framing main functions of leading technology innovations: developing vision, supporting integration, ensuring the accountability of technology initiatives (Voogt et al., 2018). Distributed and transformational leadership perspectives support the vision of education (described in the first section) in a number of ways:

- To be able to support the variety of paths in which any person and/or community could develop, the leadership influence should be inclusive and participatory. When leadership influence and power are distributed to several individuals, groups and teams without formal leadership positions (usually teachers) individual capacities are enhanced through social interactions. In this distribution the meaning, knowledge and leadership practice are socially constructed and are co-performed (Dexter, 2018; Harris, 2009).
- To be able to transform, this influence should be able to guide and motivate all to strive for individual, school and system improvements. The main role of transformational leaders is to motivate followers to open to self-actualisation and lead to a vision for the school (Leithwood and Jantzi, 2005; Bush and Glover 2003).
- To be able to be effective, this influence should have a strong vision on teaching and learning and students' engagement and achievements in the core of these processes, and distributed leadership has positive effects on teacher effectiveness, student outcomes and student engagement (Harris, 2009; Leithwood et al., 2008).
- To be able to be accountable, this influence should have strong responsibility towards the public good it serves and the public resources it uses, "ensuring the accountability of technology initiatives in terms of ownership and outcomes" (Voogt et al., 2018).
- Finally, to be able to consider the larger context of a global, environmental and digital transformations (Peters and Jandrić, 2017), this influence should take into account the visions of many different people and open the space for their opinions to be exchanged and new meanings and knowledge created in this exchange, but from the perspective of *interactions* rather than actions and states (Harris and De Flaminis, 2016).

To support the liberating and transformative process of a person's development, we claim that less directive and more bottom-up approach to educational

leadership is needed, the one that supports different paths of personal and social transformation and goals achievements. We argue that two leadership approaches should be combined if this influence is to bring fruits: transformational and distributed leadership.

Distributive leadership offers the co-created inclusive methodology while transformational leadership inspires, guides and motivates all to strive for individual and school improvements through cognitive and social development of a person and of a community. In this distributed process of influencing, which is not a directive but inclusive, two-way process, different visions are shared, and common meanings transformed through co-creation. This shared school vision is mindful about both personal and community's potential and is adjusted to changing needs and environment while influenced by school staff, students and community stakeholders.

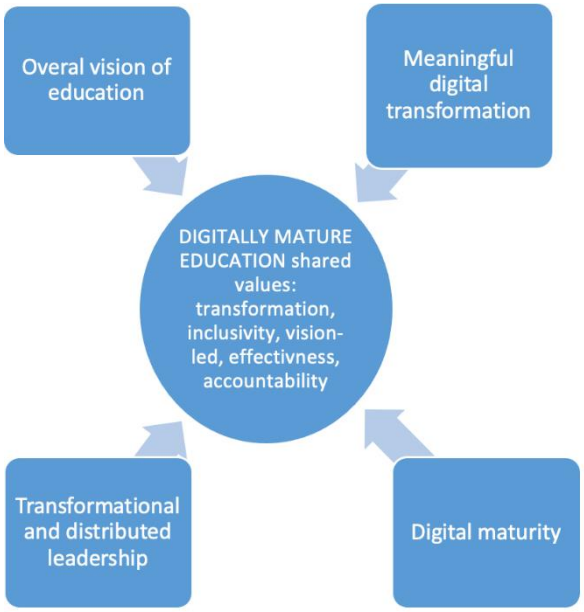


Figure 1. Digitally mature education concept

5 Conclusion

We have asked ourselves what the role of education is when considering global, environmental and digital transformations happening today. We acknowledge that education should be mindful of the wider changes but should always continue to support cognitive and social development of a person who is then able to become active and contributing citizen of the same changing and troubled world.

We acknowledge digital transformation as an important lever for supporting the aforementioned vision of education. We ask what kind of meaningful digital transformation can support a cognitive and social development of a person and her community? We propose to use digital maturity as an organisational improvement framework and a reflective tool designed to guide meaningful digital transformation of education from the system-wide ecological perspective into transformative, inclusive, visionary, effective and accountable process of cognitive and social development of individuals and communities (Harrison et al., 2014; Marshal, 2010, 2012; Teichert, 2019; Underwood and Dillon, 2004; Volungevičienė et al., 2021).

Lastly, we asked what kind of leadership schools need to digitally mature in a way that is meaningful for them and their community? We argue that if this process is to be meaningful, we need educational leadership that is an exchanging process of influence and relationship leading to a shared vision and a common purpose for the school, that is mindful about personal and community's potential and is adjusted to changing needs and environment, co-created by school staff, students and community stakeholders. To be able to support the variety of paths in which any person and/or community could develop, the leadership influence should be inclusive and distributed. To be able to transform, this influence should be able to guide and motivate all to strive for individual, school and system improvements.

We conclude with the idea that distributing leadership to all stakeholders (including teachers and students) in reflecting and devising meaningful digital transformation using digital maturity concept should move schools and educational systems closer towards digitally mature (school) education, in which shared transformative and inclusive vision of education, mindful of wider global, environmental and digital changes leads to cognitive and social development of a person and of a community. Digitally mature education concept, presented in Fg.1., is deeply rooted in the transformative, inclusive, visionary, effective and accountable vision of education, and secondarily supported by meaningful digital transformation. It is based on the shared values between overall vision and role of education on one side, and the role of meaningful digital transformation in education on the other. The shared values are transformation, inclusivity, vision-led, effectiveness and accountability. Both education and digital transformation should share these values and strive towards the same vision of education, and that is to support cognitive and social development of a person and her community. Education and digital transformation should not have disjointed goals and practices, on the contrary, we should strive to make digital transformation increasingly a meaningful part of education and training system.

We propose that digitally mature education is to be co-created and co-lead with transformative and distributed practices in schools while using digital maturity as an improvement and reflective framework for guiding meaningful digital transformation. Leading

towards digitally mature education should be done through transformational and distributed practices, using clear guidelines of digital maturity.

We propose further research that will investigate the digital maturity as an improvement and a reflective framework for meaningful digital transformation of education, leading to a digitally mature education. In doing this, the further research should aim at better understanding the role of leadership in digital maturation of (school) education.

References

Anderson, J., & Boyle, C. (2020). "Good" Education in A Neo-Liberal Paradigm: Challenges, Contradictions and Consternations. In C. Boyle, J. Anderson, A. Page, & S. Mavropoulou (Eds.), *Inclusive Education: Global Issues and Controversies*. Leiden ; Boston Brill Sense. https://doi.org/10.1163/9789004431171_003

Aristoteles. (2001). *On the soul and On memory and recollection* (J. Sachs, Trans.). Green Lion Press.

Balaban, I., Redjep, N. B., & Čalopa, M. K. (2018). The analysis of digital maturity of schools in Croatia. *International Journal of Emerging Technologies in Learning*, 13(6), 4–15. https://doi.org/10.3991/ijet.v13i06.7844

Bush, T., & Glover, D. (2014). School leadership models: What do we know? *School Leadership and Management*, 34(5), 553–571. https://doi.org/10.1080/13632434.2014.928680

Bush, T., & Glover, D. (2003). School Leadership: Concepts and Evidence. National College for School Leadership, (Spring), 1–42.

CooperGibson Research. (2022). *Education technology : exploring digital maturity in schools*. Retrieved August 16, 2022 from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1061797/Exploring_digital_maturity_in_schools.pdf

European Commission. (2020a). *Digital Education Action Plan 2021-2027 Resetting education and training for the digital age.* Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Retrieved August, 16, 2022 from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0624&from=EN

European Commission. (2020b). *Commission Staff Working Document.* Accompanying the document Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions Digital Education action Plan 2021-2027 Resetting education and training for the digital age. Retrieved August, 16, 2022 from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020SC0209

Daniëls, E., Hondeghem, A., & Dochy, F. (2019). A review on leadership and leadership development in educational settings. *Educational Research Review*, 27(March), 110–125. https://doi.org/10.1016/j.edurev.2019.02.003

Dewey, John (1916). *Democracy and Education: An Introduction to the Philosophy of Education*. New York: Macmillan. Retrieved from https://archive.org/details/democracyandedu00dewegoog/page/n6/mode/2up?view=theater

Dexter, S. (2018). The Role of Leadership for Information Technology in Education: Systems of Practices. In J. Voogt, G. Knezek, R. Christensen, K-W. Lai (Eds.), *Second Handbook of Information Technology in Primary and Secondary Education*. Springer International Handbooks of Education.

European Commission. (2020). Digital Education action Plan 2021-2027: Resetting education and training for the digital age. In *Commission Staff Working Document*.

Freire, Paulo (1970). Pedagogy of the oppressed. New York: Herder and Herder.

Gumus, S., Bellibas, M. S., Esen, M., & Gumus, E. (2018). A systematic review of studies on leadership models in educational research from 1980 to 2014. *Educational Management Administration and Leadership*, 46(1), 25–48. https://doi.org/10.1177/1741143216659296

Harris, A. (Ed.). (2009). *Distributed Leadership: Different Perspectives* (Studies in Educational Leadership, Vol. 7). Springer Science+Business Media B.V. 2009.

Harrison, C., Tomás, C., & Crook, C. (2014). An e-maturity analysis explains intention-behavior disjunctions in technology adoption in UK schools. *Computers in Human Behavior*, 34, 345–351. https://doi.org/10.1016/j.chb.2013.10.042

Hodges, C.B., Moore, S., Lockee B.B., Trust, T. and Bond, M.A. (2020). *The Difference Between Emergency Remote Teaching and Online Learning*. EDUCAUSE Review. Retrieved August, 16, 2022 from https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Jaakkola, E. (2020). Designing conceptual articles: four approaches. *AMS Review*, 10(1–2), 18–26. https://doi.org/10.1007/s13162-020-00161-0

Kampylis, P., Devine, J., Punie, Y., & Newman, T. (2016). *Supportinig school to go digital: From a conceptual model towards the design of a self-assessment tool for digital-age learning*. ICERI2016 Proceedings. 9th International Conference of Education, Research and Innovation. November, 4–16. 2017. Seville, Spain.

Leithwood, K., & Jantzi, D. (2005). A review of transformational school leadership research 1996–2005. Leadership & Policy in Schools, 4(3), 177–199. https://doi.org/10.1080/15700760500244769

Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. School Leadership & Management, 28(1), 27–42. https://doi.org/10.1080/13632430701800060

Marshall, S. (2012). Improving the quality of e-learning: Lessons from the eMM. *Journal of Computer Assisted Learning*, *28*(1), 65–78. https://doi.org/10.1111/j.1365-2729.2011.00443.x

Marshall, S. (2010). A Quality Framework for Continuous Improvement of E-learning: The E-learning Maturity Model. *Journal of Distance Education Revue De L'Éducation À Distance*, 24(1), 143–166.

Michelsen, G., Fischer, D. (2017). Sustainability and Education. In M.V.Hauff and C. Kuhnke (Eds.), *Sustainable Development Policy. A European perspective* (pp. 135-158). London: Routlege. Retrieved June 13, 2022 from https://www.routledge.com/Sustainable-Development-Policy-A-European-Perspective/Hauff-Kuhnke/p/book/9781138284999

Peters, M. A., & Jandrić, P. (2017). Dewey's Democracy and Education in the age of digital reason: the global, ecological and digital turns. *Open Review of Educational Research*, 4(1), 205–218. https://doi.org/10.1080/23265507.2017.1395290

St. John, C. (Ed.). (1931). *Ellen Terry and Bernard Shaw: A Correspondence.* G.P.Putnam's Sons, Ltd.

Stogdill, R. M. (1950). Leadership, membership and organization. *Psychological Bulletin*, 47, 1-14

Tarabini, A. (2021). The role of schooling in times of global pandemic: a sociological approach. *International Studies in Sociology of Education*, *00*(00), 1–19. https://doi.org/10.1080/09620214.2021.1966825

Teräs, M., Suoranta, J., Teräs, H., Curcher M. (2020). Post-Covid-19 Education and Education Technology 'Solutionism': a Seller's Market. *Postdigit Sci Educ 2*, 863–878 (2020). https://doi.org/10.1007/s42438-020-00164-x

Toh, Y., & So, H. J. (2011). ICT reform initiatives in Singapore schools: A complexity theory perspective. *Asia Pacific Education Review*, 12(3), 349–357. https://doi.org/10.1007/s12564-010-9130-0

Townsend, T. (2011). School leadership in the twenty-first century: Different approaches to common problems? *School Leadership and Management*, 31(2), 93–103. https://doi.org/10.1080/13632434.2011.572419

Volungevičienė, A., Brown, M., Greenspon, R., Gaebel, M. & Morrisroe, A. (2021). *Developing a High-Performance Digital Education System: Institutional Self-Assessment Instruments*. European University Association absl. Retrieved from https://eua.eu/downloads/publications/digi-he desk research report.pdf

Voogt, J., Knezek, G., Christensen, R., & Lai, K.-W. (2018). Developing an Understanding of the Impact of Digital Technologies on Teaching and Learning in an Ever-Changing Landscape. In J. Voogt, G. Knezek, R. Christensen, & K.-W. Lai (Eds.), *Second Handbook of Information Technology in Primary and Secondary Education*. Springer International Handbooks of Education.

Underwood, J., & Dillon, G. (2004). Capturing complexity through maturity modelling. *Technology, Pedagogy and Education*, *13*(2), 213–225. https://doi.org/10.1080/14759390400200181

ICT Entrepreneurship and Innovation

Miroslav Mandić, Iva Gregurec, Ema Galović
Employee's Perception of CRM in the Croatian Telecommunications Industry

Eva Malichová, Peter Márton, Emese Tokarčíková
Blended Mobility on Entrepreneurship – Intensive Programme for ICT Students in the HEI Curriculum

Employee's Perception of CRM in the Croatian Telecommunications Industry

Miroslav Mandić
Faculty of Economics and Business
University of Zagreb
Trg J.F. Kennedy 6, 10000 Zagreb, Croatia
{mmandic}@efzg.hr

Iva Gregurec
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{iva.gregurec}@foi.unizg.hr

Ema Galović
FIVE, an Endava company
Heinzelova ulica 33, 10000 Zagreb, Croatia
{emagalovic13}@gmail.com

Abstract. *Companies in today's market, especially in the service sector, require constant monitoring and adaptation to the more dynamic needs and desires of their customers. Therefore, companies use all their resources to achieve their goals and competitive advantage. This paper shows a theoretical review of CRM concept with a focus on the telecommunications industry and the results of quantitative research conducted through a questionnaire, on a purposive sample, a sample of experts. The aim of this paper is to show, whether there is a difference in the perception of CRM by 170 employees at different managerial levels in selected telecommunications companies in Croatia.*

Keywords. Customer Relationship Management (CRM), technological solution, customer-oriented strategy, employees, telecommunications sector

1 Introduction

Customers are becoming more powerful, especially with the increasing use of the Internet, which gives them access to a wealth of information (Chen & Popovich, 2003) and thus increasing their needs and wants. So, marketing, since the industrial revolution, has changed from transactional marketing to relationship marketing, which instead of a single transaction and mass access focuses on an individual customer and a personalized approach. And because of that, the imperative for companies has become to build more intimate and valuable relationships with their customers, which the entire company, and all departments of a company, needs to work on in an integrated way. Customer Relationship Management (CRM) represents one concept that can help to achieve that. Because it aims to build and maintain long-term, mutually profitable relationships, which is a very effective business strategy because it gives a company the opportunity to serve its customers satisfactorily, which makes them satisfied and loyal to the company. A satisfied and loyal customer brings the company many advantages, and building a relationship with them brings a competitive advantage for a company. Therefore, it is important for a company to consistently be focused on maintaining customers and building relationships with them. This can be achieved only if everyone in the company shares the same value, same vision, same pursues the same goals and the same strategy.

Therefore, it is necessary to emphasize that the acquisition of real customers is just as important as the acquisition of real employees, who will meet the company's goals through communication and interaction with customers and lead to mutual satisfaction. The impression that employees leave on customers, and the mutual coordination of everyone in the company in order to achieve the company's goals, are important. Even though each employee works for themselves, in their field of specialization, the joint effort of each individual and the team within which they work is crucial. Employees in the "backstage", i.e., in the organizational sense "back office" are those whose work is invisible to customers, and who provide support to everyone in the company in their daily activities. Departments such as finance, administration, logistics, human resources, etc., are included here (Chen & Popovich, 2003). These employees are responsible for the successful functioning of all "background" processes that contribute to better business operations. Employees in the "frontstage", i.e., in the organizational sense of the "front office" are directly in front of the eyes of customers, so their honesty, dedication, discipline, and communication skills are key to successful interactions with customers. In this case, departments such as marketing, sales, and

customer support are involved (Chen & Popovich, 2003). Both aim to deliver quality and achieve customer satisfaction, while sharing customer information between different departments and decision-making levels. The literature states that both, the back and front office, have in common the result, i.e. the perception of quality and satisfaction of the customers with the service (and product), deserves the teamwork of all (employees in the company) involved in the performance (Grove & Fisk, 2004) (Xu & Geodegebuure, 2005) add that successful and coordinated teamwork is what will generate satisfaction if it includes strong leadership, effective communication in the company and support from back-stage employees. Building long-term relationships with customers and creating a unique customer experience are strongly linked to building long-term relationships with employees through their identification with key business values and business culture. Therefore, it is important to create a unique culture within the company and develop internal relationships because every external "contact with the customer was influenced by internal relationships in the company (Triznova et al., 2015).

This paper shows a short theoretical review of CRM concepts, and its implementation in the telecommunications industry. For the purposes of this paper, CRM is viewed from two different angles or dimensions, as a technological solution and as a business strategy, i.e. company culture. Also, the results of quantitative research conducted through a questionnaire, as an instrument for collecting primary data, on a purposive sample, a sample of experts at different managerial levels (the employees at the bottom of the decision pyramid who on daily bases communicate directly with end customers and the one at the top of the pyramid who are not in contact with their customers as much) are shown at the end of this paper.

2 Customer Relationship Management

Although the concept of CRM first appeared in the mid-1990s in the IT sector (Dirsehan, 2017) and the reason why CRM in the past has been identified with technology (Chang et al., 2010), customer relationship as a concept dates back to the 1950s when Drucker (1954) put forward the idea that customers should be the foundation and reason for the existence of each company (taken from Payne, 2005). The idea has been accepted in scientific circles, which has led to the development of what is now called the marketing concept and which advocates building customer relationships as an essential part of day-to-day business management (Coltman, 2007). So, over time, a holistic approach to the definition has become more prevalent, according to which a company's CRM strategy and cross-functional process aimed at building and maintaining long-term relationships with customers and creating their loyalty, and technology is considered a valuable tool for easier and more efficient maintenance relationships (Chang et al., 2010) (Eid, 2007) (Zablah et al., 2004) (Nahar & Dhaka, 2014) (Becker et al., 2009). But it should be pointed out that the term CRM is used in many ways and several industries and has multiple meanings. There is no consensus on a specific definition of the term CRM (Nemeček, 2013) (Zablah et al., 2004) states that there are approximately 45 different definitions of CRM in the literature, and it is most commonly characterized as a process, strategy, philosophy, capability or technology.

From the very beginning of defining the term CRM, relationship marketing is mentioned, which expresses the importance of creating long-term relationships between companies and customers, getting to know and adapting products and services to their needs and wants. Relationship marketing theory dictates that customers who perceive that a company has built a relationship with them behave in a way that is profitable for the company. Because they are offered exactly what they want, such customers are willing to buy more products or use more services, spend more, and are loyal to the company, and are unlikely to outperform the competition (Verhoef, 2003). A key component of relationship marketing is precisely the transformation of the customer into the customer by ensuring their loyalty. The fact is that the cost of acquiring a new customer is significantly higher than maintaining the satisfaction of the existing one (Ramya et al., 2013). Relationship marketing is a new level in the development of marketing and is the opposite of transactional marketing, where the focus is primarily on individual transactions, while relationship marketing is focused on retaining customers and long-term added value (Dukić & Gale, 2015).

While some authors equate relationship marketing with the concept of CRM (Zablah et al., 2004), others find relationship marketing as a key component of CRM, along with management and modern information and communication technology (Dukić & Gale, 2015). (Verhoef, 2003) outlines a key difference between them. And according to him, CRM does not focus exclusively on establishing and maintaining close, mutually beneficial relationships, he considers CRM as a strategically oriented process related to the “production” of an ideal mix of customer relationships, while according to him relationship marketing focuses only on the tasks needed to build sustainable customer relationships (Verhoef, 2003). From the above, it can be concluded that CRM is a broader concept than relationship marketing, but that the overall concept of relationship marketing is an important part of the entire CRM. Several authors cite relationship marketing as the starting point or philosophy of CRM (Zablah et al., 2004) (Ryals & Knox, 2001).

Today in practice, many companies have realized that the critical success factors are no longer individual transactions, but building long-term relationships with customers, which CRM allows them to achieve (Kamrul Islam Shaon & Rahman, 2015). The basic idea of CRM is to use technology and human resources to understand customer behavior, values, and attitudes (Eid, 2007). (Ryals & Knox, 2001) stated that if relationship marketing is adopted as an organizational philosophy that emphasizes customer retention, CRM can be considered representative of the organizational implementation of such a philosophy. So, for this paper it is important to now the difference if CRM is considered as a technological solution or as a business strategy, i.e. company culture.

3 Usage of CRM in the telecommunications industry

Since the service sector is characterized by numerous specifics, the most prominent of which is intangibility, the human factor is extremely important in the customer's perception of service quality and creating relationships with the company. Therefore, it is the employees who represent the points of contact between the company and its customers, and their focus on customers is crucial when communicating and interacting, in order to deal with specific inquiries (Xu & Geodegebuure, 2005) (Eichorn, 2004) (Hennig-Thurau, 2004). Therefore, their behavior, attitude, knowledge, and honesty directly affect the customers' perception of the service and their satisfaction level (Hennig-Thurau, 2004). (Fleming et al., 2005) state that any interaction an employee has with a customer represents an opportunity to build or reduce emotional connection. Employees, therefore represent, as a human factor of the company, a key component of CRM success (Malthouse et al., 2013) (Triznova et al., 2015).

The telecommunications industry is a pure service sector, so its success depends entirely on customer satisfaction and loyalty (Rahaman et al., 2011). And because of that, the ultimate goal of CRM in the telecommunications sector is to provide a comprehensive set of applications that will enable greater productivity and customer satisfaction as it coordinates and synchronizes customer interactions across all touchpoints including web, contact center, field, and distribution channels (Mohammed, 2013). Buttle (2004) states that CRM is not just a method of running service companies to achieve competitive advantages, but it is necessary for their survival (taken from Rootman et al., 2007). (Foss et al., 2008) state that telecommunications companies must reduce customer churn while increasing value for customers through up-selling and cross-selling tools is considered as 'Icing on the cake'.

Authors (Rahaman et al., 2011) pointed out difficult times in the telecommunications industry. The reason for that is the fact that markets are saturated, debts are growing and innovations are becoming more expensive. Therefore, it is necessary to create a strategy whose goals will be higher profitability, a lower outflow of customers (customer churn), achieving customers greater satisfaction, and reduced costs. Since the telecommunications sector is characterized by strong competition, retaining existing customers is becoming more important than acquiring new ones. In addition to being more profitable for the company, a satisfied and loyal customer brings many other benefits as well. These include, among other things, a very low probability of competition, a positive "word of mouth", a willingness to buy more and generate more revenue and provide valuable inputs for further business decisions (Kamrul Islam Shaon & Rahman, 2015).

To be successful and competitive in the market, the telecommunications company needs to operate more effectively than its competition, employees need to be more productive, and services in line with the needs and wants of their customers. At the strategic level, the right decisions need to be made faster than the competition. (Nikolić et al., 2014) state that this requires information - about business activities and about existing and potential customers. For the effective handling of this information, companies choose a technological solution, i.e. a software tool that will support customer-oriented business and customer relations. The importance of the good flow of information through all departments in the company has been mentioned before, so that information obtained from customers. For example, the call center can reach everyone relevant, from other sales departments, marketing, customer support and, ultimately analytics and reporting. CRM as a technological solution facilitates achieving customer satisfaction because it enables a better perception of customers, their greater accessibility and more effective interaction, which (Brodarić, 2010) pointed out as a key of build long-term relationships with customers. He adds that the basic idea of CRM is not only the orientation of the company towards services and the quality of service in the technical sense but the care for the customer and the relationship with them (Brodarić, 2010).

CRM as a technological solution enables communication, facilitates the handling of information about customers, and is a central place for data collection and storage, which enables the delivery of relevant information for business decisions. Organizational culture in the company is important for the quality implementation of the CRM strategy, which aims to build and maintain long-term relationships with customers. Customer-oriented organizational culture means a quality flow of information throughout the company, the involvement of employees at all levels of management and their mutual quality communication

and focuses on the same goal. This can be characterized by a unique view of all employees in the company on CRM, both on the technological solution and the business philosophy of the company.

4 Research study

4.1 Research methodology

The subject of this research is the application of CRM as a technological solution or business strategy in the telecommunication industry in the Croatian market. At the same time, the goals of the research were set: examine the perception of CRM by employees in selected telecommunications companies in Croatia, determine whether there is a difference in perception between different levels of management, determine whether there is a difference in the perception of CRM as a technological solution and the perception of CRM as customer-oriented business culture and management (strategy). The goals were set based on the collection of secondary data sources as a precondition that if there are differences in perception of employees there is no holistic approach of the company to CRM in the telecommunications sector because at different levels of management employees perceive different technological solutions and business culture.

Quantitative research was conducted through a questionnaire as an instrument for purpose of collecting primary data. A questionnaire was sent via e-mail to employees on different levels of management in selected telecommunication companies in Croatia, so research was conducted on a purposive sample, a sample of experts. A total of 170 respondents fully responded to all the questions of the questionnaire and their answers were analyzed below in this paper.

For the purposes of writing this paper, management levels were initially defined according to (Severović, 2013) and (Dukić & Gale, 2015) under the frequency of contact with the end customers (Table 1.):

Table 1. Managerial levels important for setting up research methodology

MANAGERIAL LEVELS	PARTICIPANTS BY LEVELS	BASIC TASKS BY LEVELS	DEGREE OF CONTACT WITH CUSTOMERS	FREQUENCY OF CONTACT WITH CUSTOMERS
Strategic level	The highest level of heads of organizational units	Increase profitability, market share and reduce costs by resource allocation	Low contact with customers except in dealing with specific requirements	Never or several times a year
Tactical level	The lower level of heads of organizational units	Shortening the time of processing customer requests, shortening the time of sales, receiving information and further distributing of information to lower and higher levels, etc.	Middle contact with customers, resolving issues of realization of individual issues when sales staff does not have the information needed for the final solution	Several times a month
Operational level	Sales units, call centers	Sales, connecting organizational units, communication with the customers	High, constant contact with customers	Daily or several times a week

Source: adapted to Severović, K. (2013) Upravljanje odnosima s klijentima kao izvor informacija za oblikovanje i poboljšanje usluge, Varaždin: Fakultet organizacije i informatike, p. 94.

Since the frequency of contact with the end customers is relevant for research, i.e. examining the difference in perception of employees who do not have contact with the end customers from those who have, for research purposes management levels are divided into strategic and operational levels, while the tactical level is partially covered by strategic management.

In the first part of the questionnaire, respondents were asked if they are familiar with the CRM concept. Although at the beginning of the research it is assumed that CRM, as a technological solution, is unified for all telecommunications companies according to secondary data, there was some possibility that some of the companies included in the research sample use other technological solutions that are not relevant to this research. According to the results, 94.1% of them stated that they are familiar with the term CRM.

The second part of the questionnaire refers to the contribution of CRM to the way of doing business. The questions in the questionnaire, i.e. the measurement scales, were taken from (Chang et al., 2010), who took the questions mainly from (Jayachandran et al., 2005). A 7-point Likert scale consisting of 16 questions was used for the examination. It referred to CRM as a technological solution consisting of four components: sales support, customer service support, analysis support, and data integration and access support. The measuring instrument proved to be reliable based on the calculated Cronbach's alpha coefficient ($\alpha = 0.896$).

The third part of the questionnaire referred to the business culture of the company, i.e. the culture that is aimed at end customers and the encouragement of such a culture by management (management system focused on a business culture that is oriented to customers). Questions and measurement scales were also taken over from (Chang et al., 2010). The Likert scale of 7 degrees was used, and this part of the questionnaire consisted of 6 questions. It referred to the organizational culture of the company focused on creating relationships with customers and the management system that encourages the creation of relationships with customers. The measuring instrument again proved to be reliable based on the calculated Cronbach's alpha coefficient ($\alpha = 0.903$).

An analysis was also performed depending on the level of management where employees are (in this case, respondents) in order to determine whether there is a difference in the perception of the contribution of CRM as a technological solution to business operations and the perception of CRM as customer-oriented business culture and management (strategy). The statistical analysis below determines whether there is a significant difference between the two independent groups of respondents. For this purpose, the independent sample t-test in the statistical tool SPSS was used.

4.2 Research results

The questionnaire was sent to employees of Iskon Internet, Hrvatski Telekom (including the Bonbon brand), Optima Telekom, Vipnet, Bnet, Tele2, BTnet and Terrakom. That means, the questionnaire covered all telecommunications companies in the Croatian market. A total of 170 full responses were collected, mostly from Iskon Internet (82.4%), Vipnet (11.8%) and Hrvatski Telekom (4.7%). A smaller number of responses was collected from BTnet (0.6%) and others (0.6%). Respondents from other telecommunications companies, despite being sent a questionnaire and a request to participate in research, were not interested and / or available to cooperate.

According to the demographic characteristics of respondents (Table 2.), a slightly higher percentage of respondents were male (55.3%), while 44.7% were female. The highest percentage of respondents were between 26 and 34 years old (41.2%), followed by respondents between 18 and 25 years old (38.2%). A smaller percentage of respondents were between 35 and 50 years old (20%) and only one respondent was between 50 and 65 years old (0.6%). When examining work experience, the current workplace was covered, as well as student jobs. The highest percentage of respondents have between 1 and 5 years of work experience (45.3%) in the telecommunications industry, followed by those between 6 and 10 years of experience (25.3%). They are followed by employees with between 11 and 15 years of experience (14.1%) and more than 15 years of experience (13.5%). The lowest percentage of employees has less than 1 year of work experience (1.8%) in the telecommunications industry.

Table 2. Demographic characteristics of the respondents

DEMOGRAPHIC CHARACTERISTICS		FREQUENCY (n)	PERCENTAGE (%)
Sex	Male	94	55,3
	Female	76	44,7
Age	18-25	65	38,2
	26-34	70	41,2
	35.50	34	20
	50-65	1	0,6
Work experience in the telecommunications industry	Less than 1 year	3	1,8
	1-5 years	77	45,3
	6-10 years	43	25,3
	11-15 years	24	14,1
	More than 15 years	23	13,5

Source: own research

The majority of respondents did not work for other telecommunications companies (59.4%), while 22.4% of them worked in another telecommunications company. 12.4% of respondents worked in 2 other telecommunications companies, while 5.9% of respondents until now worked in 3 or more other telecommunications companies. And it is important to mention that a higher percentage of employees work in operational management (58.8%), which

communicates daily or several times a week with end customers of the company.

When examining the contribution of CRM to the company's business, the respondents were given statements related to CRM as a technological solution. The overall assessment of the respondents on the CRM dimension of the technological solution was obtained using the average value of the corresponding statements (Table 3.). Respondents gave the highest average grade to CRM-related statements when providing relevant information about end customers and interacting with them, primarily sales and sales support, and then employees in general. They are followed by statements related to monitoring the delivery of products and services. Respondents gave average ratings to CRM-related statements when introducing sales and sales support to potential additional sales to existing customers, integrating all information gathered from different customer communication channels, and measuring loyalty, retention, and customer value for the company. Slightly lower average grades were given by respondents to CRM-related statements when comparing information on customer transactions with information from external sources and determining the profitability of products and services. Respondents gave the lowest average grade to the functionality of the CRM as a technology solution when accessing the knowledge base by answering frequently asked questions (FAQs) and providing relevant information about the competition.

Table 3. Dimension – CRM as a technological solution

STATEMENTS RELATED TO CRM AS A TECHNOLOGICAL SOLUTION	AVERAGE GRADE	MEDIAN	STANDARD DEVIATION
CRM provides sales and sales support information about customers.	5,8	6	1,3
CRM allows employees to view relevant information about customer.	5,6	6	1.3
CRM provides sales and customer support with access to information about communication (interaction) with the customer.	5,5	6	1,5
CRM monitors service delivery.	5,2	6	1,5
CRM allows you to track product delivery.	5,1	5	1,5
CRM introduces sales and sales support to potential additional sales to existing customers.	4,7	5	1,5
CRM integrates information about customers from different channels of communication with customers (e.g. e-mail, telephone, web, fax, etc.).	4,6	5	1,5
CRM provides insight into customer loyalty.	4,4	5	1,8
CRM provides sales and sales support with a specific offer tailored to each customer.	4,3	4	1,6
CRM provides insight into the degree of customer retention.	4,3	4	1,7
CRM allows you to predict the future needs of the customer.	3,8	4	1,8
CRM calculates the value of each customer for the company (Customer Life Value - CLV).	3,6	4	1,7
CRM compares customer transaction information with information from external sources.	3,4	4	1,6
CRM determines the profitability of products and services.	3,4	4	1,7
CRM provides access to a knowledge base of answers to frequently asked questions (FAQs).	3,3	3	1,8
CRM enables the sale of relevant competition information.	3,2	3	1,7

Source: own research

When examining the business culture of the company, the respondents were given statements related to the business culture of the company in which they are employed. The overall assessment of the respondents on the business culture dimension of the company was obtained using the average grade of the corresponding statements (Table 4.). Respondents gave the highest average grade to statements related to the company's attitude that retaining existing customers is a priority and that long-term customer relationships are valuable assets. Respondents gave average ratings to the statements related to encouragement of employees to focus on building long-term customer relationships, emphasizing the importance of long-term customer relationships, and providing an employee education program that would

contribute to better interactions with existing customers. Respondents gave the lowest average grades to statements related to providing incentives to employees based on the satisfaction of existing customers and assessing employee contact with customers based on the quality of relationship building.

Table 4. Dimension - CRM as customer-oriented business culture and management (strategy)

STATEMENTS RELATED TO CRM AS CUSTOMER-ORIENTED BUSINESS CULTURE AND MANAGEMENT	AVERAGE GRADE	MEDIAN	STANDARD DEVIATION
The company considers retaining existing customers a priority.	5,5	6	1,4
The company considers long-term relationships with an existing customer to be a valuable asset.	5,4	6	1,4
The company encourages employees to focus on creating long-term relationships with an existing customer.	5,1	5	1,5
The company emphasizes the importance of long-term relationships with existing customers.	5,1	5	1,5
The company provides an educational program for employees in order to improve the quality of interaction with existing customers.	4,7	5	1,7
The company provides incentives to employees based on the satisfaction of existing customers (e.g. bonuses, commendations, awards, etc.)	4,5	5	1,9
The company evaluates employees for customer contact based on the quality of building relationships with existing customers	4,3	4	1,7

Source: own research

Out of a total of 170 respondents who participated in this research, 100 (58.8%) of them represent operational management, while the remaining 70 (41.2%) represent tactical and strategic management. Table 5. shows the descriptive values of management levels on the dimensions of CRM as a technology solution and CRM as customer-oriented business culture and management. This represents an input for conducting the independent samples t-test, which will determine whether there is a significant difference in the perception of employees on the observed dimensions according to the level of management.

Table 5. Descriptive values of management levels based on dimensions set in the research

OPERATIONAL / TACTICAL AND STRATEGIC MANAGEMENT		SAMPLE SIZE	AVERAGE GRADE	STANDARD DEVIATION	STANDARD AVERAGE ERROR
CRM as a technology solution	Operational management	100	4,2544	1,04459	0,10446
	Tactical and strategic management	70	4,5518	0,93378	0,11161
CRM as a business culture	Operational management	100	5,0586	1,19817	0,11982
	Tactical and strategic management	70	4,7918	1,32160	0,15796

Source: own research

Table 6. shows the results of the conducted t-test of independent samples. Since the values obtained by performing Levene's test of equality of variance are greater than 0.05, the t-test of independent samples assumes equal variance and the reference values are those of the above order. On the CRM as technological solution dimension, the result of the t-test analysis is marked in red (t-test = - 1.907), and the p-value in green (p-value = 0.058 → 0.058> 0.05). According to the results of the t-test analysis, it is determined that there is no statistically significant difference between the levels of management in the perception of CRM as a technological solution. In the business culture dimension of the company, the result of the t-test analysis is marked in orange (t-test = 1.369), and the p-value in blue (p-value = 0.173 → 0.173> 0.05). According to the results of the t-test analysis, it is determined that there is no statistically significant difference between the levels of management in the perception of the business culture of the company.

Table 6. Independent samples t-test

LEVENE'S TEST OF EQUALITY OF VARIANCE		T-TEST OF EQUALITY OF ARITHMETIC MEANS OF TWO POPULATIONS						
F-test	P-value	T-test	Degrees of freedom	P-value	Difference in arithmetic means	Std. Error difference	95% confidence interval difference	
							Lower end	Upper end
0,602	0,439	-1,907	168	0,058	-0,29741	0,15593	-0,60525	0,01042
		-1,946	158,217	0,053	-0,29741	0,15287	-0,59933	0,00451
0,590	0,443	1,369	168	0,173	0,26673	0,19485	-0,11794	0,65141
		1,345	139,138	0,181	0,26673	0,19826	-0,12526	0,65873

Source: own research

4.3 Research discussion

In accordance with the results of respondents on both observed dimensions, it can be concluded that employees perceive the greatest contribution to the business of telecommunications companies that CRM as a technology solution allows them access to relevant customer information, while the lowest grade, i.e. the perception of the smallest contribution of CRM access to a knowledge base of frequently asked questions and providing information on the competition. Also, according to the given assessments, the employees perceive that the telecommunications company in which the employees are focused on building relationships with end customers consider these relationships as a valuable asset. Employees gave the lowest average grade to the incentives provided by the company in building and improving relationships with existing customers.

According to the analyzed research results, it was found that there is no significant difference in the perception of CRM on both observed dimensions. The results indicate the existence of a unified perception in selected companies in the telecommunications sector in Croatia. The lack of a significant difference in the perception of CRM as a technological solution and CRM's contribution to the business of telecommunications companies and the perception of business culture at different levels of management further indicates that employees at both levels of management perceive CRM as business culture. Employees at the operational level of management who represent points of contact with end customers and employees who work on background processes and make strategic decisions about the company's business perceive CRM in both observed dimensions without significant differences. This further means that in the telecommunications sector, a unique view of all employees is provided on the common goal of the company's business, which will be perceived by end customers through operational management and interaction with employees. Si, the goals that were set before research was conducted were achieved.

4.4 Research limitations and recommendations for future research

The limitations of the research should be considered when concluding. Since such research has never been conducted in that way before, and to conduct this research authors combined different approaches and methodologies from the available literature (Chang et al., 2010) (Becker et al., 2009) (Vadadi & Magham, 2014) (Ramya et al., 2013), which are tailored to the needs of this research the research limitation can be seen. Also, management levels are arbitrarily defined based on the available literature, and it should be noted that there may be some other division that could be used which might result in different research results. Also, combining tactical and strategic decision-making for research purposes can be considered a limitation. But, the distribution of respondents according to demographic characteristics may be the most important limitation of this research. Especially, the low turnout and the predominance of several companies, with the largest percentage of Iskon Internet employees, indicates that the sample of respondents does not represent the reference sample of the entire market, which calls into question the possibility of generalizing research results.

Recommendations for further research are to conduct research on a larger sample in order to get a broader picture and also to examine the business culture in the telecommunications industry in Croatia, in order to determine whether it is aimed at existing customers and to build long-term relationships with them. Also, it would be desirable to examine the reasons for the high percentage of outflows of customers, as well as the reasons for their dissatisfaction that leads them not to remain loyal to one telecommunications company for a long time period.

5 Conclusion

The services sector is extremely important in the economy of each country, and it is specific that due to

the intangibility of the service, employees have an extremely important role to play in companies' overall success. The employees are the ones who give the end customer perception of the quality of the delivered service. The CRM allows companies to focus on existing customers in order to treat them in a way that is satisfactory to them, which will ultimately lead to loyalty. And it is well known that loyal customer is an extremely valuable resource for the company, and that relationship with those customers generates valuable information for business decisions and competitive advantage. CRM in companies is the integration of employees, a unique company culture focused on the needs and wants of customers, business processes, and technology that automates these processes and uses them to facilitate the treatment and communication with customers as well as storage and handling of information about them. It is certainly not a one-dimensional structure and an isolated project, but a business philosophy permeated through all departments in the company. Therefore, it is important that everyone in the company shares the same values and goals and follows the same vision and mission and that company has to have in mind that CRM cannot be successfully implemented within the company without quality employees.

If companies in the telecommunications industry adopt a CRM and focus their business on the end customer by creating satisfaction and loyalty, the same perception of all employees can more easily achieve a consistent business culture, permeated through all departments within the company. In this way, they can achieve the benefits of implementing a CRM as a technological solution through a business culture aimed at end customers, which is the basis of the company's survival in the modern market. According to the findings from the available literature, retaining existing customers is certainly more cost-effective than acquiring new customers, and brings many other benefits for the company. CRM as a technology solution significantly facilitates the implementation of a business strategy aimed at existing customers if it is implemented in a cooperative environment where all employees are focused on the same goal. According to the results of the research conducted, for the purpose of writing this paper, there is no difference in the perception of CRM at different levels of management, so the telecommunications industry has a good basis for implementing a CRM strategy and creating long-term, mutually beneficial relationships with customers.

So, to conclude, management in telecommunications companies should continue to work on building and maintaining quality relationships and communication culture, both inside and outside the company. Because there is a consensus among employees that retaining existing customers is a company priority and that long-term customer relationships are a valuable asset, which is an extremely good basis for successful CRM implementation. Opportunities for improvement lie in the development of a CRM as a technology solution that will, in addition to valuable customer information, provide employees with information about the market and competition and based on that, the answers to frequently asked questions. This will make the technology solution even easier for employees to communicate with customers or make business decisions. In addition, management should work on evaluating and rewarding employees based on building and improving relationships with existing customers as this will further foster a customer-oriented business culture in the company.

References

Becker, J. U., Greve, G., & Albers, S. (2009). The impact of technological and organizational implementation of CRM on customer acquisition, maintenance, and retention. *International Journal of Research in Marketing*, *26*(3), 207–215. https://doi.org/10.1016/j.ijresmar.2009.03.006

Brodarić, A. (2010). Upravljanje odnosima s korisnicima kroz procese CRM-a. *18. Telekomunikacioni Forum TELFOR 2010*, 98–101.

Buttle, F. (2001). Customer retention: A potentially potent marketing management strategy. *Journal of Strategic Marketing*. https://www.academia.edu/10719492/Customer_retention_a_potentially_potent_marketing_management_strategy

Chang, W., Park, J. E., & Chaiy, S. (2010). How does CRM technology transform into organizational performance? A mediating role of marketing capability. *Journal of Business Research*, *63*(8), 849–855. https://doi.org/10.1016/j.jbusres.2009.07.003

Chen, I. J., & Popovich, K. (2003). Understanding customer relationship management (CRM): People, process and technology. *Business Process Management Journal*, *9*(5), 672–688. https://doi.org/10.1108/14637150310496758

Coltman, T. (2007). Why build a customer relationship management capability? *The Journal of Strategic Information Systems*, *16*(3), 301–320. https://doi.org/10.1016/j.jsis.2007.05.001

Dirsehan, T. (2017). The Evolution of CRM from the Perspective of an Emerging Country: Past, Present and Future. *UBT International Conference*, 49–52.

Dukić, B., & Gale, V. (2015). UPRAVLJANJE ODNOSIMA S POTROŠAČIMA U FUNKCIJI ZADRŽAVANJA POTROŠAČA. *Ekonomski vjesnik/Econviews - Review of Contemporary Business, Entrepreneurship and Economic Issues*, *28*(2), 583–598.

Eichorn, F. L. (2004). Internal Customer Relationship Management (IntCRM): A Framework for Achieving Customer Relationship Management from the Inside Out. *Problems and Perspectives in Management*, *2*(1). https://www.businessperspectives.org/index.php/journals/problems-and-perspectives-in-management/issue-5/internal-customer-relationship-management-intcrm-a-framework-for-achieving-customer-relationship-management-from-the-inside-out

Eid, R. (2007). Towards a Successful CRM Implementation in Banks: An Integrated Model. *The Service Industries Journal*, *27*(8), 1021–1039. https://doi.org/10.1080/02642060701673703

Fleming, J. M., Coffman, C., & Hartner, J. K. (2005). Manage your human sigma. *Harvard Business Review*, *34*(1), 1–10.

Foss, B., Stone, M., & Ekinci, Y. (2008). What makes for CRM system success—Or failure? *Journal of Database Marketing & Customer Strategy Management*, *15*(2), 68–78. https://doi.org/10.1057/dbm.2008.5

Grove, S. J., & Fisk, R. (2004). *Service theater: An analytical framework for services marketing* (pp. 78–87).

Hennig-Thurau, T. (2004). Customer orientation of service employees: Its impact on customer satisfaction, commitment, and retention. *International Journal of Service Industry Management*, *15*(5), 460–478. https://doi.org/10.1108/09564230410564939

Jayachandran, S., Sharma, S., Kaufman, P., & Raman, P. (2005). The Role of Relational Information Processes and Technology Use in Customer Relationship Management. *Journal of Marketing*, *69*(4), 177–192. https://doi.org/10.1509/jmkg.2005.69.4.177

Kamrul Islam Shaon, S. M., & Rahman, H. (2015). A Theoretical Review of CRM Effects on Customer Satisfaction and Loyalty. *Central European Business Review*, *4*(1), 23–36.

Malthouse, E. C., Haenlein, M., Skiera, B., Wege, E., & Zhang, M. (2013). Managing Customer Relationships in the Social Media Era: Introducing the Social CRM House. *Journal of Interactive Marketing*, *27*(4), 270–280. https://doi.org/10.1016/j.intmar.2013.09.008

Mohammed, M. F. (2013). Customer Relationship Management (Telecommunication Industry) Comparison Between (Airtel) And (Zain). *International Journal of Business and Management Invention*, *2*(11), 52–58.

Nahar, P., & Dhaka, V. P. (2014). A Review-: Benefits and Critical Factors of Customer Relationship Management. *Global Journal of Computer Science and Technology*. https://computerresearch.org/index.php/computer/article/view/142

Nemeček, J. (2013). Customer relationship management influence on sales of selected companies. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, *60*(2), 231–236. https://doi.org/10.11118/actaun201260020231

Nikolić, D. M., Stanković, M. A., & Simić, N. D. (2014). CRM as a management tool for improving services. *Journal of Process Management and New Technologies*, *2*(1), 81–90.

Payne, A. (2005). *Handbook of CRM: Achieving Excellence in Customer Management* (1st ed.). Elsevier; Butterworth-Heinemann.

Rahaman, M. M., Ferdous, M. I., & Rahman, M. Z. (2011). Role of CRM in Profitability of Service Organizations: A Case of a Leading Telecommunication Company in Bangladesh. *Journal of Economics and Sustainable Development*, *2*(4), 91–102.

Ramya, S., Rajam, K., & Sivasubramanian, D. (2013). A study on employees' perception of customer relationship management in e-banking. *Asia Pacific Journal of Research*, *1*(9), 9–17.

Rootman, C., Tait, M., & Bosch, J. (2007). The influence of bank employees on bank customer relationship management. *Acta Commercii*, *7*(1), 181–192. https://doi.org/10.10520/EJC17001

Ryals, L., & Knox, S. (2001). Cross-functional issues in the implementation of relationship marketing through customer relationship management. *European Management Journal*, *19*(5), 534–542. https://doi.org/10.1016/S0263-2373(01)00067-6

Severović, K. (2013). *Upravljanje odnosima s klijentima kao izvor informacija za oblikovanje i poboljšanje usluga* [Sveučilište u Zagrebu, Fakultet organizacije i informatike Varaždin]. https://urn.nsk.hr/urn:nbn:hr:211:214090

Triznova, M., Maťova, H., Dvoracek, J., & Sadek, S. (2015). Customer Relationship Management Based on Employees and Corporate Culture. *Procedia Economics and Finance*, *26*, 953–959. https://doi.org/10.1016/S2212-5671(15)00914-4

Vadadi, A., & Magham, S. (2014). Investigating the effect of organizational capacity on CRM performance. *Management Science Letters*, *4*(8), 1625–1630.

Verhoef, P. C. (2003). Understanding the Effect of Customer Relationship Management Efforts on Customer Retention and Customer Share Development. *Journal of Marketing*, *67*(4), 30–45. https://doi.org/10.1509/jmkg.67.4.30.18685

Xu, Y., & Geodegebuure, R. (2005). *Employee satisfaction and customer satisfaction: Testing the service-profit chain in a Chinese securities firm*. https://openrepository.aut.ac.nz/handle/10292/2082

Zablah, A. R., Bellenger, D. N., & Johnston, W. J. (2004). An evaluation of divergent perspectives on customer relationship management: Towards a common understanding of an emerging phenomenon. *Industrial Marketing Management*, *33*(6), 475–489. https://doi.org/10.1016/j.indmarman.2004.01.006

Blended Mobility on Entrepreneurship – Intensive Programme for ICT Students in the HEI Curriculum

Eva Malichová, Peter Márton, Emese Tokarčíková
Faculty of Management Science and Informatics
University of Žilina
Univerzitná 8215/1, 01026 Žilina, Slovakia
{eva.malichova, peter.marton, emese.tokarcikova}@uniza.sk

Abstract. *The main objective of the SmartSoc project was to create and implement an intensive interdisciplinary course in the field of ICT-based entrepreneurship considering the transnational and intercultural context. This project enforced ICT student teams to deal with economic, societal and environmental challenges defined by strategic documents 'Transforming our world: the 2030 Agenda for Sustainable Development' and 'Horizon Europe' programmes. It included an important message about smart specialisation strategies and their importance for regional development in the EU. The project results confirm the increasing economic and societal knowledge of the participants. They obtained them throughout out the development of start-up business plans based on selected entrepreneurial cases in line with essential sustainability aspects.*

Keywords. entrepreneurship, ICT, HEI curriculum, interdisciplinary

1 Introduction

The 'Education of Future ICT Experts Based on Smart Society Needs (SmartSoc)' project was developed by the consortium of twelve universities and one registered voluntary organisation from nine EU countries, coordinated by the University of Žilina, Slovakia. The main objective of the SmartSoc project was to enforce ICT student teams to deal with economic, societal and environmental challenges of current strategic documents of the UN and the EU when creating own business plans for virtual start-ups.

The advancement of HEI education in ICT encompasses different aspects such as learning methods, tools, business and social skills and work in the environment as closely as possible to a future workplace in high technology, i.e., in interdisciplinary international team. Therefore, the key issue of the project was the collaboration of students from different European universities on a common task (an entrepreneurial case) related to recognising needs of society and providing a smart solution by using ICT. Specific entrepreneurial cases from Smart Society domain were elaborated following principles of project-based learning, student teamwork, blended mobility and intensive courses.

The SmartSoc project brought novelties regarding topics under consideration, i.e. ICT expertise related to smart society needs, with a special emphasis on entrepreneurship and skills related to evaluation of business perspectives and societal implications of the proposed solution, respecting relevant smart specialisation strategies defined in the EU at different geographical levels.

The project duration (three years) enabled the project consortium to organise three blended mobilities for about 100 students. The blended mobilities structure and content were adapted based on participants' surveys and unexpected situations caused by covid-19 and the war in Ukraine.

2 The needs analysis – preparation phase

When preparing the SmartSoc project, the project consortium relied on the results of the needs analysis. An online survey was conducted in February and March 2019. There were two questionnaires prepared for two different target groups - ICT companies and students of ICT oriented study programmes.

The first survey involved 28 IT companies focusing on software development. 70% were companies operating in Slovakia, further companies in the Czech Republic, Germany and Hungary. More than 53% of companies being stated had more than 250 employees. In the 25% of the companies involved in the survey, 50 to 249 employees were declared. 93% of companies answered they are interested in new education programmes aimed not only at improving ICT skills but also at basic business and personal skills. Based on the responses, it can be noted that when they are recruiting new employees, they attach great importance, in addition to IT skills, to the transfer of

skills and business skills. Other factors included digital skills, networking and language skills.

The second survey involved 329 students mainly from Slovakia (59%), Hungary (25%), following by students from France, Croatia, Ireland and Spain. All the respondents studied ICT focused study programme. Students were asked if they had any possibility to attend lessons, courses, or seminars where they could gain knowledge related to business, management or communication (Business and Social Skills). The most frequent answers were Economics courses (69%), as well as they had an opportunity to study courses such as Sociology (25%), Psychology and Rhetoric and E-learning possibilities, e.g. MIT, Coursera (20%). Most of the students stated their position in programming. Among the most important areas of the ICT skills, students mainly presented Topical knowledge, Analytical skills, and Awareness of new IT trends. Among the most important areas within Business Skills, they selected Project Management and Planning, Delegation and Time Management, Team Management, Communication and Negotiation skills. The differences within the priorities in the extent of bidding options were minimal, but networking was the least important according to students. The answer to the question 'What are the most important areas within the Personal Skills?' was mainly the problem solving, flexibility, responsibility, and motivation. Interpersonal skills were considered to be of the least importance. From all respondents more than half said they were interested in an educational exchange abroad, especially in the frame of the Erasmus+ programme.

3 SmartSoc Blended Mobilities

3.1 Basic concept

Blended learning (also known as hybrid learning) could be described as a study type that involves learning in a combination of modes, often used more specifically to refer to courses which use a combination of traditional face-to-face teaching workshops or seminars and distance learning techniques online (such as the internet, television, conference calls) (Erasmus+ Programme Guide, Version 1 2018, 25/10/2017). This is a definition published in the previous Erasmus+ programme (2014-2021). Blended learning was supported by the Key Action 2 activities in the frame of Strategic partnerships of HEI institutions.

The SmartSoc project goal was to develop basic (i.e. Maths, Science and Technology) knowledge and transversal (i.e. critical thinking, problem-solving, collaborative working in intercultural teams, management) skills of HEI students using innovative methods (i.e. teaching and learning based on case studies and intercultural collaboration) implemented through blended mobility and intensive course. A special focus in the project curricula was put on technical, economic and legal aspects of entrepreneurial education.

The basic blended mobility idea of the SmartSoc project was to motivate the participants (students and teachers) to identify new ICT trends which emerged during the last years. Student teams supported by their supervisors (teachers) worked on the entrepreneurial cases together. These case studies took a twofold perspective on ICT and analysed not only innovations that happened within ICT, but as well as how ICT fostered innovations in other fields (e.g. 'green energy', 'smart transport', 'inclusive society', 'smart specialisation strategy of the specific region'). The original plan was to develop 8 ICT-based entrepreneurial cases every year (2020, 2021, 2022), which will be solved by student teams during the blended mobility. Description and results of those entrepreneurial cases were used to create the first project's intellectual output. This intellectual output was presented to a wider audience at the multiplier events in the last project months. The general learning outcome of blended mobilities was defined by the following words – 'identify, understand and apply current ICT trends to become an entrepreneur'.

The preparation phase of each blended mobility started 6 months before its real start. Teachers from all participating organisations were invited to prepare and apply their entrepreneurial case proposals. It was realised approximately in the middle of October. The final list of the entrepreneurial cases used for the following blended mobility was announced in November. The call for students was opened approximately at the same time to receive enough applications before the winter holidays. The final list of student participants was known in January. It was used to create student teams and to announce their assignment to the entrepreneurial case in a few months.

The work in the student teams always started in the middle of February. Unfortunately, it was not possible to get the beginning of cooperation between students and teachers to start at the beginning of the second semester of the current academic year. However, this was possible with most partners. The blended mobility itself was adapted to the duration of one semester at the project coordinator, the University of Žilina. Two phases of blended mobility were planned for the 13-week semester - virtual and physical. The virtual part lasted 10 weeks and the physical part 2 weeks. There was one spare week to complete the tasks from the virtual part, if a student team needed it, between them. The blended mobility completion was planned for late May.

The virtual mobility involved working on a joint task - an entrepreneurial case. The student team was created so that it cannot include students from the same institution and country. Normally, group of 3 to 5 students worked in each student team. Each student team was provided with guidance by teachers (supervisors) in three areas - technical, business and societal.

3.2 Participants

The project consortium was created by 12 institutions as mentioned in the introduction (Figure 1 and Table 1).



Figure 1. SmartSoc project – participating countries

Table 1. SmartSoc partner institutions

Country	Institution
Slovakia	University of Žilina
Slovakia	Technical University of Košice
Bulgaria	Technical University of Sofia
Romania	University of Oradea
Hungary	University of Debrecen
Hungary	University of Győr
Croatia	University of Zagreb
Croatia	University of Osijek
Germany	EIAB institute
Finland	University of Jyvaskyla
Finland	Seinajoki University of Applied Sciences
France	IMT Atlantique
Spain	Technical University of Valencia

The participation of three students and two teachers from each partner institution was planned for each blended mobility. Students and teachers were nominated through selected contact person. The numbers of participants differed in each year of the project. Thanks to the EIAB partner from Germany, it was possible to involve students from institutions that were not directly involved in the project consortium. In the last year of the project, students from Cameroon and Albania were also invited to blended mobility, as well as students admitted to study mobility at the University of Žilina (including students from Taiwan and Kazakhstan).

There is an overview of number of students nominated for blended mobilities in table 2. Some students cancelled their participation during the virtual mobility. It shows how covid-19 and massive implementation of online teaching influenced motivation and workload of students at all participation institutions. There is an apparent influence of the workshop venue in all project years as well – Žilina (2020), Valencia (2021), Žilina (2022). The 2022 SmartSoc Workshop finally took place in Valencia because of war in Ukraine.

Table 2. Statistics of participants

	2020	2021	2022
Virtual mobility	35	51	36
Workshop	17	32	22

3.3 Entrepreneurial cases

An entrepreneurial case was the most important task formulation for the student teams. All assignments were prepared by teachers from the participating institutions. Some of them in cooperation with the industry partners. These are the examples of the solved SmartSoc entrepreneurial cases prepared for 2021:

- Smart Museums
- A Self-Driving Tractor Is Changing the Future of Food Business
- Safe, Effective, and Sustainable covid-19 Protection Operations in Congested Areas
- Smart Recycling in Latest Technologies
- Cyber Security
- Creation of Identification Methods and Tools to Distinguish between Original or Counterfeit Products
- Health on the Move
- Smart Package Solutions
- 5G Open Radio Access Networks in Event Management

There were nearly 30 entrepreneurial cases prepared and 23 out of them were used within the SmartSoc project.

3.4 Cooperation tools being used within blended mobility

The cooperation tools and platforms enabled effective cooperation of the student teams during virtual mobility. It was not expected to use any cooperation and communication platform during physical mobility (two weeks workshop).

For the first project year, before covid-19, it was decided to use Moodle LMS as the essential cooperation platform. All study materials were shared there with students and teachers. All teams were invited to decide about their own additional communication tool - a FB group, Messenger group, Skype, etc. This combination of cooperation tools and platforms was used only during the first project year (2020).

The blended mobilities were influenced by covid-19 during all three project years. Unfortunately, the first workshop, the physical part of the first blended mobility, could not be prepared according to the original plan. May 2020 was completely unsuitable for physical mobility, and the consortium then decided to move the workshop to September 2020, hoping that covid-19 would be defeated by that time. Finally, the workshop was held in an abbreviated version, for six days and only online. Since this event, Microsoft Teams has become the main communication platform of the project. Moodle LMS has begun to lose its importance, especially as a place to share teaching materials.

In 2021 and 2022, all student team collaboration moved to the Microsoft Teams application environment. In this environment, each team had their private channel with their own content and place to communicate with each other. The LMS Moodle environment was created in 2021, but it was no longer used by students at all. In 2022, all student teams used only the Microsoft Teams application. It was used during the second online workshop in May 2021 as well.

3.5 Blended mobility structure – learning and teaching materials

The content of the virtual part of the blended mobility was carefully prepared by the staff of the Department of Macro and Microeconomics of the project coordinator. A few tasks were prepared for the students. For the processing of each task, documents were prepared in the form of documents or links to videos publicly available on the Internet. Teachers assigned to lead student groups regularly consulted on all assignments so that students could understand and process them. This is what the content of virtual mobility in the first year looked like (Figure 2):

1. Kick-off meeting – Teamwork – How to create a successful team
2. The first version of a Business Model Canvas (BMC)
3. The first version of a Business Plan for a start-up
4. The second version of the BMC and the Business Plan
5. Pecha Kucha presentation
6. The final version of the BMC and the Business Plan

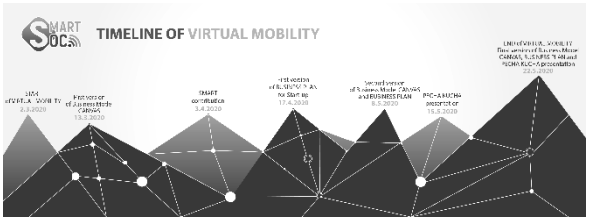


Figure 2. Virtual mobility timeline (2020)

The students always received feedback on each version of their BMC and business plan. The very last changes in the business plan were made by the student teams subsequently during the workshop. The workshop, originally planned as a physical meeting of students and teachers, contained mainly lectures focused on current topics of ICT and business in ICT. The lectures of the teachers of the partner institutions were supplemented by lectures by industry experts. A few excursions, even virtual ones, were organised.

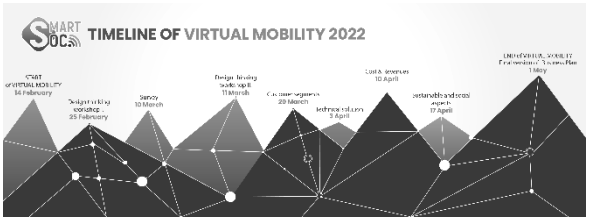


Figure 3: Virtual mobility timeline (2022)

The organisation and structure of virtual mobility were modified and improved based on surveys. Students and teachers used this possibility to give feedback to the project organisational team. The third version of virtual mobility content was the following (example of five weeks – for a timeline see Figure 3):

- Week 01 – Task 01: Agreement on weekly meetings / Task 02: What do you like? / Task 03: Find information about your entrepreneurial case, current solutions, trends
- Week 02 – Task 01: Weekly discussion / Task 02: Design Thinking workshop – Part I.
- Week 03 – Task 01: Weekly discussion / Task 02: Competition analysis / Task 03: Survey
- Week 04 – Task 01: Weekly discussion / Task 02: Team roles / Task 03: Design Thinking workshop – Part II.
- Week 05 – Task 01: Weekly discussion / Task 02: Definition of your product/service / Task 03: Customer segments, Persona and Value proposition

One of several activities was to propose and create a start-up name, acronym and logo. There are examples of student teams' logos in the Figure 4.



Figure 4: Start-up logos (2021)

4 Public access to the SmartSoc project results

The SmartSoc project was supported by the European Union in the frame of Erasmus+ KA2 programme. All learning and teaching materials and project intellectual outputs are accessible and visible for public use at different places. The project consortium uses several ways to multiply the project impact and to share the project life cycle, e.g.:

- Project web – https://www.smartsoc.uniza.sk/
- FB account – https://www.facebook.com/smartsocproject
- Moodle LMS – 2020 – https://moodle.uniza.sk/public/course/index.php?categoryid=11
- Moodle LMS – 2021 – https://moodle.uniza.sk/public/course/index.php?categoryid=12
- Moodle LMS – 2022 – https://moodle.uniza.sk/public/course/index.php?categoryid=13
- SmartSoc YouTube channel – https://www.youtube.com/channel/UC0SsORwOMyxxjHgVMGmgG3A

Above mentioned links allow to access the content and structure of all blended mobilities, including videos of lectures prepared for the project workshops and final presentations of student teams.

8 Future work

Blended intensive programmes (BIP) were introduced in the frame of new Erasmus+ programme period (2022-2029). Blended mobilities support was moved from the KA2 action directly to the KA1 actions to be part of different options for individual mobilities of students and teachers.

The SmartSoc project consortium prepared a well-developed basis for the future cooperation in or out of the current consortium. The project results could be implemented in the new BIPs. Basic conditions to prepare and organise a BIP are quite similar but there are some differences. For example, it is not possible to cover the participation costs of students and teachers in the same event any longer. Only 'learners' costs are eligible. But there is an option of the individual KA1 teaching mobility combination.

Other challenges should be faced and solved to be able to 'switch' to the new way of organising blended mobilities in the frame of the Erasmus+ programme. They are related to the effective organisation of the virtual part of the BIP. Virtual mobility was very challenging when organising online team meetings for students and teachers from different institutions. Semester duration, own courses schedule, free time activities – all of these should be considered. There is possibly one solution that may be understood like very unrealistic – the creation of a joint course for all participating institutions with predefined time slots in the course schedules. This may avoid other time planning discords for students and teachers. There are some initiatives to enable students to participate in a mobility with duration of a semester – 'mobility window'. The support of increasing internationalisation of higher education environment stated in the latest Higher Education Act of 2022 in Slovakia (bulletin SAIA, April 2022). It is time to think about a similar concept to enable the virtual part of BIPs.

Acknowledgements

This paper was supported by the European Commission in the framework of 'Education of Future ICT Experts Based on Smart Society Needs (2019-1-SK01-KA203-060789)' project.
The European Commission accepts no responsibility for the information contained in this paper.

References

Čo prináša novela zákona o vysokých školách z pohľadu internacionalizácie vysokoškolského vzdelávania ? (What does the amendment to the Higher Education Act bring from the point of view of the internationalization of higher education?) Retrieved from https://www.saia.sk/_user/documents/bulletin/bulletin2022/BS_04_2022.pdf

Erasmus+ Programme Guide 2022. Retrieved from https://erasmus-plus.ec.europa.eu/programme-guide/erasmusplus-programme-guide

Intelligent Energy

Zvonimir Šimić, Danijel Topić, Anto Rogalo
Modeling and Simulation of AC, DC, and Hybrid AC-DC Microgrid with Battery Energy Storage

Modeling and Simulation of AC, DC, and Hybrid AC-DC Microgrid with Battery Energy Storage

Zvonimir Šimić, Danijel Topić
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Josip Juraj Strossmayer University of Osijek
Kneza Trpimira 2b, 31000 Osijek, Croatia
{zvonimir.simic, danijel.topic}@ferit.hr

Anto Rogalo
Primary School Ivanja Reka
Ivanjorečka cesta 1b, 10373 Ivanja Reka, Zagreb, Croatia
anto.rogalo@skole.hr

Abstract. *A microgrid is a part of a distribution grid that can operate connected to the main grid and independent from the main grid. There are AC microgrids, DC microgrids, and hybrid AC-DC microgrids. The difference between these three topologies is the number of AC-DC converters. Modeling and simulation of these three main microgrid topologies and a comparison of simulation results are presented in this paper. The microgrid model consists of the photovoltaic power plant, wind turbine, battery energy storage, and two loads. Simulation results are presented so microgrid topologies can be compared and the most suitable topology can be selected.*

Keywords. Microgrids, AC microgrid, DC microgrid, Hybrid AC-DC microgrid, Battery energy storage

1 Introduction

Trends in power systems are transformations from large-scale power grids where electrical energy is produced in large power plants and then transported over the transmission and distribution lines to customers to small-scale power grids where electrical energy is produced in distributed generation and used locally. Distributed generation (DG) is renewable energy sources (RES) based, which leads to reducing global warming and environmental pollution impact. In most cases, photovoltaic (PV) power plants which are suitable for mounting on the roofs of houses and buildings, are used. Production and use of electrical energy locally lead to reducing power losses due to energy transportation on long distances. Generation and use of electrical energy at the same location make a microgrid. The microgrid consists of a generation, consumption, and energy storage system, usually battery energy storage (BES).

There are different types of microgrids. Based on the number of energy sources, microgrid configurations can be classified as microgrid configurations based on two, three, or four energy sources (Rezkallah et al., 2019). There are centralized microgrids where all consumers are connected to one microgrid with a central PV power plant and energy storage and standalone nano grids where every consumer is equipped with their PV power plant and energy storage (Werth et al., 2015):

Microgrids can be classified considering six criteria (Sahoo et al., 2021):

- Power form: AC and DC microgrids.
- Microgrid application: utility, industrial and remote microgrid.
- Microgrid structure: single-stage and two-stage.
- Supervise control: unified control and decentralized control.
- RES integration: converter and conventional DG-based microgrids.
- Operation mode: grid forming and grid following.

Microgrids can be classified according to three main criteria (Cabana-Jiménez et al., 2022).

- Demand criteria: simple, multi DG, and utility.
- Capacity criteria: simple, corporate, feeder area, substation area, and independent microgrids.
- Power criteria: DC, AC, and hybrid DC-AC.

This paper deals with the comparison of microgrids according to the power criteria and there are AC, DC, and hybrid AC-DC microgrids. Current distribution networks are AC so switching to an AC microgrid is very simple as it is with the development of microgrids in a rural area (Zhu et al., 2019). The AC microgrid is a suitable solution for grid operation mode and when the backup generator is used (Minemura et al., 2020). PV and batteries are DC while the utility grid and consumers are AC so there is a hybrid AC-DC microgrid that connects DC and AC sides with a bidirectional AC-DC converter (Gao et al., 2020). Fuel cells that have the potential to be a great power source in the future are DC, like batteries, and PV so hybrid AC-DC microgrids will be a suitable solution in the future (Liang et al., 2018). Many devices today are charged with USB which is DC, LED lighting is DC,

and data centers require DC power distribution so, with a high amount of DC load, production from PV power plants and BES has the sense to make a DC microgrid (Ali et al., 2021). DC microgrids are suitable for commercial buildings with computers, data centers, LED lighting, EV charging stations, batteries, and PV power plants (Zhang et al., 2015).

The contribution of this paper is modeling and simulation of AC, DC, and hybrid AC-DC microgrid topologies for the same microgrid components sizes and comparison of results to find the most suitable topology for selected component sizes.

This paper consists of five chapters. The first chapter gives an introduction. Microgrids and microgrid topologies are described in chapter 2. Modeling of microgrid components is presented in chapter 3. Chapter 4 presents simulation results. The last chapter gives the conclusion of this paper.

2 Microgrids and microgrid topologies

A microgrid is a part of the distribution grid that can operate connected to the main grid and independent from the main grid. A microgrid is usually on a low voltage distribution grid but sometimes can be on a medium voltage distribution grid. It can be only one building with a PV power plant on the roof or more buildings located closely. A microgrid can be for residential, commercial, or industrial consumers. Microgrids are suitable for rural areas or remote areas which are not connected to the A microgrid connects DGs and grid users at the same location in a small power system.

Microgrids have a positive impact on the environment and the end customers. There are several advantages of microgrids as follows:

- Reducing greenhouse gas emissions by producing electricity from RES.
- Reducing power losses due to the short distance the between production of electricity and consumption. There are no long power lines.
- Resistance to network disturbances.
- Ability to power consumption when the main grid is not available if there is enough energy storage capacity or backup generator.

Microgrid operation independent from the main grid presents a great challenge to ensure enough energy to cover whole consumption. There are several disadvantages of microgrids as follows:

- Requirements for battery energy storage systems, which are still very expensive.
- Requirements for the backup generator which is very expensive to operate due to high fuel prices.
- Variability and intermittency of solar and wind energy.
- Load reducing if there is not enough battery capacity or backup generator when the main grid is not available.

According to the power criteria, there are AC, DC, and hybrid AC-DC microgrids. Differences between these three topologies are the type of power between components, the number of converters, the rated power of converters, and the efficiency of converters.

2.1 AC microgrid

The distribution grid and most of the buildings are AC so it is a very easy switch to an AC microgrid. In an AC microgrid, components are connected to a central AC bus. AC components like a wind turbine, AC load, and grid are connected to the AC bus directly while DC components like PV power plants, batteries, and DC loads are connected to the AC bus over DC-AC converters. The battery is connected over the bidirectional DC-AC converters. The topology of the AC microgrid is presented in Fig. 1.

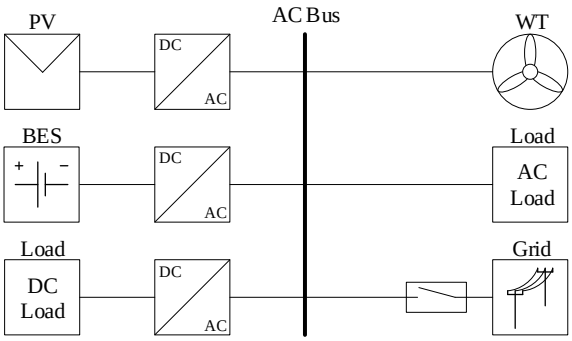


Figure 1. The topology of the AC microgrid

2.2 DC microgrid

DC microgrids are suitable for commercial buildings like data centers which require DC power distribution. In the DC microgrid, components are connected to the central DC bus. DC components like PV power plants, batteries, and DC loads are connected directly to the DC bus while AC components like wind turbines and grids are connected to the DC bus over AC-DC converters, and the AC load is connected over the DC-AC converter. The topology of the DC microgrid is presented in Fig. 2.

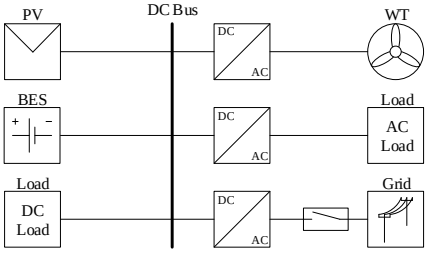


Figure 2. The topology of the DC microgrid

2.3 Hybrid AC-DC microgrid

In building with AC and DC loads and generation, hybrid AC-DC microgrids are a suitable solution. There are two buses, the AC bus, and the DC bus. AC components like a wind turbine, AC load, and grid are connected to the AC bus directly. DC components like PV power plants, batteries, and DC loads are connected directly to the DC bus. AC bus is connected to the DC bus over the bidirectional converter. The topology of the hybrid AC-DC microgrid is presented in Fig. 3.

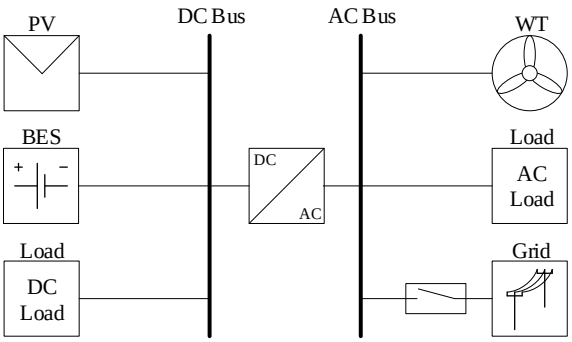


Figure 3. The topology of the hybrid AC-DC microgrid

3 Modeling of microgrid

In this chapter models of microgrid and microgrid components are described. The microgrid model in this paper consists of the PV power plant, wind turbine (WT), battery energy storage, AC load, DC load, and the utility grid. These components are connected to the AC, DC, and hybrid AC-DC microgrids.

3.1 PV power plant

Generation from PV power plant is modeled according to measured solar irradiation data for the sunny day presented in Fig. 4. Generation from PV power plant starts at 4:00 and peak is at 12:00. Generation from PV power plant ends at 21:00.

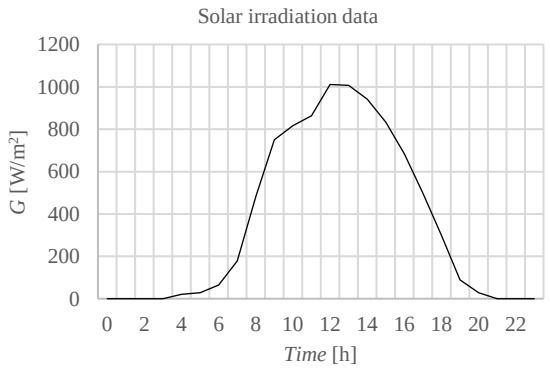


Figure 4. Solar irradiation data

One PV power plant consists of 40 PV modules with a maximum power of 250 W so the total installed power of one PV power plant is 10 kW. Module efficiency is 15.3% and dimensions are 1649x991 mm. When solar irradiation data is prepared, PV generation can be estimated according to eq. 1.

$$P_{\mathrm{pv,i}} = n \cdot G_{\mathrm{i}} \cdot A \cdot \eta \tag{1}$$

Where:

- n – number of PV modules;
- $P_{\mathrm{pv,i}}$ [W] – PV generation in hour i;
- G_{i} [W/m^2] – solar irradiation in hour i;
- A [m^2] – area of one PV module;
- η – efficiency of a PV module.

3.2 Wind turbine

Generation from a wind turbine is modeled according to measured wind speed data for one day presented in Fig. 5. The rated power of the wind turbine used in this model is 10 kW. Wind turbine height is 27 m, and wind speed is measured at 10 m so wind speed data needs to be calculated at a height of 27 m according to eq. 2. Output power of the wind turbine can be estimated according to the wind turbine power curve presented in Fig. 6. The cut-in speed is 2 m/s, rated power is achieved at 9 m/s and the cut-off speed is 30 m/s.

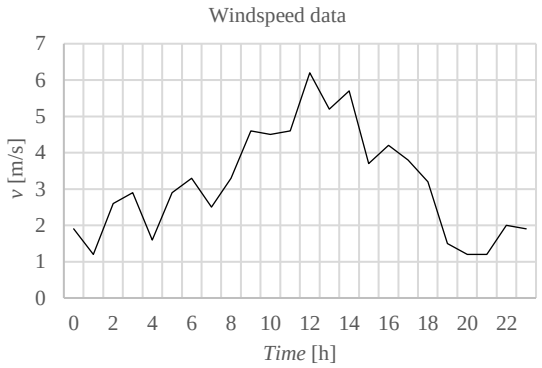


Figure 5. Wind speed data

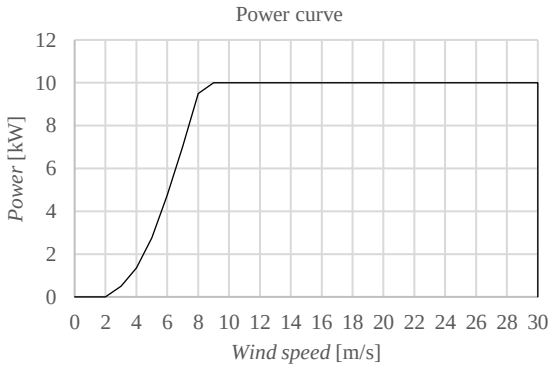


Figure 6. Wind turbine power curve

$$v_i = v_{0,i} \cdot (h / h_0)^{\alpha} \quad (2)$$

Where:

- v_i [m/s] – wind speed at height h in the hour i;
- $v_{0,i}$ [m/s] – wind speed at height h_0 in the hour i;
- h [m] – the height of the wind turbine;
- h_0 [m] – the height of wind speed measurement;
- α – friction coefficient.

3.3 Load

It is assumed that the microgrid model is for a residential area. There are two types of loads in the model, households and charging stations for electric vehicles (EV). AC load is modeled with measured load data which is scaled to microgrid size. The daily load diagram of AC load is presented in Fig. 7.

Minimal load demand is at 9:00 and peak load demand is at 20:00. The total daily AC load demand is 600 kWh. DC load is an EV charging station with a rated power of 50 kW. The daily load diagram of the EV charging station is presented in Fig. 8. Electric cars are connected to the EV charging station from 6:00 to 10:00, then from 16:00 to 18:00, and from 22:00 to 0:00. Electric cars are connected to the charging station 8 hours during the day so the daily load demand for the EV charging station is 400 kWh. The total daily load demand is 1000 kWh.

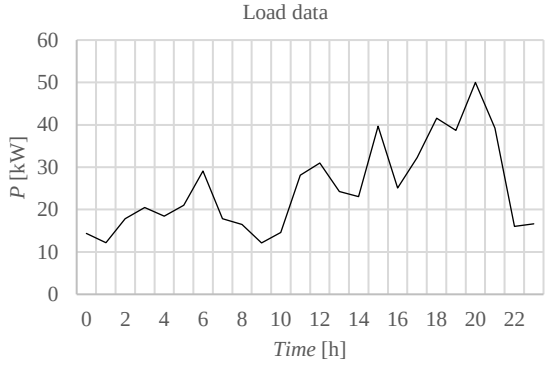


Figure 7. Load data

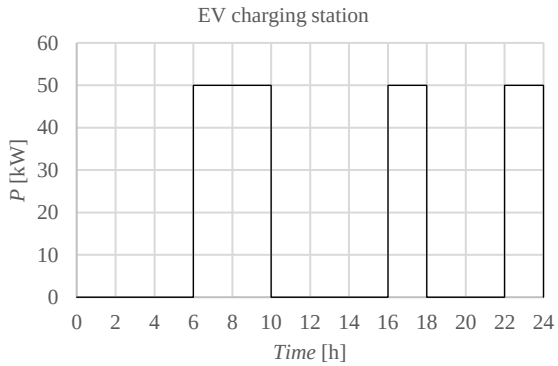


Figure 8. EV charging station

3.4 Battery energy storage

Battery storage is very important in microgrids to ensure enough energy to cover load demand during periods without solar and wind energy and if the main grid is not available. Battery energy storage used in this model is battery storage with a rated power of 50 kW and an installed capacity of 192 kWh.

Battery storage can be charged when there is an excess of energy generated from PV power plants or wind turbines and discharged when generation from PV power plants or wind turbines is lower than load demand. Battery storage also can be charged with electricity from the main grid during a period of low electricity prices.

4 Simulation results

The results of the simulations will be presented in this chapter. Simulations are done for AC, DC, and hybrid AC-DC microgrids during a working day. Generation from the PV power plant is presented in Fig. 9, electricity generation from a wind turbine is presented in Fig. 10, and load profiles for AC and DC load are the same for all three microgrid topologies. Battery charging and discharging power, battery state of charge (SOC), and power received from the main grid are different for AC, DC, and hybrid AC-DC microgrids.

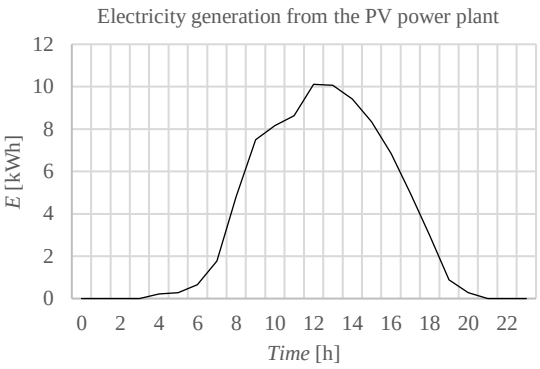


Figure 9. Electricity generation from PV power plant

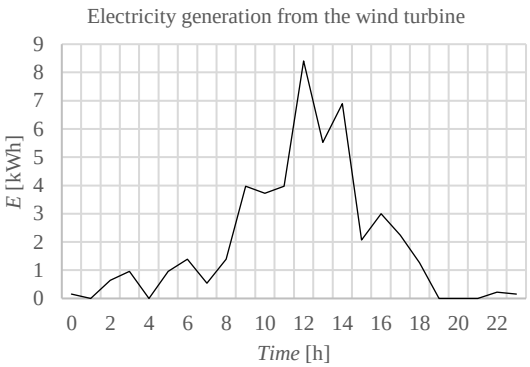


Figure 10. Electricity generation from wind turbine

Installed power of PV power plant and wind turbine, the number of PV power plants and wind turbines, daily generation, and total daily generation are presented in Table 1.

Table 1. The electricity generation in the microgrid

	PV power plant	Wind turbine
Power [kW]	10	10
Number of units	4	4
Daily generation of one unit [kWh]	85.945	47.472
Total daily generation [kWh]	343.78	189.89

The total daily generation from one PV power plant is 85.945 kWh. There are four PV power plants in the observed microgrid model so the total daily electricity generation from the PV power plants is 343.78 kWh. The total daily generation from one wind turbine is 47.472 kWh. There are four wind turbines in the observed microgrid model so the total daily electricity generation from the wind is 189.89 kWh. Two tariff electricity pricing model is considered in this microgrid simulation model. There is a higher day tariff with a price of 0.16 EUR/kWh and a lower night tariff with a price of 0.08 EUR/kWh. Day tariff starts at 8:00 and ends at 22:00 while night tariff starts at 22:00 and ends at 8:00.

4.1 Simulation of AC microgrid

In the AC microgrid there is a DC-AC converter that connects the DC PV power plant to the AC microgrid, then a bidirectional DC-AC converter that connects battery storage to the microgrid and enables charging and discharging of battery storage, and an AC-DC converter that connects EV charging station to the AC microgrid. The efficiencies of the DC-AC converter and bidirectional DC-AC converter are set to 0.92 and the efficiency of the AC-DC converter is set to 0.87 (Atia et al., 2016).

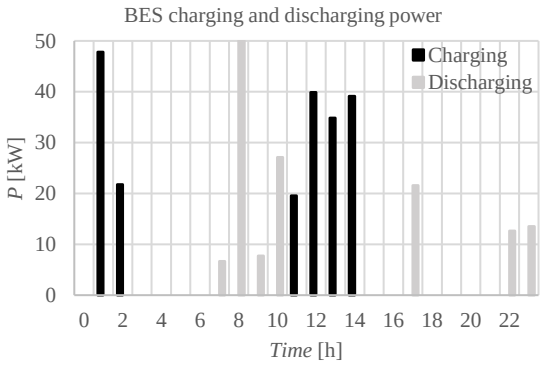


Figure 11. Charging and discharging power of BES in AC microgrid

The charging and discharging power of BES in the AC microgrid is presented in Fig. 11. It can be seen that the battery is charging during the night when the electricity price is low and during the day when there is a peak of generation from PV power plants and wind turbines. The battery is discharging in the morning when electric cars are connected to the EV charging station and when the electricity price is high.

4.2 Simulation of DC microgrid

In a DC microgrid, there is an AC-DC converter that connects the wind turbine to a DC microgrid, a DC-AC converter that connects the AC load to the DC microgrid, and an AC-DC converter that connects DC microgrid to the main grid which is AC. The efficiencies of AC-DC converters are set to 0.87 while the efficiency of the DC-AC converter is set to 0.92.

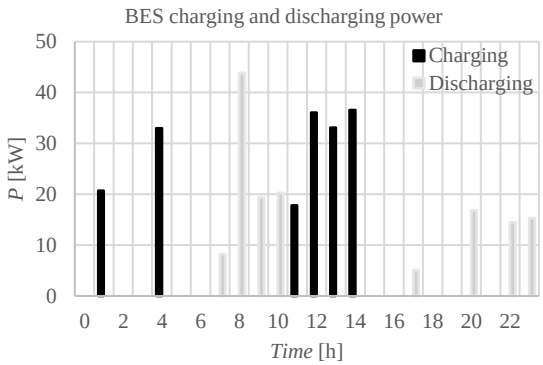


Figure 12. Charging and discharging power of BES in DC microgrid

The charging and discharging power of BES in the DC microgrid is presented in Fig. 12. It can be seen that the battery is charging during the night when the electricity price is low and during the day when there is a peak of generation from PV power plants and wind turbines. The battery is discharging in the morning to cover DC load demand and at 20:00 to cover peak AC load demand.

4.3 Simulation of hybrid microgrid

In a hybrid AC-DC microgrid, there is one large bidirectional AC-DC converter that connects the DC part of the microgrid with the AC part of the microgrid. The efficiency of the bidirectional converter is 0.9.

The charging and discharging power of BES in the hybrid AC-DC microgrid is presented in Fig. 13. It can be seen that the battery is charging during the night when the electricity price is low and during the day when there is a peak of generation from PV power plants and wind turbines. The battery is discharging in the morning to cover DC load demand and at 20:00 to cover peak AC load demand. The battery is also discharging at 22:00 and 23:00 when the car is connected to an EV charging station.

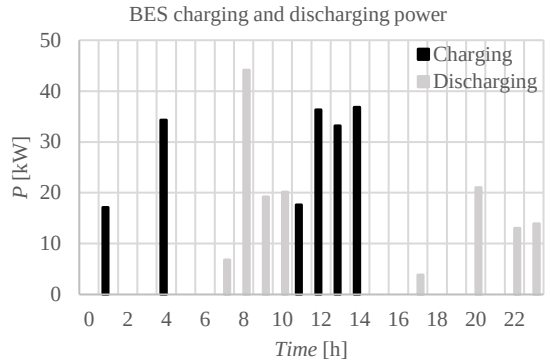


Figure 13. Charging and discharging power of BES in hybrid AC-DC microgrid

4.4 Comparison of simulation results

When simulations for AC. DC and hybrid AC-DC microgrids are done, results of the simulation can be compared. Fig. 14 presents a comparison of battery SOC for AC, DC, and hybrid AC-DC microgrids. Battery SOC at the start of the day is 50%. The battery in the AC microgrid is charged to 80% from 3:00 to 6:00 while batteries in DC and hybrid microgrids are charged up to 75%. Batteries in all three topologies are discharging from 6:00 to 10:00 and then charging from 10:00 to 14:00 when there is excess energy from PV power plants and wind turbines. The battery in all three topologies is charged to 80% from 14:00 to 16:00 and then the battery in the AC microgrid is discharged to 66% while batteries in DC and hybrid microgrids are discharged to 77% and then discharged at 20:00 to cover peak load demand. Battery SOC at the end of the day is 50%.

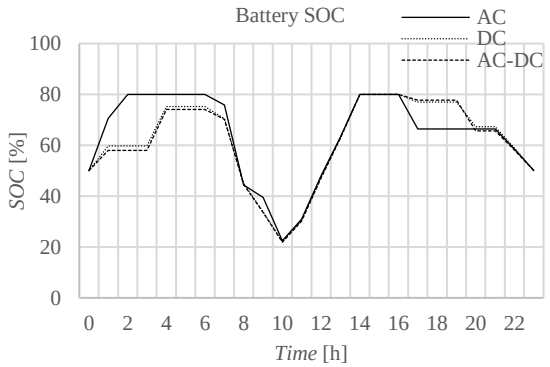


Figure 14. Comparison of battery SOC for AC, DC, and hybrid AC-DC microgrid

Fig. 15 represents a comparison of electricity received from the grid for AC, DC, and hybrid AC-DC microgrids. Most electricity is received from the grid during the night when the electricity price is low and in the evening during periods of peak load demand. There is no electricity received from 10:00 to 14:00 due to electricity generation from PV and wind.

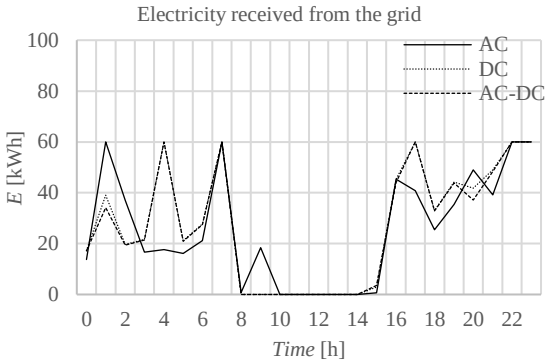


Figure 15. Comparison of electricity received from the grid for AC, DC, and hybrid AC-DC microgrid

A comparison of electricity received from the main grid, electricity cost, losses in a microgrid, and overall microgrid efficiency for AC, DC, and hybrid AC-DC microgrids is presented in Table 2. Energy losses in microgrids occur because AC-DC and DC-AC power converters are not ideal. Energy losses in a microgrid can be calculated according to eq. 3.

$$E_{\text{losses}} = (E_{\text{pv}} + E_{\text{wt}} + E_{\text{grid}}) - E_{\text{load}} \qquad (3)$$

Where:

- E_{losses} [kWh] – microgrid losses;
- E_{pv} [kWh] – generation from PV power plants;
- E_{wt} [kWh] – generation from wind turbines;
- E_{grid} [kWh] – electricity received from the grid;
- E_{load} [kWh] – total load demand;

The overall efficiency of a microgrid is a good indicator that is suitable for microgrid comparison (Manandhar et al., 2016). The overall efficiency of the microgrid can be calculated according to eq. 4.

$$\eta_{\text{microgrid}} = E_{\text{load}} / (E_{\text{pv}} + E_{\text{wt}} + E_{\text{grid}}) \qquad (4)$$

Where:

- $\eta_{\text{microgrid}}$ – overall microgrid efficiency.

Table 2. Comparison of microgrid topologies

	AC	DC	Hybrid
PV [kWh]	343.78		
WT [kWh]	189.89		
AC load [kWh]	600		
DC load [kWh]	400		
Electricity [kWh]	617.43	663.06	650.35
Cost [EUR]	69.80	75.15	73.62
Losses [kWh]	151.10	196.73	184.02
Efficiency [%]	86.87	83.56	84.46

In the AC microgrid, 617.43 kWh of electricity is received from the utility grid, in the DC microgrid, 663.06 kWh of electricity is received from the utility grid, and in the hybrid AC-DC microgrid 650.35 kWh of electricity is received from the utility grid. The cost for electricity in the AC microgrid is 69.80 EUR, in the DC microgrid cost for electricity is 75.15 EUR and in a hybrid AC-DC microgrid the cost for electricity is 73.62 EUR.

Due to different numbers of power converters, the amount of power loses is different in all three microgrid topologies. A high number of power converters, the higher rated power of converters, or the lower efficiency of converters result in higher power losses in the microgrid. The low number of power converters, the lower rated power of converters, or the higher efficiency of converters results in lower power losses in the microgrid. Higher power losses mean that more electrical energy is required to cover the whole load demand and power losses which results in the lower overall efficiency of the microgrid topology. Lower power losses mean that less electrical energy is required to cover the whole load demand and power losses which results in the higher overall efficiency of the microgrid topology.

Losses in AC microgrid are 151.10 kWh, in DC microgrid losses are 196.73 kWh and in hybrid AC-DC microgrid, losses are 184.02 kWh. The overall efficiency of the AC microgrid is 86.87%, the overall efficiency of the DC microgrid is 83.56% and the overall efficiency of the hybrid AC-DC microgrid is 84.46%.

According to the simulation result, AC microgrid topology is most suitable for selected system components. It is because of high AC load demand, significant generation from wind turbines, and a significant amount of electricity received from the grid. DC microgrid topology would prove suitable if DC load demand was higher than AC load demand. Hybrid AC-DC microgrid topology has good efficiency due to significant generation from PV power plants and significant DC load demand.

5 Conclusion

Microgrids will be more and more present as a suitable solution for connecting distributed generation with consumers at same the location in a small independent power system. Three main topologies of microgrids according to the power between microgrid components of microgrid systems are AC microgrid, DC microgrid, and hybrid AC-DC microgrid. The difference between microgrid topologies is the number of AC-DC converters that are not ideal so there are power losses during power conversion.

AC microgrids are suitable for existing consumers which are largely AC and the transition from the existing power system to the AC microgrid is very simple and fast. DC microgrids are suitable for applications with battery storage, generation from DC sources like PV power plants, and significant DC consumers like EV charging stations, computers, and LED lighting. Hybrid AC-DC microgrids with one bidirectional AC-DC converter are suitable for applications where DC generation and consumers as well as AC generation and consumers are equally represented in the microgrid.

Acknowledgments

This work was funded by the European Union through the European Regional Development Fund's Operational Programme Competitiveness and Cohesion under project KK.01.1.1.04.0034 Connected Stationary Battery Energy Storage.

References

Ali, S., Zheng, Z., Aillerie, M., Sawicki, J. P., Péra, M. C., & Hissel, D. (2021). A review of DC microgrid energy management systems dedicated to residential applications. *Energies*, *14*(14), 1–26. https://doi.org/10.3390/en14144308

Atia, H. R., Shakya, A., Tandukar, P., Tamrakar, U., Hansen, T. M., & Tonkoski, R. (2016). Efficiency analysis of AC coupled and DC coupled microgrids considering load profile variations. *2016 IEEE International Conference on Electro Information Technology (EIT)*, *2016-Augus*, 695–699. https://doi.org/10.1109/EIT.2016.7535324

Cabana-Jiménez, K., Candelo-Becerra, J. E., & Santos, V. S. (2022). Comprehensive Analysis of Microgrids Configurations and Topologies. *Sustainability (Switzerland)*, *14*(3), 1–25. https://doi.org/10.3390/su14031056

Gao, F., Wang, X., Yang, P., Kou, S., & Sun, M. (2020). Research and Simulation of Hybrid AC/DC Microgrid. *2020 4th International Conference on HVDC (HVDC)*, 1276–1280.

Liang, B., Kang, L., Zhang, Z., Hu, B., Zhao, Y., Zhang, Z., & Yao, N. (2018). Simulation Analysis of Grid-connected AC/DC Hybrid Microgrid. *2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, 1–6.

Manandhar, U., Ukil, A., & Jonathan, T. K. K. (2016). Efficiency comparison of DC and AC microgrid. *2015 IEEE Innovative Smart Grid Technologies - Asia, ISGT ASIA 2015*. https://doi.org/10.1109/ISGT-Asia.2015.7387051

Minemura, K., Owaki, D., Yukita, K., Goto, Y., Ota, T., Miyoshi, H., Beibei, W., Yang, L., & Hirose, K. (2020). Study of a microgrid using a private power generator during a utility grid failure. *2019*

IEEE Third International Conference on DC Microgrids (ICDCM), 1–6. https://doi.org/10.1109/icdcm45535.2019.9232677

Rezkallah, M., Chandra, A., Singh, B., & Singh, S. (2019). Microgrid: Configurations, Control and Applications. *IEEE Transactions on Smart Grid*, *10*(2), 1290–1302. https://doi.org/10.1109/TSG.2017.2762349

Sahoo, B., Routray, S. K., & Rout, P. K. (2021). AC, DC, and hybrid control strategies for smart microgrid application: A review. *International Transactions on Electrical Energy Systems*, *31*(1), 1–53. https://doi.org/10.1002/2050-7038.12683

Werth, A., Kitamura, N., Matsumoto, I., & Tanaka, K. (2015). Evaluation of centralized and distributed microgrid topologies and comparison to open energy systems (OES). *2015 IEEE 15th International Conference on Environment and Electrical Engineering, EEEIC 2015 - Conference Proceedings*, 492–497. https://doi.org/10.1109/EEEIC.2015.7165211

Zhang, F., Meng, C., Yang, Y., Sun, C., Ji, C., Chen, Y., Wei, W., Qiu, H., & Yang, G. (2015). Advantages and challenges of DC microgrid for commercial building: A case study from Xiamen university DC microgrid. *2015 IEEE 1st International Conference on Direct Current Microgrids, ICDCM 2015*, 355–358. https://doi.org/10.1109/ICDCM.2015.7152068

Zhu, X., Premrudeepreechacharn, S., Sorndit, C., Meenual, T., Kasirawat, T., & Tantichayakorn, N. (2019). Design and Development of a Microgrid Project at Rural Area. *2019 IEEE PES GTD Grand International Conference and Exposition Asia, GTD Asia 2019*, 877–882. https://doi.org/10.1109/GTDAsia.2019.8716001

Intelligent Systems

Saša Brdnik, Boštjan Šumak
Adaptive and intelligent user interfaces and how they adapt to users: A systematic mapping study

Ivan Magdalenić, Neven Vrček, Boris Tomaš, Karl Sölvi Guðmundsson
Localization of drone operator using Unmanned Aerial Vehicles: Preliminary research

Ena Barcic, Igor Tomičić, Mario Harjač
A survey on face template protection methods

Marko Horvat, Ljudevit Jelečević, Gordan Gledec
A Comparative study of YOLOv5 models performance for image localization and classification

Jože M. Rožanec, Patrik Zajec, Jelle Kaizer, Elena Trajkova, Blaž Fortuna, Bor Brecelj, Beno Šircelj, Dunja Mladenić
Enhancing manual revision in manufacturing with AI-based defect hints

Jože Rožanec, Tadej Krivec, Vid Keršič, Larsen Cundrič, Blaž Stojanovič, Marko Zeman, Ivan Bratko
Bicycle Sharing Systems meet AI: forecasting bicycles availability and decision-making

Josip Kunsabo, Jasminka Dobša
A Systematic Literature Review on Topic Modelling and Sentiment Analysis

Ljiljana Zekanović-Korona, Jurica Grzunov
Facilitating Smart Tourism Destination Development: Towards a Model of Information System for Smart Tourism

Adaptive and intelligent user interfaces and how they adapt to users: A systematic mapping study

Saša Brdnik, Boštjan Šumak
Faculty of Electrical Engineering and Computer Science, Institute of informatics
University of Maribor
Koroška cesta 46, 2000 Maribor, Slovenia
{sasa.brdnik, bostjan.sumak}@um.si

Abstract. *This paper presents a systematic mapping study on Intelligent (IUI) and Adaptive (AUI) user interfaces to (i) get an overview of existing research topics and authors, (ii) identify publication trends, (iii) uncover potential gaps in research and (iv) observe the use of terms in academia. Terms AUI and IUI are used concurrently in papers; 30% of observed AUIs include some form of artificial intelligence. The focus for both is in solution proposals and validation research. Interfaces mostly adapt to history of users' behaviour, context changes, and users' actions.*

Keywords. Intelligent user interfaces, Adaptive user interfaces, IUI, AUI

1 Introduction

The idea of Intelligent user interfaces (IUI) has been around for decades, while the goal of developing computer systems which have the ability to accommodate themselves to different users, their skill level, experience and needs, is still very relevant today. Intelligent user interfaces (Maybury, 1998) are human-machine interfaces that aim to improve the efficiency, effectiveness, and naturalness of human-machine interaction. They achieve this by representing, reasoning, and acting on different models (eg. user, domain, task, discourse, and media models).

The IUI field (Alvarez-Cortes, Zayas-Perez, et al., 2007) is multi-disciplinary, interchanging the ideas from different areas, with emphasis on three core disciplines - Artificial Intelligence (AI), User Modelling (UM) and Human-Computer Interaction (HCI). A holistic meta-study (Völkel et al., 2020) was conducted on the use of intelligence in IUI research in the last 25 years. This sprouted our interest in conducting similar overview of the research area while focusing on how the interfaces are adapted to users and what terms are used to describe these interfaces in practice. Systematic mapping study was conducted in order to present comprehensive results. We hope to add some more insight to the debate on what makes user interfaces intelligent.

The main contributions of this paper are (i) Presenting an overview of primary and secondary studies on IUI, (ii) Conducting a review on methods used for IUI review form HCI point of view and (iii) Presentation of discovered research gaps in the field. The paper is structured as follows. Section 2 briefly discusses taxonomy in the field. Section 3 presents related work of secondary studies, concluded in this area. Section 4 presents the research method used to address the research objectives. Results of the systematic mapping study are presented in Section 5, while section 6 presents the conclusion and future work implications.

2 Adaptive and Intelligent User Interfaces

Adapted and adaptable user interfaces are not a point of interest in this study, but are mentioned for taxonomy clarification. Adapted user interfaces are UIs, that are adapted to the end user at design time, while adaptable user interfaces offer the end user to change (ie. to adapt) the characteristics or functionality of the UI (Schlungbaum, 1997).

Intelligent user interfaces "are human-machine interfaces that aim to improve the efficiency, effectiveness, and naturalness of human-machine interaction by representing, reasoning, and acting" (Maybury, 1998). They adapt at run-time and can make decisions about what, when, why and how to communicate with the user (Maybury, 1998). They decrease task complexity, bringing expertise to the user (in the form of expert critiquing, task completion, coordination), or simply providing a more natural environment for users to interact. IUI field is an overlap of multiple different fields, with Human-computer interaction and Artificial intelligence being the major ones. HCI provides efficient designing techniques for user interfaces and AI is used to automate or build intelligence in said interfaces (Shaikh et al., 2017).

In practice, the terms intelligent user interface and adaptive user interface are used concurrently in papers. The use of AI techniques to design adaptive in-

terfaces has been an accepted approach applied in various domains and applications (Alvarez-Cortes, Zayas-Perez, et al., 2007). Such adaptive interfaces, using AI methods and techniques are considered intelligent user interfaces. The term *adaptive intelligent user interfaces* (Ahmad et al., 2004, Raheel, 2016) has sometimes been used in papers proposing user interfaces with intelligent adaptive mechanisms, which are capable of monitoring the user behavior and then adapting the interface accordingly, beyond of the scope of predefined rules. Tahir et al. (Tahir, 2015) recognized *user-adaptive intelligent user interfaces* as a branch or a subset of IUIs. Volkel et al. (Völkel et al., 2020) observed the concurrence in their meta study of 504 papers on Intelligent user interfaces and confirmed that the researchers might call an entity both "intelligent" and "adaptive". However it is clear that intelligent systems with interfaces cannot be labeled as IUI, if they are only intelligent in machine point of view, but not in user or HCI point of view (Shaikh et al., 2017).

The terms are often confused and intertwined in papers (e.g. Patrick, 2003, Alvarez-Cortes, Zarate, et al., 2009, Alvarez-Cortes, Zayas-Perez, et al., 2007), as some researchers declare the terms as synonyms, even though not all adaptive user interfaces can be labeled as intelligent. By definition, IUIs are aimed to improve the efficiency, effectiveness and naturalness of human machine interaction by representing, reasoning and acting on models of users, domain, task, discourse and media. Therefore, "intelligent systems must be data driven, self-aware and have the capability to learn over time from experience" Mezhoudi and Medina, 2015. In this study we were interested in both, intelligent and adaptive user interfaces.

3 Related work

Similar secondary studies have already been conducted and are presented in this section. Gonçalves et al. (Gonçalves et al., 2019) conducted a systematic literature review on intelligent user interfaces in 2019, though only presented preliminary results with little data extraction and interpretation were presented. They investigated the design of IUIs in the context of contemporary software systems. Authors identified context model, dialog model and user model as most used models in IUI design. They also report IUIs mostly adapt to context of use, user's actions, user's preferences, context changes, device characteristics, platform characteristics, user behaviors and users´ characteristics. The research questions focused on domains where IUIs are used, trends regarding use if IUI in contemporary software systems and IUI evaluation were left as future work within this SLR.

In 2017 Sanchez et al. (Sanchez et al., 2017) conducted a similar systematic mapping study on IUIs, with focus on Ambient Assisted Living (AAL) technologies. The study included 43 papers and observed type of used media models (graphics, natural language, gesture), type of solution (application, mobile system, website), relationship between IUIs and Ambient assisted living and overview f applications that have adopted IUIs. Authors highlighted the importance of designing user-friendly interfaces.

Volkel (Völkel et al., 2020) conducted a Meta-Analysis of articles, published in ACM's database in the span of last 25 years. Study was conducted as text analysis and focused on intelligent aspects of IUIs from the view of the IUI community. Paper focused on what researchers deem intelligent, what characteristics are used to describe IUIs and how else are they characterized. Authors noted that considerable amount of articles never use the term 'intelligen*'. Authors conducted it has remained elusive what exactly renders an interactive system or user interface deserving of the term intelligent.

Miraz et al. (Miraz et al., 2021) conducted a survey on universal usability of IUI and AUI on 165 papers. They focused on whether the UI should be adaptive under system control or be made adaptable under user control, and conducted a performance evaluation of plasticity. They concluded adaptations can improve usability, but the trade-offs should be further analyzed. A taxonomy proposition for comparison of the various AUIs is also presented in their work.

To the best of our knowledge, none of the recently conducted secondary studies focus on the similarities and differences in the use of terms AUI and IUI in academia.

4 Method

The goal of this study is to determine current state-of-the-art of the research area, observe how are adaptive and intelligent user interfaces designed and developed, and to discover some of the possible research gaps. The study was conducted by following the guidelines for conducting a systematic mapping study, written by Petersen et al. (Petersen et al., 2008). The process was extended with iterative improvement of classification scheme during the pilot study, conducted on 20 papers. The data extraction and mapping process was based on the review of the whole paper, not just abstract and keyword screening.

4.1 Defining research questions

The objective of this study was to obtain a comprehensive overview of current research in the field of Intelligent User Interfaces, with emphasis on the assessment of human-computer interaction aspect. This lead to the following research questions:

- RQ1: What are key bibliometric facts of included publications?
  - RQ1.1: What are the publication counts in recent years?

  - RQ1.2: Which journals and conferences include papers on IUI and AUI?
  - RQ1.3: Who are the contributors in the field, who scientifically promote AUI and IUI topics?
- RQ2: How is research in the IUI area conducted?
  - RQ2.1: What types of papers are published in the area?
  - RQ2.2: Which domains do the identified publications address?
- RQ3: How is artificial intelligence included in user interfaces?
  - RQ3.1: What actions do intelligent and adaptive user interfaces perform?
  - RQ3.2: Which algorithms and methods are used to include the intelligence?
- RQ4: How are AUI and IUI designed and developed?
  - RQ4.1: What adaptive criteria are taken into account?
  - RQ4.2: What models are used in interface design?
  - RQ4.4: What programming languages, technologies, frameworks and tools are used?

4.2 Conducting search and screening

After identifying the research questions we defined the appropriate keywords for finding all published papers with topics from IUI. As we wanted to provide a general overview of the research area, broad keywords were used. As Volkel (Völkel et al., 2020) already pointed out, it is evident the articles in the Intelligent user interface research area do not always refer to their solutions with the terms "intelligen*". Therefore we broadened our search to also include term "adaptive", which was one of the most commonly used other terms used to describe them. The final query used was following: *"intelligent user interface" OR "intelligent user interfaces" OR "IUI" OR "adaptive user interface" OR "adaptive user interfaces" OR "AUI"*

There are many possible digital libraries available for conducting a literature search. We chose a wide range of the established ones: IEEE, ACM, ArXiv, Web of Science, Science Direct and Scopus. With pilot preliminary search, 5022 pieces of literature, fitting this criteria were found. Search was then limited with the exclusion criteria, presented in Table 1. The search was conducted on 19th and 20th April 2021.

Number of papers we found with automatic search in the first stage, by following the selected query, is presented in Table 2. All together 5022 possibly relevant papers were acquired. The highest number of possibly relevant studies was found in Scopus digital library.

The first stage of screening process consisted of acquiring the possibly relevant literature automatically from selected databases. In the second stage of screening, exclusion criteria E2 (year), E3 (literature type) and E7 (field) were applied, which reduced the number of possibly relevant papers to 1384. In the third phase, we screened the papers by abstract and applied

Table 1: Inclusion and exclusion criteria

	Criteria	**Description**
IC1	Field	Include studies, addressing IUI or AUI.
IC2	Language	Studies must be written in English.
IC3	Availability	Studies must be accessible electronically.
IC3	Area	Computer science and informatics or HCI literature.
EC1	Area	Exclude any non-computer science or non-HCI literature.
EC2	Year	Exclude any literature, published before 2010.
EC3	Literature type	Exclude any literature which is not journal or conference article (eg. books and theses).
EC4	Language	Exclude any studies, not published in English.
EC5	Duplicates	Exclude any duplicated studies found in multiple databases.
EC6	Availability	Exclude any studies that are unavailable online.
EC7	Field	Exclude studies outside of the IUI or AUI research area.
EC8	Short articles	Exclude articles with number of pages < 4.

Table 2: Papers retrieved from selected digital libraries

Database	Nr. of papers
ACM	298
ArXiv	18
IEEE	501
Science Direct	297
Scopus	2802
Web of Science	1095
Together	5022

E1 (Area) and E7 (Field). In the fourth stage, duplicates were removed (E5) in the chronological order of screening databases (IEEE, ACM, ArXiv, Web of Science, Science Direct and Scopus, respectively). In the fifth stage, we screened the content of 448 potentially relevant papers and further reduced the number of potentially relevant papers to 226.

4.3 Sampling

After obtaining the final number of relevant papers, we conducted sampling. Total number of relevant studies remaining after screening was 226, which represents a fairly small population size. A sampling strategy suggested by Israel (Israel, 1992) was followed - using a sample size of a similar study. Sanchez et al. (Sanchez et al., 2017) conducted the systematic mapping study IUIs in Ambient Assisted Living (AAL) technologies on the population of 43 papers, while Guerino et al. (Guerino and Valentim, 2020), who was observing evaluation and user experience of natural user interfaces, included 56 papers. After observ-

ing similar studies in the field, the final sample of 80 studies was included in our systematic mapping study. Papers were randomized before selection, with every paper having the same probability of being included.

4.4 Classification and data extraction

This section includes detailed overview of our classification scheme, which is later summed up in Table 3. First version of classification scheme was tested on 20 studies. We implemented some iterations during the piloting phase, mostly focused on changing the grouping the selected variables, while also some new variables and some additional possible values for existing variables were brought to our attention. After the iteration, the classification scheme was finalized, and was used as a data extraction tool for all the collected papers.

5 Results

This section presents and discusses the results of classification and data extraction of 80 studies. The objective of this study was to provide an overview on intelligent user interface research field and provide an overview in differences between what is labeled as AUI or IUI in practice. Gathered data is also available in a shared file. [1]

5.1 Key bibliometric facts

This section aims to answer RQ1. As noted in the previous chapters, we included intelligent and adaptive user interfaces in our study, as they are often used concurrently. We therefore firstly noted the terms primarily used for observed user interface in each paper. Use of term in title and in keywords was also weighted. Even if authors did not follow the definitions of IUI and AUI distinction, we noted the term used in the paper. Most papers (63%) used the term adaptive user interface, while 35% used the term intelligent user interface. One paper primarily used the term multi-modal user interface (reffered as MUI in this text). The results with citations to papers are displayed in Figure 4.

The selected papers were published between 2010 and 2021. As illustrated in Figure 1, the number of studies seems to be relatively constant, with some noticeable negative deviations every second year between 2010 and 2018, which could present a trend or could be a result of our sampling. Number of publications has slightly decreased in 2019 and 2020, compared to the trends in previous years. The representation of papers, published in 2021 is incomplete, since the initial query was conducted in March 2021. Number of papers primarily using the term AUI is higher compared to the number of studies, primarily using the term IUI.

[1] https://univerzamb-my.sharepoint.com/:x:/g/personal/sasa_brdnik_um_si/EQCj5P6nJW5PijMi61_VUqABkLDp5L8v12vcRmAH3TEWnQ?e=bjiXu3

Table 3: Research evaluation scheme

ID	Variable	Descriptors
E1	Term used	IUI, AUI, MUI
E2	Lit. type	Conference, Journal
E3	Research type	Validation, Evaluation, Solution proposal, Philosophical Paper, Opinion paper, Experience paper, Lit. review
E4	Source type	Primary, Secondary study
E5	Time scope	Cross-section, Longitudinal
E6	Standpoint	Software engineering/AI, Human-computer interaction
E7	Methodology	Case Study, Experiment, Grounded theory, Literature review, Mapping study, Meta-analysis, Observing, Other, Prototype, Questionnaire, Interview, Systematic literature review, Systematic mapping study, Field study
E8	Domain	Health, Education, Software development, Entertainment/Games, Industry, Communication, Cartography, Transportation, Accessibility, Security, Military, Tourism
E9	Engineering phase	Testing, implementation, design, analysis
E10	Action performed by intelligent entity	Assistance, Adaptation, Analysis, Communication, Creation (of content, etc.), Decision, Detecting or capturing information, Illustration, Improvement, Interacting with user, Learning Modeling, Monitoring, Perform, Prediction, Processing, Recommendation, Selection, Understanding, Utilisation
E11	Adaptivity criteria	Context changes, Device characteristics, Platform characteristics, History of users' behavior, Users' actions, Users' characteristics, Users' needs, Users' preferences, Task, User's emotions
E12	Models used	Context, Dialog, User, Device, Layout, Ontology, Platform, presentation, Task, Domain
E13	Proposed solution	Agent, Algorithm, Application, Approach, Architecture, Concept, Framework, Guidelines, Interface, Model, Not applicable, Other, Software, System, Technique / Method
E14	Environment	Open description of software environment

To get a broader view on who are the contributors in the field, who scientifically promote IUI topics, we observed where are first author's primary institutions located. Most of the first authors are active in Germany (13 authors), India (6 authors), and in Spain, China and Canada (4 first authors each). The largest part of the researchers (46,3%) were active in Europe, while further 35% were active in Asia. Further 11,3% of authors were from North or South America, while

Table 4: Term used

Term	%	Papers
IUI	36%	S3, S6, S8, S9, S10, S15, S16, S17, S18, S20, S22, S27, S29, S36, S37, S38, S49, S53, S59, S61, S62, S64, S66, S68, S70, S72, S76, S77
AUI	63%	S1, S2, S4, S5, S7, S11, S12, S13, S14, S16, S19, S21, S23, S24, S25, S26, S29, S30, S31, S32, S33, S34, S35, S39, S40, S41, S42, S43, S44, S45, S46, S47, S50, S51, S52, S54, S55, S56, S57, S58, S60, S63, S65, S67, S69, S71, S73, S74, S75, S78, S79, S80
MUI	1%	S48

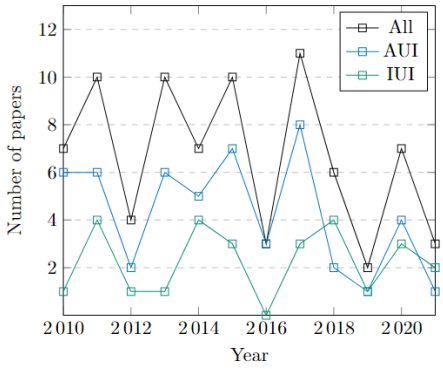


Figure 1: Number of publications by years

a smaller percentage (3,8%) were based in Australia and in Africa. We observed which authors are the most active in the observed fields and promote them scientifically. The most prominent author included in three papers is Dr. Sucheta V. Kolekar from India.

The main observed journals, where observed papers on Intelligent user interfaces were published are Interacting with Computers, Procedia Computer Science, ACM Transactions on Interactive Intelligent Systems and Personal and ubiquitous computing, all but one represented with two papers in our study. All of the mentioned journals are peer-reviewed. The main conferences observed are International Conference on Intelligent User Interfaces with four papers included in our sample, International Conference on Human-Centered Software Engineering (HCSE), with 3 papers included in our sample and ACM SIGCHI Symposium on Engineering Interactive Computing Systems, which is also represented with three papers in our sample. Most of the included papers were conference papers and proceedings (77,5%), while journal papers represented 22,5% of our sample. Division based on the term used is similar, conference papers represent most of the included AUI (78%) and IUI (75%) papers, similar ratio of IUI (25%) and AUI (22%) papers are published in journals.

5.2 Research data

This subsection aims to answer the RQ2 (How is research in the IUI area conducted?). Research type classification per publication year is presented in Figure 2. It is visible that most of the observed studies were solution proposals and validation researches. These representative trend of these two research types is visible in all observed years. No experience or opinion papers were included in our sample, while a small number of literature reviews, evaluation research and philosophical papers were observed (3 papers of each type).

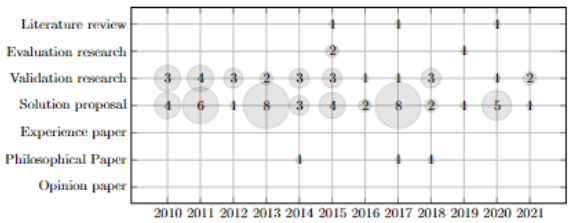


Figure 2: Publications per year by research type

Most of the papers included were solution proposals or validation researches, as visible from Figure 3. Closer look in division of papers by dominantly used term shows solution proposals present similar percentage of AUI and IUI papers (59% and 54% respectively). Same goes for validation research (29% of AUI and 36% of IUI). No philosophical papers or evaluation research were in included for IUI, while we noted one literature review for AUI and two for IUI. Most of the studies (50 papers) have been focused on software development area, with 29 studies offering solution proposals, while 15 of them were classified as validation research. Further, health presented the second most represented application domain with 7 papers. Most common contribution in observed papers was a framework (18 papers: 11 AUI and 7 IUI), followed by an interface (14 papers: 6 AUI and 8 IUI) and an approach proposition (5 papers; 2 AUI and 3 IUI). We have further observed the engineering phase of presented contributions in 68 studies (two literature reviews presented no additional solutions in the paper and were excluded). Most of the proposed solutions were in the testing phase (33 papers) or in the implementation stage (31 papers). Only a few papers presented the contribution in the analysis phase. Ratio between engineering phases is very similar for AUI and IUI papers. Majority of both are in testing phase (45% of AUI and 46% of IUI), while a third are in implementation phase (33% of AUI and 32% of IUI). Further 18% of AUI and 7% of IUI papers presented solutions in design phase, while some papers presented work in the analysis phase (2% of AUI and 11% of IUI).

5.3 Intelligence in user interfaces

In this section we aim to answer the RQ3 (How is artificial intelligence included in user interfaces?). Many

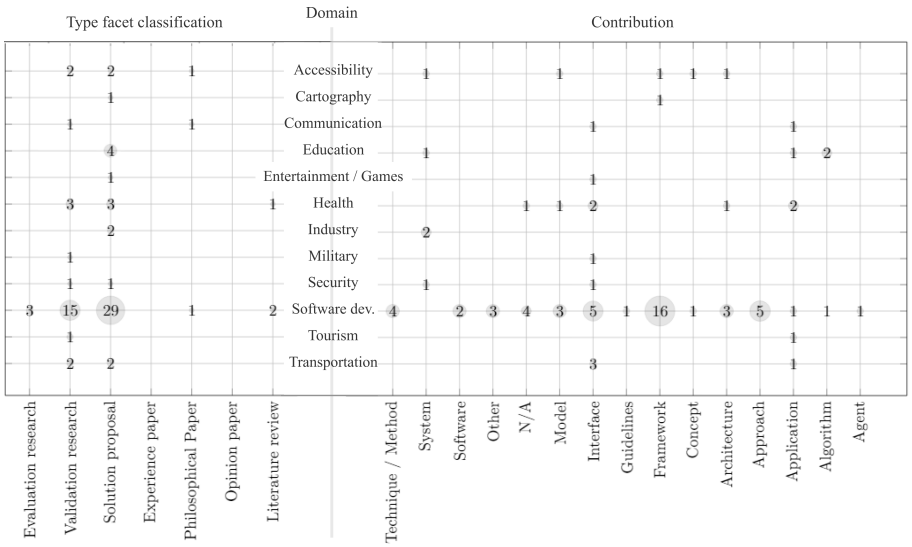


Figure 3: Research facet and type of contribution by domain

publications observed user interfaces perform more than one action. Most commonly performed action by intelligent entities was adaptation, which was described in 48 papers, followed by detecting or capturing information (18 papers), providing assistance (13 papers), recommendation (6 papers), interacting with user (4 papers) and analysis or understanding (3 papers each). Performed actions of user interfaces are visualized in Figure 4 (only primary studies included). The division between AUI and IUI shows they both perform adaptations of user interface or its elements, however, the action is more common with AUI papers. Detecting or capturing information, recommendation, analysis and understanding were observed in the same amount in IUI and AUI papers. More AUI papers focused on providing assistance, while more IUI papers covered creation of content (S49 - algorithm creates tables with unique color and digit combination) and illustration (S8 -virtual illustration of agents movement), conducted modeling and monitoring. It is interesting actions of prediction (S31 - predicting what to offer user next, S34 - predicting the contacts that a user will most likely call next on his mobile phone) and decision making (S31- deciding on UI adaptations) can be observed in AUI papers.

5.4 Design and development process

In this section we aim to answer RQ4 (How are AUI and IUI designed?) with overview of the adaptive criteria taken into account in the design process. In the process off designing the IUIs, there are generally various adaptive criteria into the account. We observed 68 papers (two secondary studies excluded), where researchers described the adaptive criteria. Mean number of adaptive criteria used in UI was 1,6. As visible in Figure 5, the most common criteria for adaptation

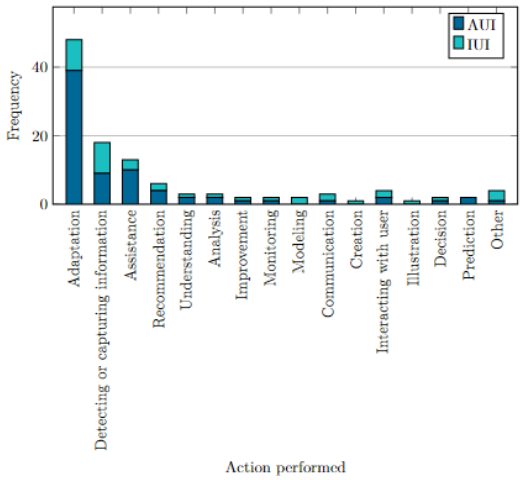


Figure 4: Actions performed by intelligent and adaptive user interfaces

is user's behaviour (25 observed papers), followed by the context change (21 papers), users actions (19 papers), users characteristics (15 papers), users needs (10 papers) and preferences (9 papers), and device characteristics (9 papers). Platform characteristics and users emotion were not as commonly used criteria. In terms of user interface type, papers describing AUI, have mostly used history of users behaviour (19 papers) and context changes (14 papers) as adaptive criteria, while papers describing IUI mostly focused on user actions (7 papers), context changes (7 papers) and history of their behaviour (6 papers). Model-driven approach has been used for the generation and development of adaptive user interfaces. User interface designer or developer specifies one or multiple models from which a UI is later developed or even generated. Researchers in both, AUI and IUI papers most commonly used the user model when they designed user interfaces. Further the task, context and dialog models were also commonly used in the observed papers.

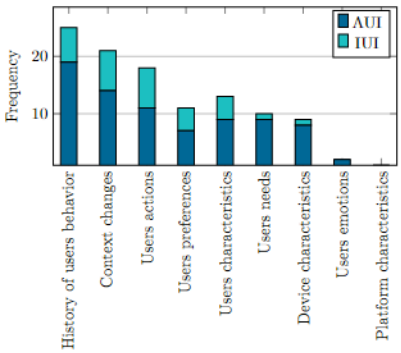


Figure 5: Adaptive criteria

To further analyze the development process of state-of-the-art IUI and AUI propositions, we noted the programming language, frameworks and tools, researchers used in their work. Not all researchers reported the metrics, and in some cases the variables were not applicable for the proposed solution. Researchers reported at least some of the variables in 49 papers. Most commonly the mobile applications were developed for Android OS (6 papers), while most commonly reported programming language was Java (6 papers). In terms of sensors, the most represented sensor was Kinect, which was used in three studies. In terms of software and development tools, three studies reported the use of MyUI Prosperity4All, 2017, an infrastructure, generating the individualized user interfaces and performing adaptations to cater the diverse user needs, devices and environmental conditions during run time. Further three papers reported use of Cedar Studio software.

6 Discussion and conclusion

In this work a systematic mapping study on Intelligent user interfaces was presented with the aim of providing the IUI research community with unbiased, objective and systematic overview of research between 2010-2021. We believe it would be beneficial to periodically conduct such studies to make researchers aware of progress, related work and possible gaps in the research. We observed that most of the research is conducted in the software development domain, with interests also showing in health, accessibility, transportation and education area. With most of the papers being solution proposals or validation research, we observe a lot of suggestions for IUI and AUI, which are already in the implementation or testing phase. Based on the number of publications by years (Figure 1), we can conclude both fields (AUI and IUI) are active and evolving further. Large number of papers (18) proposed new frameworks, which we interpret as development of the area and development process in a more structured and formal manner. Some of them focus on automatic adaptation of user interface (S13), on optimizing the design (S3, S65) and development process (S24, S45), while others focus on accessibility (S67, S79) or data mining (S16). We observed little difference in research approached in AUI and IUI papers. The research trends in research type, research phase, proposed solutions and UX and usability assessment is very similar in both fields. We also note a lot of observed AUI papers (16 papers - 33% of all AUI papers) included artificial intelligence techniques or algorithms. We therefore conclude that the AUI and IUI fields overlap and researchers should aim to separate the terms in future in order to achieve higher level of unification and standardization in the field.

Some of the papers mentioned Maes's (Maes, 1995) list of three major challenges of IUI field, which are presentation (computer-human interaction phase of IUIs), competence (AI methods or techniques) and trust (includes intelligibility and privacy challenges). In our overview we have gathered some additional challenges in IUI area. Savidis (S2) mentions lack of support and standards for including adaptive behavior to the existing non-adaptive systems, while Todi et al. (S12) adds cost of adaptation as a concern for adapting user interfaces. Contradictory to challenges presented by Maes, Ruijten et al. (S70) suggest the use of IUI as means of increasing the trust in autonomous vehicles. Researchers (S15) also expressed concern over potential high costs, that could be imposed by a carelessly picked adaptation and further advocate for strong policies on adaptation. Peissner et al. (S73) expressed concern over possible usability problems, including disorientation and the feeling of losing control in adaptive user interfaces. They advocated for transparency and controllability of automatic adaptations.

We observed no longitudinal studies in our sample, which is surprising, as the most commonly used model in observed IUIs was user model and the most commonly used adaptive criteria was history of user's behaviour, which must obtain at least some user-related data before it can commence the task of personalising the interface. It would be interesting to observe usability assessments over longer periods of time, from users' first interaction with the UI until the point in use when they are comfortable with the UI and when IUI or AUI have enough data to create an integrated user model and can further adapt itself to the user.

6.1 Threats to validity

Threats to the validity are possible in identification of relevant primary research, sampling, data extraction and classification. Some of the studies were excluded in our data query (eg. multi-modal interfaces). A large number of older, ground-setting studies were excluded, as we were interested only more in the recent trends. With randomized sampling some of the important recent studies could have been excluded. Data extraction and classification process was challenging due to wide range of the research area, interdisciplinarity and varying quality and complexity of the observed studies.

To improve validity of initially proposed classification scheme, we conducted a pilot study and improved the scheme over various iterations.

6.2 Future work

One of the main identified challenges is lack of consensus on when is user interface intelligent. Future research will focus efforts towards establishing that. As we observe a lack of evaluation research in the field, an overview of usability and user experience assessments of IUIs would be beneficial to further investigate the state of the research field and current IUIs.

7 Acknowledgments

The authors acknowledge the financial support from the Slovenian Research Agency (Research Core Funding No. P2-0057).

References

Ahmad, A.-R., Basir, O., & Hassanein, K. (2004). Adaptive user interfaces for intelligent e-learning: Issues and trends. *ICEB*.

Alvarez-Cortes, V., Zarate, V., Ramirez Uresti, J., & Zayas, B. (2009). Current challenges and applications for adaptive user interfaces. In-Teh. https://doi.org/10.5772/7745

Alvarez-Cortes, V., Zayas-Perez, B. E., Zarate-Silva, V. H., & Ramirez Uresti, J. A. (2007). Current trends in adaptive user interfaces: Challenges and applications. *Electronics, Robotics and Automotive Mechanics Conference (CERMA 2007)*, 312–317. https://doi.org/10.1109/CERMA.2007.4367705

Gonçalves, T. G., Kolski, C., de Oliveira, K. M., Travassos, G. H., & Strugeon, E. G.-L. (2019). A systematic literature review on intelligent user interfaces: Preliminary results. *Proceedings of the 31st Conference on l'Interaction Homme-Machine: Adjunct*. https://doi.org/10.1145/3366551.3370344

Guerino, G., & Valentim, N. (2020). Usability and user experience evaluation of natural user interfaces: A systematic mapping study. *IET Software*, *14*, 451–467. https://doi.org/10.1049/iet-sen.2020.0051

Israel, G. (1992). Determining sample size. https://books.google.si/books?id=hm-xDAEACAAJ

Maes, P. (1995). Agents that reduce work and information overload. In R. M. BAECKER, J. GRUDIN, W. A. BUXTON, & S. GREENBERG (Eds.), *Readings in human–computer interaction* (pp. 811–821). Morgan Kaufmann. https://doi.org/https://doi.org/10.1016/B978-0-08-051574-8.50084-4

Maybury, M. (1998). Intelligent user interfaces: An introduction. *Proceedings of the 4th International Conference on Intelligent User Interfaces*, 3–4. https://doi.org/10.1145/291080.291081

Mezhoudi, N., & Medina, J. P. (2015). An agile adaptation framework for context-aware user interfaces. *International journal of artificial intelligence*, *2*.

Miraz, M. H., Ali, M., & Excell, P. S. (2021). Adaptive user interfaces and universal usability through plasticity of user interface design. *Computer Science Review*, *40*, 100363. https://doi.org/https://doi.org/10.1016/j.cosrev.2021.100363

Patrick, E. (2003). Intelligent user interfaces: Introduction and survey.

Petersen, K., Feldt, R., Mujtaba, S., & Mattsson, M. (2008). Systematic mapping studies in software engineering. *Proceedings of the 12th International Conference on Evaluation and Assessment in Software Engineering*, 68–77.

Prosperity4All. (2017). Myui. Retrieved July 26, 2021, from http://www.prosperity4all.eu/related-projects/myui/

Raheel, S. (2016). Improving the user experience using an intelligent adaptive user interface in mobile applications. *2016 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET)*, 64–68. https://doi.org/10.1109/IMCET.2016.7777428

Sanchez, C., Cedillo, P., & Bermeo, A. (2017). A systematic mapping study for intelligent user interfaces - iui. *2017 International Conference on Information Systems and Computer Science (INCISCOS)*, 361–368. https://doi.org/10.1109/INCISCOS.2017.34

Schlungbaum, E. (1997). Individual user interfaces and model-based user interface software tools. *Proceedings of the 2nd International Conference on Intelligent User Interfaces*, 229–232. https://doi.org/10.1145/238218.238330

Shaikh, S., Sawand, M., Khan, N., & Solangi, F. B. (2017). Comprehensive understanding of intelligent user interfaces. *International Journal of Advanced Computer Science and Applications*, *8*.

Tahir, R. (2015). Analyzing the intelligence in user interfaces. *2015 SAI Intelligent Systems Conference (IntelliSys)*, 674–680. https://doi.org/10.1109/IntelliSys.2015.7361213

Völkel, S. T., Schneegass, C., Eiband, M., & Buschek, D. (2020). What is "intelligent" in intelligent user interfaces? a meta-analysis of 25 years of iui. *Proceedings of the 25th International Conference on Intelligent User Interfaces*, 477–487. https://doi.org/10.1145/3377325.3377500

Localization of drone operator using Unmanned Aerial Vehicles: Preliminary research

Ivan Magdalenić, Neven Vrček, Boris Tomaš
Faculty of Organization and Informatics
University of Zagreb
Varaždin, Pavlinska 2
{ivan.magdalenic, neven.vrcek, boris.tomas}@foi.unizg.hr

Karl Sölvi Guðmundsson
Faculty of Electrical and Computer Engineering
University of Iceland
Sæmundargötu 2, 102 Reykjavík
karlsg@hi.is

Abstract. *The paper presents the results of a preliminary study of the localization of drone operator using other multiple Unmanned Aerial Vehicles. The emphasis is put on hardware testing, to be mounted on UAVs, in order to detect, identify and track drone signal and to enable localization of drone operator. Since UAVs can carry limited cargo mass, the challenge is to find optimal hardware in terms of ratio of mass and receiving signal capabilities and processing power. The paper describes results of the experiment where hardware, to be mounted on UAVs, is used to measure the signal strength of the intruder drone operator.*

Keywords. Localization, UAV, drone

1 Introduction

The popularity of drones is growing day by day and more and more drones of different sizes and capabilities are filling the sky above us. The main difference between Unmanned Aerial Vehicles (UAVs) and drone is that UAV needs to have autonomous flight capabilities, whereas drones do not. In addition to business and recreation, drones could be used for illegal activities such as unauthorized entry into prohibited airspace and for invading privacy. In the rest of the paper, a drone that does illegal activities will be called an intruder drone and the person who controls it will be called the intruder drone operator. Although the intruder drone itself is visible in the sky, it is difficult to locate the intruder drone operator to prevent him to continue or repeat illegal activities.

The idea of our research is to locate a drone operator by measuring the strength of its wireless signal at several locations at the same time and using the obtained data to calculate its location. There are several challenges that need to be addressed. The first challenged is to identify the right wireless signal. The radio frequency spectrum is often saturated and it is not clear which signal belongs to the intruder drone and its operator. The situation is particularly pronounced in urban areas. The second challenge is to track the intruder drone as it is very fast and can change its location quickly. Our answer to both challenges is to use several UAVs (swarm) to get close to the intruder drone, track it and to catch the strongest signal in the area. Once the signal emitted by the intruder drone is identified, the next step is to identify the intruder drone operator signal by searching the same pattern in other signals in the same frequency band. The other advantage of this approach is to have multiple measured values of signal power on different location which makes it easier to calculate the exact location of the intruder drone operator, which moves significantly slower than drone

The paper is organized as follows: the second chapter brings related work, the third chapter describes our proposal of drone localization using UAVs, followed by a description of the first experiment, and the last part brings discussion, conclusion and future work.

2 Related work

Localization of signal sources by measuring signal strength and calculating position is not new and has been used for decades especially for military purposes. (Jawhar, 2017) and (Hayat, 2016) provide an overview of the communications and architectures used in UAVs. It was pointed out in the papers that a smaller number of drones use the pure IEEE 802.11 standard for communication, while most drones use their own modified version of the IEEE 802.11 standard. We single out here two papers (Bisio, 2018) and (Ezuma, 2020) dealing with the detection and classification of drones, based on information available from captured communication between the drone and the drone-controlling system. Learning techniques were used to identify and classify drones. The paper (Ezuma, 2020) lists the equipment used: 6 GHz bandwidth Keysight MSOS604A oscilloscope with a maximum sampling frequency of 20 GSa / s, 2 dBi omnidirectional antenna

(for short-range detection), and 24 dBi Wi-Fi network antenna (for detection) at greater distances). Antennas are used in the 2.4 GHz spectrum. Common to both approaches is the use of stationary equipment of large dimensions and mass. Such approach is not practical nor suited for real life situations. That is why our approach is based on small, lightweight and fast constellation of drones controlled by only one operator who is driving entire swarm by the software we developed as a part of our research.

Similar research to our approach was conducted by usage of fixed-wing unmanned aerial vehicles (UAVs) and described in (Frey et all. 2008) and (Fray, 2008) were different flight parameters were investigated to locate a moving target. Our approach differs in that we use UAVs with multiple rotors that do not have the ability to stay in the air for a long time and that cannot carry more additional weight. Therefore, flight parameters and detection algorithms must be adapted specifically for such equipment. Compared to fixed-wing, multirotor UAVs have a greater manoeuvrability, thus allowing more flexibility in terms of constellation design and search patterns.

3 Localization of drone operator using UAVs

The basic idea of localization of the intruder drone operator is shown in Figure 1. Several UAVs need to approach the intruder drone as shown in Figure 2.

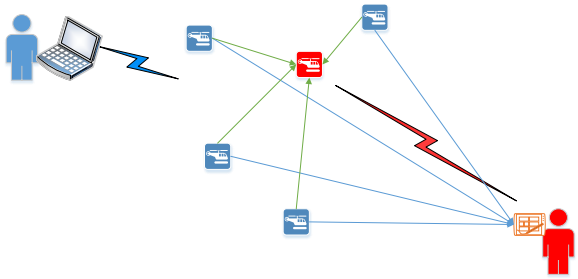


Figure 1. Localization of the intruder drone

We expect to find the intruder drone signal to be the strongest in the vicinity of the drone. The next step is to search the signal of the same pattern that belongs to the intruder drone operator. That signal is received by each UAV and we expect to measure different values of the signal receive power. The final step is to calculate the position of the intruder drone operator based on different location of UAVs and its measured values.



Figure 2. UAV is approaching to intruder drone

Each UAV has equipment mounted for searching and recording of radio signals as shown in Figure 3. The equipment consists of: Raspberry Pi 4 model B with 8 GB RAM, Wireless network adapter Alfa AWUS036ACH, GPS module Ublox NEO M8U and LTE modem HUAWEI E3372. We chose this network adapter because it can be set to monitor mode, which means it can receive all frames on a certain channel. By researching the available technology, we concluded that network adapters that have implemented one of the listed chipsets can be set in monitor mode: Atheros AR9271, Ralink RT3070, Ralink RT3572, Ralink RT5572, Realtek RTL8812AU and Ralink RT5370N. In future research, we plan to acquire other smaller network adapters based on one of the listed chipsets.

We divided our project into two phases. In first phase, we are dealing with intruder drones that are communicating with pure 802.11 protocol. This phase is the proof of concept of our approach and test of our equipment. In the second phase, we plan to use machine learning to find intruder drone signal pattern in radio signal.

For the first phase, we made a software in Python consisting of server part on each drone and desktop application on laptop. Figure 4 shows part of GUI of desktop application. Both, intruder drone and intruder drone operator can be identified by their MAC addresses, which makes it easier to distinguish their radio signal. Each UAV can scan different Wi-Fi channel which speeds up the search for intruder drone signal. After the radio signal is identified, channel is locked and all UAVs are listening to the same Wi-Fi channel.



Figure 3. Equipment for searching and recording of radio signals



Figure 4. Part of GUI of desktop application.

The challenge is to find the best constellation of UAVs in order to speed up the search of radio signal in first step and then to measure the signal strength in the second step. After the measured values are sent to server, the location of the intruder drone operator can be calculated.

Distance will be calculated using the Friis transmission equation (Friss, 1946) where frequency is going to be set to middle of receiving channel, Power component is going to be set to a constant and ignored considering the same signal is received by multiple UAVs. The plan is to use at least four UAVs to exactly calculate the location of the intruder drone. Additional experiments has to be made to determine if more UAVs will increase accuracy and speed of localization.

4 Experiment

The aim of the first experiment was to test equipment and software shown in Figure 3 and to check the feasibility of our approach.

The intruder drone operator is actively operating the intruder drone from its well-known location. GPS coordinates of the intruder drone operator were determined at the start of the experiment for the purpose of later distance calculation. The equipment and software are turned on and are slowly moving away from intruder drone operator. The radio signal is measured until the equipment is so far away that the signal from intruder drone operator is too weak to be received.

The experiment is repeated several times and dependence of received signal strength related to distance is presented in Figures 5, 6 and 7.

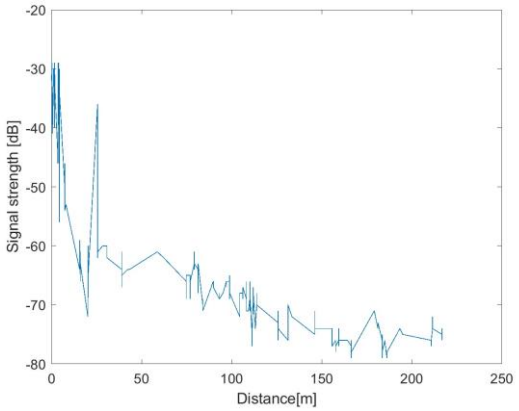


Figure 5. Dependence of received signal strength related to distance – first case

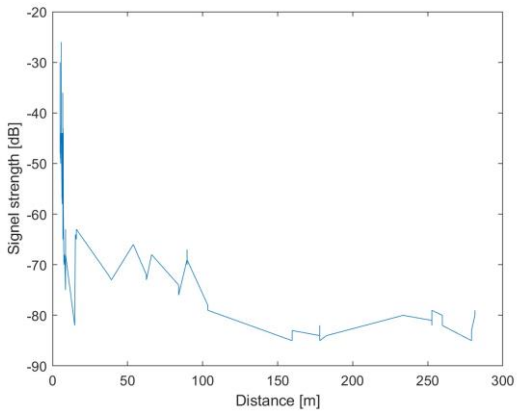


Figure 6. Dependence of received signal strength related to distance – second case

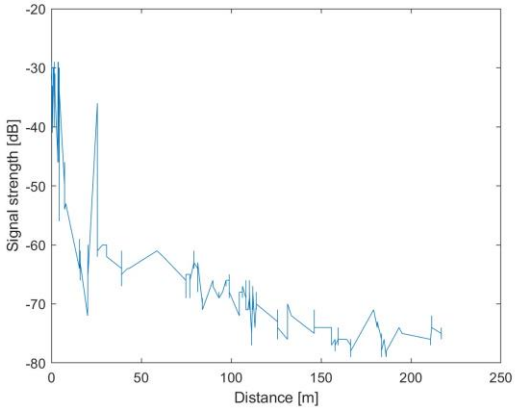


Figure 7. Dependence of received signal strength related to distance – third case

As shown in figures, there is a clear relation between distance of the source of the signal and signal strength and that can be used to determine the location of the source of the signal. In the first 150 meters signal strength is satisfactory and large number of frames are received. After 200 meters, the signal strength is becoming too weak to be received and only a small number of frames are received.

In figures there is a clear limitation of used hardware regarding maximal distance of the receiver from the signal source, and it is about 250 meters. However, it has to be noted that in real scenarios signal might be partially obscured by buildings, trees, etc. This means that this relationship is not that precise. Our future work is directed towards research of such use cases and the development of algorithm capable to compensate eventual signal obstruction.

Since this research is focused on testing the equipment and evaluating the feasibility of our approach, no more experiments were done to determine the ultimate capabilities of radio signal detection. It is planned to do this on equipment mounted on UAVs.

5 Discussion

The advantage of using several UAVs to locate the intruder drone operator, is in short time for identifying and tracking of a target and in flexibility to adopt to the current circumstances in the area. There are also several disadvantages of this approach and the most significant are the short time the drone stays in the air, the size and weight of the signal receiver and the processing power limited by battery power. Although, there are much powerful receivers available on the market, most of them are too big and too heavy to be mounted on a small UAV. The next step is to find an optimal constellation shape of UAVs that will adopt to different scenarios on the ground and that will annul shortcomings of mounted hardware on UAVs by flexible and fast placement in the area.

The conducted experiment has shown some physical and hardware limitations of our equipment. Although the source of our signal in the experiment transmitted packages with constant rate, the rate of the received packages was decreasing with the distance. This means that we received fewer packets with lower power at a greater distance from the source of the radio signal. Although we received packages at distance of 250 meters from the source, we expect that at this distance there will be too few packages received by all four UAVs to calculate intruder drone location. We are planning to mount the equipment used in this experiment on the UAVs and to repeat measurements in order to find exact limitations of our equipment and our approach.

5 Conclusion

The paper presents the results of the experiment where the special hardware is used to measure the signal strength of the intruder drone operator in order to test ability to use those measured values to calculate intruder operator exact location. The experiment shows that it is possible to combine light and small hardware which can be mounted on UAVs and which has enough receiving capabilities and processing power to detect, identify and track radio signal of the intruder drone operator.

The future work is going to follow the next steps on the project and those are mounting the hardware on the UAVs and finding the optimal constellation of UAVs in order to locate the intruder drone operator.

Acknowledgments

This work has been fully supported by Croatian Science Foundation under the project IP-2019-04-4864.

References

Bisio I., Garibotto C., Lavagetto F., A. Sciarrone, and S. Zappatore,. “Unauthorized amateur UAV detection based on WiFi statistical fingerprint analysis,” IEEE Commun. Mag., vol. 56, no. 4, pp. 106–111, Apr. 2018.

Ezuma M., Erden F., Kumar Anjinappa C., Ozdemir O. and Guvenc I., "Detection and Classification of UAVs Using RF Fingerprints in the Presence of Wi-Fi and Bluetooth Interference," in IEEE Open Journal of the Communications Society, vol. 1, pp. 60-76, 2020, doi: 10.1109/OJCOMS.2019.2955889

Frew E.W., Lawrence D. A., Morris S. Coordinated standoff tracking of moving targets using Lyapunov guidance vector fields. Journal of

guidance, control, and dynamics 31 (2), 290-306, 2008.

Frew E.W. Sensitivity of cooperative target geolocalization to orbit coordination. Journal of Guidance, Control, and Dynamics 31 (4), 1028-1040, 2008.

Friis H. T., "A Note on a Simple Transmission Formula," in Proceedings of the IRE, vol. 34, no. 5, pp. 254-256, May 1946, doi: 10.1109/JRPROC.1946.234568

Jawhar I. et al. Communication and Networking of UAV-Based Systems: Classification and Associated Architectures. Journal of Network and Computer Applications · February 2017, DOI: 10.1016/j.jnca.2017.02.008

Hayat S. et al. Survey on Unmanned Aerial Vehicle Networks for Civil Applications: A Communications Viewpoint. IEEE Communications Surveys & Tutorials (Volume: 18 , Issue: 4 , Fourthquarter 2016), DOI: 10.1109/COMST.2016.2560343

A survey on face template protection methods

Ena Barčić, Igor Tomičić, Mario Harjač
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, Varaždin, Croatia
{enbarcic, itomicic, mharjac}@foi.unizg.hr

Abstract. *Biometric authentication has become one of the most popular ways of protecting information, with techniques such as "face unlocking" becoming more and more prevalent. For the purpose of authentication, a face image needs to be saved in the form of a template. To ensure that our information is protected, biometric face template protection methods have been created. In this paper we will present an overview of template protection methods, show their effectiveness and their uses.*

Keywords. Face template protection, Key Binding, Key Generating, Cancellable Biometrics, Hybrid Methods, Homomorphic Encryption

1 Introduction

In today's world, conventional forms of protection like PINs, passwords, tokens, etc. are stolen relatively easily, disclosed and reused by hackers (Sarkar, Singh 2020). To get the highest possible degree of security, biometric traits are often used as keys for unlocking information. Biometric attributes or biometric traits like fingerprint, iris, face (2D, 3D), etc. are extremely individual. The human face is one of the most distinct features, and also one of the most popular physiological characteristics used in commercial biometric systems.

In this paper, we will show the most common biometric facial template methods, determine which ones are the most commonly used today and present their advantages and disadvantages.

2 Biometrics and biometric templates

According to ISO/IEC 30136:2018 Information technology — Performance testing of biometric template protection (BTP) schemes (Shahreza et al. 2022), each scheme needs to include four main properties:

- Cancelability: if compromised, we have to be able to cancel the enrolled template and replace it with a new one.
- Unlinkability: There should be no link between different protected templates from the same unprotected (original) biometric template.
- Irreversibility: It should be difficult or impossible to recover the original biometric template from the protected one.
- Recognition Performance: The protected templates should allow for accurate recognition, without accuracy degradation.

Among most notable problems with biometrics are non-revocability and privacy compromises. Non-revocability means that biometric information in the form of a biometric database/template, if compromised, cannot be replaced or revoked. Privacy compromises consist of three categories: secrecy of biometric data, secrecy of biometric information and privacy of a user's identity. Secrecy of biometric data means that stored templates can be used for template recovery. Secrecy of biometric information means that various biometric traits can be stored together, so if one is compromised, another one can be used. Privacy of a user's identity means that biometric templates, stored in different biometric databases, can be used for cross matching.

Biometric systems are vulnerable to several types of attacks at different stages, the four most common ones being (Sarkar, Singh 2020):

1. Attacks at the user interface– the sensor is incapable of differentiating among fraudulent and real biometric traits.
2. Attacks between the interfaces between two modules – a strategically placed jammer block a wireless interface and intercepts or alters the template
3. Attacks on the software infrastructure – the modification of a module during its execution stage in a way that returns the values programed by an intruder.
4. Template database attacks – templates within databases are attacked. The most common attacks include replacement and spoofing.

The main advantage of biometric templates is that they are easy to use, convenient and reliable, but despite that, there still are some risks to security and privacy (Sandhya and Prasad 2017):

1. Impersonation: a thief gains access to a person's accounts/services.
2. Sensitivity: template contains a lot of personal and/or sensitive information.
3. Linkability: Cross matching of databases needs to be prevented.
4. Loss of biometrics is permanent: Theft can render the trait useless for the user's entire lifetime.

3 Performance estimation

False Acceptance Rate (FAR) is the percentage of times the authentication systems recognized an imposter as a legitimate user. This is also known as the false match rate (FMR) (Sarkar, Singh 2020).

$$\text{FAR}(\%) = \frac{\text{Number of false accept}}{\text{Number of imposters tested}} \text{x } 100 \quad (1)$$

False Rejection Rate (FRR) is the percentage of times the authentication systems recognized a legitimate user as an imposter. This is also known as false non-match rate (FNMR) (Sarkar, Singh 2020).

$$\text{FRR}(\%) = \frac{\text{Number of rejections}}{\text{Total number of users tested}} \text{x } 100 \quad (2)$$

Equal Error Rate (EER) shows that the balance of false acceptances is related to the balance of false rejections. To achieve the best possible performance of biometric authentication systems, the EER needs to be as small as possible (Sarkar, Singh 2020).

$$\text{EER} = \text{FAR where FAR} = \text{FRR} \quad (3)$$

Genuine accept rate (GAR) is the percentage of numbers of genuine registered users being recognized by the authentication infrastructure (Sarkar, Singh 2020).

$$\text{GAR}(\%) = \frac{\text{Number of geniune user accepted}}{\text{Total number of genuine trials}} \text{x } 100 \quad (4)$$

$$\text{GAR}(\%) = 100 - \text{FRR}(\%) \quad (5)$$

4 Research methodology and a literature overview

The scientific databases Web of Science, IEEE and Scopus were used in this paper. The main keywords used in the search were "biometric face template protection". The main keywords were also used in combination with the following additional keywords: Fuzzy Commitment, Fuzzy Vault, Quantization schemes, Secure sketch, Fuzzy extractor, Salting, Non-invertible Transforms, Hybrid Methods, Homomorphic Encryption and CNN. Additional criteria were also applied, mainly the number of citations and relevance to the topic of this paper. An additional criterion was that the searched approach needed to me the only method used or in cases where there wasn't a sufficient number of papers the searched method needed to be primary method with secondary methods being used to improve it. We also only considered papers that clearly show results. The publication year of the analysed papers was not restricted. The search for papers was carried out in the period from May 1 to June 1. Results of the search together with the relevant keywords are shown and summarized within Table 1.

Table 1. Search criteria

Keywords		WoS	Scopus	IEEE
Biometric face template protection	Base	168	264	133
	J/C*	163	236	130
	Eng	163	230	130
+ Fuzzy Commitment		12	20	9
+ Fuzzy Vault		21	24	11
+ Quantization schemes		10	14	6
+ Secure sketch		8	11	6
+ Fuzzy extractor		10	12	4
+ Salting		2	3	3
+ Non-invertible Transforms		6	8	4
+ Hybrid Methods		5	7	1
+ Homomorphic Encryption		6	13	7
+ CNN		7	9	4

+ - The main keywords "Biometric face template protection" were combined with one of the additional keywords
J/C* – journal, conference

4.1 Biometric template protection

BTP is one of the most important tasks when creating a database. It can be secured at three different levels: hardware level, protocol level and software level (Kaur, Khanna 2016). At the hardware level, options like tamperproof hardware, smart cards, etc. can be used, while the protocol level options include private information retrieval, multiparty communication, and many more. This article focuses on the software level of BTP.

Biometric template security techniques must satisfy the following criteria (Sarkar, Singh 2020):

1. Diversity: The protected template should not permit comparison with other templates stored in the same database.

2. Revocability: Template protection techniques must be able to generate different templates based on the user's original biometric data and cancel old compromised templates.
3. Security: The authentic biometric template can never be obtained using the secured template.
4. Performance: Despite containing various strategies for BTP, the performance of the system should not suffer or slow down during the identification of the authenticated user.

BTP methods can be categorized into biometric cryptosystems, cancellable biometrics, hybrid methods, homomorphic encryption and others. In continuation we are presenting an overview of the methods, their accuracy and the databased they were tested on.

4.2 Biometric cryptosystems

Biometric cryptosystems are systems that bind a key to a biometric feature or generate it from a biometric feature (Sandhya, Prasad 2017). An important characteristic of biometric cryptosystems is their use of "helper data". Based on how "helper data" is derived, biometric cryptosystems can be further classified into key binding and key generating systems.

4.2.1 Key Binding Biometric Cryptosystems

In a key binding cryptosystem, the "helper data" is obtained by binding a user-specific chosen key to a biometric template (Sandhya, Prasad 2017). The combined key and biometric template from a secure template. Key binding cryptosystem can be further classified into fuzzy commitment schemes and fuzzy vault schemes.

4.2.1.1 Fuzzy Commitment Schemes

Fuzzy commitment schemes combine cryptography and error correcting codes (ECC) (Sandhya, Prasad 2017). In the enrolment stage, a random key is chosen and encoded using ECC, generating a random code word. Next, a XOR operation is carried out between the biometric feature vector and a codeword, resulting in an encrypted template.

Wang et al. (Wang et al. 2015) present a multi-biometrics template protection scheme based on fuzzy commitment and chaotic systems. The authors capture thermal facial images and generate fuzzy commitmentfrom corporation of ECC and fusion binary features. They achieve EER = $1.163 * 10^{-1}$ on NVIE face.

Elrefaei and Mohammadi (Elrefaei and Mohammadi 2019) present a fuzzy commitment scheme combined with a machine vision gait-based biometric system to enhance system security. The proposed biometric cryptosystem has two phases: enrolment and verification. They achieve FAR = 0% and FRR = 0% on CMU MoBo and CASIA A.

Gilkalaye et al. (Gilkalayeet al. 2019) present a key-binding cryptographic template security scheme based on a lattice structure and sphere packing in Euclidean space. The proposed scheme can be applied to real-value feature vectors, making it more compatible with recent face recognition methods. They achieve TPR = 0.20-0.91 and FPR = 0.07-0.0004 on LFW and VGG.

4.2.1.2 Fuzzy Vault Schemes

In fuzzy vault schemes (FVS), a key *k* is locked by an unordered set *A*, resulting in a vault *VA* (Sandhya, Prasad 2017). During the enrolment stage, a polynomial *p* encodes key *k*, *A* is projected onto *p* and chaff points are added. During the authentication stage, if another set *B* overlaps *A*, key *k* is reconstructed (Sandhya, Prasad 2017).

Wu and Yuan (Wu and Yuan 2010) propose a face-based FVS for online authentication. The transformed template and key are generated from a password and provided to the server. Fuzzy vault encoding is implemented using both the key and transformed template. They achieve FAR = 5.26%-15.38 and FRR = 23.0%-48.5% on ORL.

Nagar et al. (Nagar et al. 2011) show a feature-level fusion framework to simultaneously protect user's multiple templates as a single secure sketch. The authors achieve this by using fuzzy vault and fuzzy commitment and present a detailed analysis of the trade-off between matching accuracy and security. They achieve GAR = 75 on CASIA v1, FVC 2002 DB2 and XM2VTS.

Kaur and Sofat (Kaur and Sofat 2017) propose a multimodal biometric system security using face and fingerprint traits with fuzzy vault template security. The proposed system focuses on feature level fusion. They achieve FAR = 0 and FRR = 8.8%.

4.2.2 Key Generating Biometric Cryptosystems

Key generating cryptosystems directly generate keys through biometric templates (Sandhya, Prasad 2017). They are further classified into quantization schemes and secure sketches.

4.2.2.1 Quantization Schemes

In quantization schemes helper data is quantized to obtain stable keys. In this scheme, intervals of feature elements are obtained by taking vectors of several biometric samples (Sandhya, Prasad 2017). Helper data consists of encoded intervals. During the authentication stage, features are calculated and mapped to the determined intervals.

Li and Chang (Li and Chang 2006) apply cryptographic operations on noisy data where objects are represented in a continuous domain, and further quantified to obtain a short authentication tag. The authors use two levels of quantization and heighten the sensitivity of the proposed framework.

Han et al. (Han et al. 2008) present a novel methodology for achieving BTP using an adaptive non-

uniform quantization (ANUQ) algorithm. This is used to eliminate the contradiction between the fuzziness of the biometric information and the hash function sensitivity. They achieve FRR = 1.07-1.58 and FAR = 0.11- 2.40

Wu et al. (Wu et al. 2010) propose a biometric cryptosystem based on face biometrics. The authors extract 128-dimensional PCA feature vectors from the face and obtain a128-bit binary vector. The distinguishable bits are selected to form a bio-key. They achieve FRR = 0.0%-97.0% and FAR = 0.0%-67.5% on ORL.

4.2.2.2 Secure Sketch

Secure sketches are used to derive a consistent cryptographic key from noisy data (Chen et al. 2014). Two main components in a secure sketch scheme are the sketch generation algorithm (encoder) and the biometric template reconstruction algorithm (decoder).

Li et al. (Li et al. 2006) examine the relative entropy loss to determine optimal parameters of additional bits that could be extracted. They present a general scheme and show the relative entropy loss due to suboptimal discretization. They achieve FAR = 0.005 and FRR = 0.045 on Essex Faces94 database.

Sutcu et al. (Sutcu et al. 2007) study how secure sketch can be applied to protect the templates by identifying several practical issues and showing the subtleties in evaluating the security of practical systems on Essex Faces94 database.

Dang et al. (Dang et al. 2013) show the constructions of a face-based authentication systems where the stored templates are protected by a secure sketch. They achieve TRR = 100% on Essex Faces94 database.

4.2.2.3 Fuzzy extractor

Fuzzy extractor is a cryptographic method that produces acryptographic key directly from different biometric features (Shahreza et al. 2022).

Blanton and Aliasgari (Blanton and Aliasgari 2013) present a study about the reusability of fuzzy sketches and extractors, as well as suggestsecurity improvements. The authors present the problem of safe reuse as well showcase improvements to the overall security.

Chen et al. (Chen et al. 2014) propose an optional multi-biometric cryptosystem based on fuzzy extractor and secret share technology.

Zhang et al. (Zhang et al. 2021) utilize dense packing feature of certain lattices to design a family of fuzzy extractors that docks well with existing neural network-based biometric identification schemes. They achieve FRR = 70% and FAR = 2.1×10^{-7} on LFW.

4.2.3 Cancellable Biometrics

Cancelable biometric (CB) systems are systems that use a key-dependent transformation function (Sandhya, Prasad 2017). CB systems can be further classified into salting and non-invariable transforms.

4.2.3.1 Salting

In salting, biometric features are transformed using an invertible function (Sandhya, Prasad 2017) and the stored key can be recalled by the user for authentication. The security of this method depends on the secrecy of the transformation key and the complexity of the transformation algorithm.

Kim et al. (Kim et al. 2007) propose a two-step method for boosting the verification performance of face biometric: using an efficient feature extraction transformation, and an error minimizing template transformation. To improve the feature extraction efficiency an extended random projection of face data is used. They achieve EER = 10.915 on AR face and BERC.

Wang et al. (Wang et al. 2007) propose an approach based on discretized random orthonormal transformation of biometrics features. The authors provide properties of zero error rate and generate revocable and non-invertible biometrics templates. They achieve FAR = 0, FRR = 0 and EER = 0 on ORL and GT.

Tarek et al. (Tarek et al. 2021) create unimodal-Bio-GAN, a reliable keyless biometric salting technique based on standard generative adversarial network (GAN). The authors use a random permuted version of biometric data as a salting key. They achieve EER = 2.1% on CASIA.

4.2.3.2 Non-invertible Transforms

Non-invertible transforms (NT) apply a one-way and irreversible process to generate a transformed template from biometric data (Sandhya, Prasad 2017). Keys are produced during the authentication stage. NT can generally be further classified into geometric transforms, robust hashing, random projections, biometric filters and random permutations.

Kaur and Khanna (Kaur and Khanna 2019) propose a template protection approach for generating revocable binary features from phase and magnitude patterns of log-Gabor filters. They apply multi-level transformations at the signal and feature level to distort the biometric data using user specific tokenized variables. They achieve EER (%) = 2.40 ± 1.12 / 3.08 ± 2.11 on CASIA-Face V5 and EER (%) = 0.99 ± 0.46 / 1.19 ± 0.19 on ORL.

Sardar et al. (Sardar et al. 2020) present a novel cancelable FaceHashing technique based on non-invertible transformation with encryption and decryption template. The system consists of four components: face preprocessing, feature extraction, cancelable feature extraction followed by the classification and encryption/decryption of cancelable face feature templates. They achieve EER (%) = 0.0000 on CASIA, IITK, CVL and FERET.

Lee et al. (Lee et al. 2021) propose a data-driven cancelable biometrics scheme, named SoftmaxOut Transformation-Permutation Network (SOTPN). The SOTPN is a neural version of Random Permutation Maxout (RPM) transform, introduced for facial

template protection. EER (%) = 3.53 on LFW, YTF and FS.

4.2.4 Hybrid Methods

Hybrid methods are a combination of cancelable biometrics and cryptosystems (Sandhya, Prasad 2017). They rely on the strengths of their component schemes and provide an integrated approach with a high degree of privacy (Jegede et al. 2017).

Sree and Radha (Sree and Radha 2016) present a system for multimodal biometric authentication based on the face and fingerprints. Biometric traits are transformed using distortion algorithm. They achieve FAR = 2%, FRR = 1.8%, GAR = 98.1%

Nguyen et al. (Nguyen et al. 2019) present a hybrid biometric template protection system which takes benefits of both feature transformation and biometric cryptosystems while preventing their limitations. The performance of the system can be maintained with a new random orthonormal project technique, reducing the computational complexity while preserving the accuracy. They achieve EER = 9%

BBousnina et al. (Bousnina et al. 2021) present a hybrid system for multimodal biometric template protection to provide robustness against template database attacks. Dual-Tree Complex Wavelet Transform Discrete Cosine Transform (DTCWT-DCT) based watermarking is employed to entrench a fingerprint sketch into a face image. EER = 0%, FRR = 0.0083 %, FAR = 0.2140%, GAR = 99.99 % on ORL & FVC2002 DB1 as well as EER = 0%, FRR = 0.0080 %, FAR = 0.2141%, GAR = 99.99 % on ORL & FVC2002 DB2 and EER = 0.12%, FRR = 0.0095 %, FAR =0.2074 %, GAR = 99.99 % on ORL & FVC2000 DB1.

4.2.5 Homomorphic Encryption

Homomorphic encryption (HE) is a method characterized by allowing a limited subset of computation on the encrypted data (Sandhya, Prasad 2017). Biometric features are encrypted using a public key during the enrolment stage.

Boddeti (Boddeti 2018) proposes a fully homomorphic encryption-based framework to secure a database of face templates, designed to preserve user privacy and prevent information leakage, while at the same time maintaining their utility through template matching directly in the encrypted domain. They achieve TAR = 90.49 @ FAR = 0.01%, TAR = 96.74 @ FAR = 0.1%, TAR = 99.11 @ FAR = 1% on LFW, TAR = 23.13 @ FAR = 0.01%, TAR = 46.07 @ FAR = 0.1%, TAR = 73.71 @ FAR = 1% on IJB-A, TAR = 25.77 @ FAR = 0.01%, TAR = 48.31 @ FAR = 0.1%, TAR = 74.58 @ FAR = 1% on IJB-B and TAR = 86.48 @ FAR = 0.01%, TAR = 90.81 @ FAR = 0.1%, TAR = 93.83 @ FAR = 1% on CASIA.

Jindal et al. (Jindal et al. 2020) propose a method based on fully HE, using one-shot enrolment and supporting operations over real valued feature vectors without quantization and supporting packing of real valued feature vectors into a single cipher text. They achieve GAR = 99.33% @ FAR = 0.01% on LFW, FEI and GTF.

Kolberget et al. (Kolberget et al. 2020) present an efficiency analysis of post-quantum-secure face template protection schemes based on HE. The authors test if the approach is compliant with the ISO/IEC IS 24745 standards and their security. They achieve a false match rate less than 2% on FERET.

4.2.6. Convolutional Neural Network (CNN)

DCNN is used to learn robust mapping from face images of users to unique binary codes (bit wise randomly generated) assigned to the users (Jami et al. 2019).

Jami et al. (Jami et al. 2019) propose a method for face template protection that improves the matching performance, provides a high level of template security and addresses re-enrolment. They compute identity/class specific perturbations to the input facial feature vectors as a function of gradients of mapping network as in targeted adversarial learning. They achieve GAR = 98% @ FAR =0% on CMU-PIE, FEI and FERET.

Mai et al. (Mai et al. 2020) present a randomized CNN to generate protected face biometric templates given the input face image and a user-specific key. Through user-specific keys, the authors introduce randomness to the secure template to strengthen its security. They achieve GAR = 78.2% @ FAR =0.1% and GAR = 81.9% @ FAR =0.1% on FRGC v2.0, CFP and IJB-A.

Kim et al. (Kim et al. 2021) present the IronMask, architecture, which can be combined with any face recognition system using angular distance metric. They achieve TAR = 99.79%, FAR = 0.0005% and TAR = 95.78%, FAR = 0% on CMU-Multi-PIE, FEI and FERET.

5 Discussion

Biometric cryptosystems can be divided in to Key-binding and Key generating biometric cryptosystems.

Key-binding biometric crypto system allows the users to bind external keys with biometric data, but matching has to be done using the ECC, making it prone to key leakage. This represents a broader category and can be further divided in to Fuzzy commitment and Fuzzy Vault Schemes. Fuzzy commitment and fuzzy vault methods are among the oldest and most reliable methods that provide consistently high template security (Wang et al. 2015 - Kaur and Sofat 2017). The methods have been upgraded including approaches like Euclidean-Distance Based Fuzzy Commitment (Gilkalaye et al. 2019) and Two-phase fuzzy vault (Kaur and Sofat 2017), providing even higher levels of security. They

are still commonly used today but not as the only form of template protection, rather they are integrated together with other approaches. There is a relatively large number of papers available for research in this fields.

Key generating biometric cryptosystems can be further divided in to Quantization Schemes, Secure Sketch and Fuzzy extractors. Using key generating, biometric cryptosystem keys can be generated directly from biometric data without any external mechanism and are cancelable, but the keys are not stable. Quantization, secure sketches and fuzzy extractor methods are often used together to provide additional security and are today often hybridized (Bousnina et al. 2021). Quantization is most commonly used with PCA (Wu et al. 2010) and non-uniform quantization (Han et al. 2008). Secure sketches are able to generate extremely low FAR, FRR and EER percentages (Kim and Toh 2007 - Tarek et al. 2021), while keeping a high level of security. Fuzzy extractor is sometimes used interchangeably with secure sketches (Sandhya, Prasad 2017) and achieves high FRR results, which can be further improved in combination with quantization and secure sketches (Blanton and Aliasgari 2013 - Zhang et al. 2021). While fuzzy extractors are still being used relatively commonly in biometric template protection and a good number of papers that use this method can be found, quantization schemes and secure sketches have a slightly lower number of papers that appeared in our searches and were more commonly used as primary methods that were being upgraded with other approaches.

Cancellable Biometrics can be divided in to Salting and Non-invertible Transforms. Cancelable biometrics in general are able to achieve low FAR percentages and generate multiple templates of the same user's biometric. However, they are user-specific and invertible in case of key loss. Salting is an older approach that is still being used and updated. If used with extended random projection (Kim and Toh 2007) and random orthonormal transformation (Wang and Plataniotis 2007), amongst others, the EER is kept low with a high FRR percentage. Today, salting can be combined with Unimodal-Bio-GAN (Tarek et al. 2021) and other neural networks for even better results. A very obvious trend with salting that can be noted is the decline of papers that use this method on its own or as a primary sores.

Non-invertible transformations are able to provide a greater diversity of biometric templates and offer cancelable biometric features, but NT features are hard to generate and it is possible to reconstruct the original trait from stored templates. The authors in (Kaur and Khanna 2019) present an example of random projection with low EER values. In (Sardar et al. 2020) the authors were able to achieve a 0% EER value using hashing and keep the FAR and FRR percentages low with a high GAR percentage in (Lee et al. 2021), using random permutation. The numbers of papers that use Non-invertible transformations I slightly higher then the number of papers that use salting but it can be seen that this is also an approach in decline.

Hybrid approaches are able to increase the template security by combining multiple schemes, but with a high chance of generating dissimilar values of the same feature set. Hybrid methods are currently the most popular ones. A good example is given in (Sree and Radha 2016), where authors present a fuzzy vault hybrid (Nguyen et al. 2019), presenting a fuzzy commitment hybrid updating older and reliable approaches and making them more secure. In (Bousnina et al. 2021), authors present a secure sketch hybrid combined with new transform approaches. Hybrid approaches are the most commonly used approaches today and the number of papers proves that.

HE is regarded as the easiest approach for complying with regulations. This approach enables user collaboration, but a disadvantage of the process time is relatively slow process time. HE is one of the most popular template protection approaches. This method constantly achieves low false match rates (Kolberg et al. 2020) and high GAR and TAR values (Boddeti 2018 - Jindal et al. 2020). HE has a consistent number of papers over the years, while the number of papers is not declining it is also nor rising rapidly, but it stays stable over the years.

CNN-based methods are the newest ones used for BTP. The most common approach for CNN based methods uses Deep CNN-s (Jami et al. 2019 - Kim et al. 2021). Their development requires larger training samples and experienced users, but produces highly secure templates achieving high TAR and GAR values (Jami et al. 2019 - Kim et al. 2021). CNN based approaches offer a great variaty of protection schemes with some examples beeing approaches that use identity/class specific perturbations for input facial feature vectors (Jami et al. 2019), randomized CNN-s (Mai et al. 2020) or an adaptable architecture that can be combined with other face recognition systems (Kim et al. 2021). CNN-based methods and their number has sharply increased over the last few years. There is a solid number of papers available that use this approach.

As can be seen from this paper, the best results are achieved using CNN and hybrid methods. An overall consensus is also that using just one protection scheme does not offer sufficient protection and combining methods is the way forward. Attacks on biometric templates are getting more and more sophisticated and to keep up with that template protection schemes need to adopt and offer better protection. New methods using AI and CNN-s are being developed daily, strengthened through hybridization, by combining them with the strongest and most secure aspects of different approaches

6 Conclusion

The aim of template protection is keeping biometric data safe. As using face images as a mean of

authentication becomes more prevalent, stronger and more robust templates need to be created to protect our biometric traits. The breach of a biometric template can have devastating consequences for the end user, as we cannot simply replace our own face.

In this paper we have given a brief overview of biometric facial template protection methods. From the presented works, we can conclude that the usage of some methods like steganography and watermarking lessens with each year, and new methods based on CNN and hybrid methods are becoming more common.

One major point we came across was the terminology problem, especially in the fields of secure sketch, fuzzy extractors and non-invertible transforms. Problems arose in some works where the terms secure sketch and fuzzy extractors were used interchangeably, as well as in non-invertible transforms and their sub-categories. We can conclude that a need for a standardized and well-defined terminology exists.

The field of template protection is still growing and evolving. With more and more people using biometric traits as authentication methods, it is conceivable that biometric authentication could replace security measures like PINs and passwords in the near future. Thus, the need for stronger template protection is greater than ever.

Acknowledgments

This research was funded by the project “Development of CSTI platform for retrieval and analysis of structured and unstructured data”. The project received funding from the European Regional Development Fund through OP Competitiveness and Cohesion 2014–2020 within the Call for Proposals “Development of the products and services arising from research and development activities Phase II”, under grant number KK.01.2.1.02.0310.

References

Shahreza, H. O., Hahn, V. K., & Marcel, S. (2022). MLP-Hash: Protecting Face Templates via Hashing of Randomized Multi-Layer Perceptron. arXiv preprint arXiv:2204.11054.

Kaur, H., & Khanna, P. (2016). Biometric template protection using cancelable biometrics and visual cryptography techniques. Multimedia Tools and Applications, 75(23), 16333-16361.

Sarkar, A., & Singh, B. K. (2020). A review on performance, security and various biometric template protection schemes for biometric authentication systems. Multimedia Tools and Applications, 79(37), 27721-27776.

Sandhya, M., & Prasad, M. V. (2017). Biometric template protection: A systematic literature review of approaches and modalities. Biometric security and privacy, 323-370.

Jegede, A., Udzir, N. I., Abdullah, A., & Mahmod, R. (2017). Cancelable and hybrid biometric cryptosystems: current directions and open research issues.

Wang, N., Li, Q., El-Latif, A., Ahmed, A., Peng, J., Yan, X., & Niu, X. (2015). A novel template protection scheme for multibiometrics based on fuzzy commitment and chaotic system. Signal, Image and Video Processing, 9(1), 99-109.

Elrefaei, L. A., & Al-Mohammadi, A. M. (2019). Machine vision gait-based biometric cryptosystem using a fuzzy commitment scheme. Journal of King Saud University-Computer and Information Sciences.

Gilkalaye, B. P., Rattani, A., &Derakhshani, R. (2019, May). Euclidean-distance based fuzzy commitment scheme for biometric template security. In 2019 7th International Workshop on Biometrics and Forensics (IWBF) (pp. 1-6). IEEE.

Wu, L., & Yuan, S. (2010, September). A face based fuzzy vault scheme for secure online authentication. In 2010 Second International Symposium on Data, Privacy, and E-Commerce (pp. 45-49). IEEE

Nagar, A., Nandakumar, K., & Jain, A. K. (2011). Multibiometric cryptosystems based on feature-level fusion. IEEE transactions on information forensics and security, 7(1), 255-268.

Kaur, M., & Sofat, S. (2017, May). Fuzzy vault template protection for multimodal biometric system. In 2017 International Conference on Computing, Communication and Automation (ICCCA) (pp. 1131-1135). IEEE.

Li, Q., & Chang, E. C. (2006, September). Robust, short and sensitive authentication tags using secure sketch. In Proceedings of the 8th workshop on Multimedia and security (pp. 56-61).

Han, Q., Wang, Z., & Niu, X. (2008). An Improved Biometric Template Protection Method based on Non-Uniform Quantization. Journal of Digital Information Management, 6(2).

Wu, L., Liu, X., Yuan, S., & Xiao, P. (2010, October). A novel key generation cryptosystem based on face features. In IEEE 10th International Conference on Signal Processing Proceedings (pp. 1675-1678). IEEE.

Li, Q., Sutcu, Y., & Memon, N. (2006, December). Secure sketch for biometric templates. In International Conference on the Theory and Application of Cryptology and Information

Security (pp. 99-113). Springer, Berlin, Heidelberg.

Sutcu, Y., Li, Q., & Memon, N. (2007). Protecting biometric templates with sketch: Theory and practice. IEEE Transactions on Information Forensics and Security, 2(3), 503-512.

Dang, T. T., Truong, Q. C., & Dang, T. K. (2013, March). Practical construction of face-based authentication systems with template protection using secure sketch. In Information and Communication Technology-EurAsia Conference (pp. 121-130). Springer, Berlin, Heidelberg.

Blanton, M., & Aliasgari, M. (2013). Analysis of reusability of secure sketches and fuzzy extractors. IEEE transactions on information forensics and security, 8(9), 1433-1445.

Chen, C., Wang, C., Yang, T., Lin, D., Wang, S., & Hu, J. (2014, August). Optional multi-biometric cryptosystem based on fuzzy extractor. In 2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD) (pp. 989-994). IEEE.

Zhang, K., Cui, H., & Yu, Y. (2021). Facial Template Protection via Lattice-based Fuzzy Extractors. Cryptology ePrint Archive.

Kim, Y., & Toh, K. A. (2007, September). A method to enhance face biometric security. In 2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems (pp. 1-6). IEEE.

Wang, Y., & Plataniotis, K. N. (2007, September). Face based biometric authentication with changeable and privacy preservable templates. In 2007 Biometrics Symposium (pp. 1-6). IEEE.

Tarek, M., Hamouda, E., & El-Metwally, S. (2021). Unimodal-Bio-GAN: Keyless biometric salting scheme based on generative adversarial network. IET Biometrics, 10(6), 654-663.

Kaur, H., & Khanna, P. (2019). Random Slope method for generation of cancelable biometric features. Pattern Recognition Letters, 126, 31-40.

Sardar, A., Umer, S., Pero, C., & Nappi, M. (2020). A novel cancelableFaceHashing technique based on non-invertible transformation with encryption and decryption template. IEEE Access, 8, 105263-105277.

Lee, H., Low, C. Y., & Teoh, A. B. J. (2021, January). SoftmaxOut transformation-permutation network for facial template protection. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 7558-7565). IEEE.

Sree, S. S., & Radha, N. (2016, January). Cancellable multimodal biometric user authentication system with fuzzy vault. In 2016 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-6). IEEE.

Nguyen, T. A. T., Dang, T. K., & Nguyen, D. T. (2019, January). A new biometric template protection using random orthonormal projection and fuzzy commitment. In International Conference on Ubiquitous Information Management and Communication (pp. 723-733). Springer, Cham.

Bousnina, N., Ghouzali, S., Mikram, M., Lafkih, M., Nafea, O., Al-Razgan, M., & Abdul, W. (2021). Hybrid multi-modal biometric template protection. IntellAutom Soft Comput, 27(1), 35-51.

Boddeti, V. N. (2018, October). Secure face matching using fully homomorphic encryption. In 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS) (pp. 1-10). IEEE.

Jindal, A. K., Shaik, I., Vasudha, V., Chalamala, S. R., Rajan, M. A., & Lodha, S. (2020, December). Secure and privacy preserving method for biometric template protection using fully homomorphic encryption. In 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom) (pp. 1127-1134). IEEE.

Kolberg, J., Drozdowski, P., Gomez-Barrero, M., Rathgeb, C., & Busch, C. (2020, September). Efficiency analysis of post-quantum-secure face template protection schemes based on homomorphic encryption. In 2020 International Conference of the Biometrics Special Interest Group (BIOSIG) (pp. 1-4). IEEE.

Jami, S. K., Chalamala, S. R., & Jindal, A. K. (2019, January). Biometric template protection through adversarial learning. In 2019 IEEE International Conference on Consumer Electronics (ICCE) (pp. 1-6). IEEE.

Mai, G., Cao, K., Lan, X., & Yuen, P. C. (2020). Secureface: Face template protection. IEEE Transactions on Information Forensics and Security, 16, 262-277.

Kim, S., Jeong, Y., Kim, J., Kim, J., Lee, H. T., & Seo, J. H. (2021). IronMask: Modular architecture for protecting deep face template. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 16125-16134).

A comparative study of YOLOv5 models performance for image localization and classification

Marko Horvat, Ljudevit Jelečević, Gordan Gledec
Faculty of Electrical Engineering and Computing, Department of Applied Computing
University of Zagreb
Unska 3, HR-10000 Zagreb, Croatia
{Marko.Horvat3, Ljudevit.Jelecevic, Gordan.Gledec}@fer.hr

Abstract. *YOLOv5 is one of the latest and often used versions of a very popular deep learning neural network used for various machine learning tasks, mainly in computer vision. The YOLO algorithm has steadily gained acceptance in the data science community due to its superior performance in complex and noisy data environments, availability, and ease of use in combination with widely used programming languages such as Python. This paper aims to compare different versions of the YOLOv5 model using an everyday image dataset and to provide researchers with precise suggestions for selecting the optimal model for a given problem type. The obtained results and the implemented YOLOv5 models are available for non-commercial use at: https://github.com/mhorvat/YOLOv5-models-comparison*

Keywords. computer vision, image classification, deep learning, deep convolutional neural networks, YOLO

1 Introduction

The rapid development of Deep Learning (DL) has accelerated the progress of related methods, algorithms and procedures in the fields of image processing and computer vision, offering a wide range of applications. Nowadays, artificial neural networks are now the standard tool for most computer vision tasks. Since 2012, with the invention of AlexNet, deep neural networks have been used in thousands of remote sensing applications (Krizhevsky, Sutskever & Hinton, 2012). Among others, methods such as R-CNNs (Region Based Convolutional Neural Networks) (Girshick et al., 2014) and Fast R-CNNs (Fast Region Based Convolutional Neural Networks) (Girshick, 2015) – representing a family of machine learning models developed for computer vision and specifically object detection, as well as, Spatial Pyramid Pooling-Nets (He et al., 2015), and "You Only Look Once" (YOLO) (Redmon et al., 2016) networks provide promising results on a variety of computer vision problems.

The YOLO is currently one of the most popular DL open-source frameworks for a wide range of machine learning (ML) tasks (Jiang et al., 2022). Additionally, YOLOv4 was the best real-time object detection algorithm in 2021 based on the MAP benchmark on the standard MS COCO dataset and the fastest real-time object detection algorithm (Lin et al., 2014; Bochkovskiy, Wang & Liao, 2020). The YOLO classifiers family is principally used in computer vision for object detection and image classification, for example (Krišto, Ivasic-Kos & Pobar, 2020; Du et al., 2021), in both off-line and real-time (Redmon & Farhadi, 2018a; Fang, Wang & Ren, 2019) settings. Apart from YOLO, other frequently used algorithms for object detection are the already mentioned R-CNN and Fast R-CNN.

The YOLO algorithm is a single-stage target recognition DL algorithm. It was first proposed in 2016 by (Redmon et al., 2016), which implements object recognition classification and localization with a single neural network. It has been widely used in object recognition ever since. YOLO has undergone development from v1 to v7 (Redmon & Farhadi, 2017; Redmon & Farhadi, 2018b; Bochkovskiy, Wang & Liao, 2020; Jiang et al., 2022). Currently, YOLOv6/v7 and YOLOv5, launched in 2022 and 2020, respectively, are the latest algorithm with many practical advantages over the previous versions (Solawetz, 2020; Nelson & Solawetz, 2020; Jiang et al., 2022). Some of the most important benefits of YOLOv5 compared to previous versions are smaller volume, higher speed, higher precision, and implementation in the ecologically mature PyTorch open-source ML framework. R-CNN belongs to the two-stage object recognition algorithms, which can effectively improve the problem of the target under test in the image, but the two-stage model is more complex and computationally intensive than the single-stage model. The YOLOv5 algorithm is commonly selected for data science projects to classify free datasets available on online repositories such as Kaggle. The

original algorithm is adapted by researchers or employed directly.

The rest of the paper is organized as follows; following the introduction, section 2 explains the YOLOv5 model. The features of the architecture and the main differences between YOLOv5 and its predecessors, in particular v3 and v4, are presented. The main differences between YOLO v5 and the latest versions v6/v7 are explained in Section 2.1. In Section 3, the experiment to comparatively study the performance of YOLOv5 models is described. The standardized experimental data set is also described in this section. The results of the conducted experiment are presented and discussed in the fourth section. The fourth section also compares the YOLOv5 models using quantitative indicators. The last section concludes the paper and provides suggestions for engineers and researchers to use the YOLOv5 model for various problems DL.

2 The YOLOv5 model

"You Only Look Once" is an object detection approach that processes input images in its entirety, unlike the older "sliding window" approach where specific image segments are individually classified often in multiple passes. This approach allows for faster processing because it does not have to perform many independent evaluations and it also improves accuracy since the global image context is made available to the whole neural network (Redmon & Farhadi, 2018a; Yang et al., 2020; Jocher, 2020a; Jocher, 2020b).

The YOLOv5 (Solawetz, 2020) is based on the YOLO detection architecture and uses several algorithm optimization strategies from the field of convolutional neural networks such as auto learning bounding box anchors, mosaic data augmentation, and the cross-stage partial network. Compared to earlier solutions the YOLO model was the first object detector to connect the procedure of predicting bounding boxes with class labels in an end-to-end differentiable network (Nelson & Solawetz, 2020).

The YOLOv5 network consists of three main components: 1) backbone, 2) neck, and 3) output. First, the input terminal performs data preprocessing tasks including mosaic data augmentation (Solawetz, 2020) and adaptive image filling. To be able to adapt to different datasets, YOLOv5 integrates adaptive anchor frame calculation on the input, so that it can automatically set the initial anchor frame size when the dataset changes.

The backbone is a convolutional neural network that aggregates and forms image features at different granularities. It mainly uses a cross-stage partial network (CSP) (Kim et al., 2019) and spatial pyramid pooling (SPP) (He et al., 2015) to extract feature maps of different sizes from the input image by multiple convolution and pooling. The BottleneckCSP architecture is used to reduce the amount of calculation and increase the speed of inference, while the SPP structure realizes the feature extraction from different scales for the same feature map, and can generate three-scale feature maps, which helps improve the detection accuracy. The neck neural network represents a series of layers to mix and combine image features and to pass them forward to the prediction. In the neck network, the feature pyramid structures of FPN and PAN are used. The FPN (Liu et al., 2016) structure conveys strong semantic features from the top feature maps into the lower feature maps. At the same time, the PAN (Wang et al., 2019) structure conveys strong localization features from lower feature maps into higher feature maps. These two structures jointly strengthen the feature extracted from different network layers in Backbone fusion, which further improves the detection capability. As a final detection step, the head output is mainly used to predict targets of different sizes on feature maps.

The YOLOv5 model includes 10 individual architectures named YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x, YOLOv5n6, YOLOv5s6, YOLOv5m6, YOLOv5l6, and YOLOv5x6 + TTA (Jocher, 2020a). However, commonly only the first five are considered for research: *nano*, *small*, *medium*, *large*, *xlarge*. The main difference between them lies in the number of feature extraction modules and convolution kernels, and subsequently – which is important from the practical perspective in working with the YOLO models – in the number of neural network parameters. Accordingly, the size of parameters' set for the YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x is 4 MB, 14 MB, 41 MB, 89 MB, and as much as 166 MB, respectively (Yang et al., 2020; Jocher, 2020b). In a previous study these five YOLOv5 models were compared using COCO val2017 corpus (Lin et al., 2014; Nelson & Solawetz, 2020). In this image classification experiment researchers employed NVIDIA V100 Tensor Core GPU and the dataset consisted of 5000 pictures separated in 80 classes. The results of this study are shown in Table 1.

For evaluation of object detection, a common way to determine if one object proposal was correct is *Intersection over Union* (IoU, IU). This method takes the set $\boldsymbol{A}$ of proposed object pixels and the set of true object pixels $\boldsymbol{B}$ and calculates:

$$IoU(\boldsymbol{A},\boldsymbol{B}) = \frac{\boldsymbol{A} \cap \boldsymbol{B}}{\boldsymbol{A} \cup \boldsymbol{B}};\ IoU(\boldsymbol{A},\boldsymbol{B}) \in [0,1] \qquad (1)$$

Commonly, the threshold $IoU > 0.5$ means that object detection was a hit; otherwise, if $IoU \leq 0.5$ it was a fail. For each object class c one can calculate the *average precision* (AP) as:

$$AP(c) = \frac{TP(c)}{TP(c) + FP(c)} \qquad (2)$$

where $TP(c)$ represents the number of true positive instances and $FP(c)$ the number of false positive instances for class c. For an arbitrary class c value $AP(c) = 1$ would represent a perfect detection and $AP(c) = 0$ the worst. Therefore, the evaluation metric *mean average precision* (mAP) over the set of all objects ($\boldsymbol{O}$) in a dataset can be expressed as:

$$mAP = \frac{1}{|\boldsymbol{O}|} \sum_{c \in \boldsymbol{O}} AP(c) \quad (3)$$

Thus, metric $mAP_{0.5:0.95}$ indicates mAP over different IoU thresholds, from 0.5 to 0.95, in 0.05 increments. Similarly, $mAP_{0.5}$ describes mAP for $IoU > 0.5$. In popular object detection challenges (i.e. COCO Object Detection Challenge) additional common performance metrics is used such as $mAP_{\text{IoU=.50:.05:.95}}$, $mAP_{\text{IoU=.50}}$ and $mAP_{\text{IoU=.75}}$ (Padilla et al., 2020).

Table 1. Comparison of YOLOv5 models on COCOval 2017 dataset. Adapted from (Nelson & Solawetz, 2020).

Model	$mAP_{0.5}$	$mAP_{0.5:0.95}$	Training duration [s]
YOLOv5n	46.0	28.4	6.3
YOLOv5s	56.0	37.2	6.4
YOLOv5m	63.9	45.2	8.2
YOLOv5l	67.2	48.8	10.1
YOLOv5x	68.9	50.7	12.1

The key benefits of the YOLOv5 compared to the previous versions are 1) simple installation requiring only the PyTorch and some lightweight python libraries, 2) fast training and reduction of experimentation costs, 3) functional inference ports on individual images, batch images, video feeds, or webcam ports, 4) intuitive layout based on standard file folder layout that is intuitive and easy to navigate while developing, and 5) easy translation to mobile devices such as smartphones and tablets with Core ML compatibility. Also, from the developers' perspective, it is essential to point out that YOLOv5 is available in the PyTorch framework using Jupyter notebooks or Google Colab (Google Colaboratory) tools.

2.1 The YOLO v6 and v7 models

The latest YOLO versions MT-YOLOv6 and YOLOv7 were published online in June and July 2022 ("YOLOv6", 2022; Wang, Bochkovskiy, Liao, 2022) after this paper was accepted for review. As reported, both versions outperform YOLOv5 in object detection accuracy, training and classification speed. However, it should be noted that v6 and v7 are not part of the official YOLO series. They were developed independently by different developer groups, so YOLOv5 is still the last official YOLO release.

According to information currently in the public domain, MT-YOLOv6 (developed by the Meituan company, hence the prefix "MT") is a single-stage object detection framework dedicated to industrial applications with a hardware-friendly, efficient design and high performance. YOLOv6 has fewer model types (lacking m/l/x). Model MT-YOLOv6-s achieved 43.1 mAP on the COCO val2017 dataset, compared to 37.4 mAP for YOLOv5-s, with similar inference speed. MT-YOLOv6 shows significant improvements with small object detection in densely packed scenarios ("YOLOv6" 2022).

Even less is known about YOLOv7, except that it is said to outperform all known object detectors in terms of speed and accuracy in the range of 5-160 FPS and has the highest accuracy 56.8% AP among all known real-time object detectors with 30 FPS or higher on GPU V100 (Wang, Bochkovskiy, Liao, 2022). Of course, to thoroughly compare the latest YOLO models, further experiments with different datasets and in different environments are needed (Wang, Bochkovskiy, Liao, 2022).

3 Image classification experiment

In the experiment reported in this paper, the YOLOv5 models' performance was compared on the Face Mask Detection image dataset (Maranhão, 2020). This dataset was created for medical face mask detection, and its properties are sufficient to compare YOLOv5 models. The presented study is preliminary, and additional datasets and performance measures that are standardized in object recognition challenges, such as the COCO Object Detection Challenge, should be used for a more comprehensive analysis.

The described experiment is conceptually similar to research (Yang et al., 2020; Sharma, 2020; Ieamsaard, Charoensook & Yammen, 2021; Jiang et al., 2021) where YOLOv3 and YOLOv5 models were trained on datasets designed to train deep neural networks to detect the human face and classify whether the person in the image is wearing a mask and wearing it properly. In (Yang et al., 2020) a system was constructed to detect a human face in a picture and classify whether the person is wearing a mask. The authors considered several different deep neural networks, Faster R-CNN, R-FCN, SDD, and YOLOv5s, and compared them. When training the YOLOv5 neural network, they used the AIZOOTech dataset (Chiang, 2020) and got slightly better results when training lasted more than 300 epochs. In the second paper, the authors trained YOLOv5 models on a face detection corpus and performed classification to identify images of a person wearing a medical mask. The dataset was created by downloading images from the Internet and manually tagging them. The classification performance was slightly worse than in the first paper. Possible reasons were training on only 100 epochs, and the training data set contained a small number of examples.

3.1 Experimental dataset

The Face Mask Detection dataset used for training is available at the Kaggle repository (Maranhão, 2020). The Face Mask Detection dataset is in Public Domain (CC0: Public Domain) and it is freely available for research, education and even commercial purposes. A sample of images in the Face Mask Detection dataset is shown in Fig. 1. The Face Mask Detection dataset images are separated in 3 labelled classes: 1) persons with mask ("with_mask"); 2) persons without mask ("without_mask"); 3) persons with mask worn incorrectly ("mask_weared_incorrect").



Figure 1. Example of images in the Face Mask Detection dataset used for the comparative study.

The image bounding boxes are available in the PASCAL VOC format. The original Face Mask Detection dataset contains 853 images. However, this was insufficient for training YOLOv5 models in this research. Therefore, all images in the dataset were mirrored along their horizontal axis to produce duplicates. Some images were discarded because they were nearly identical to the originals. In the end the experimental dataset consisted of 1677 images.

3.2 Model training

All YOLOv5 architectures in this experiment were trained using the Kaggle platform and the NVIDIA Tesla P100 GPU accelerator, which is freely available online via the Google Colab platform. This environment is particularly popular for data mining and computer vision research. To objectively compare the models, a Jupyter notebook was developed using the *ultralytics* library[1] to integrate with the YOLO model. The Jupyter notebook was reused for all 5 models included in the analysis and was therefore adapted to the Google Colab and Kaggle platforms. Also, the developed Jupyter notebook implements compression of outputs in ZIP format to expediate results download. The ability to resume data processing, load and store previously trained network parameters was implemented because of various usage restrictions of online platforms.

The developed Juypter notebook is freely available for non-commercial, educational and research purposes at: https://github.com/mhorvat/YOLOv5-models-comparison. We believe that the reported results and utilized real-world experimental settings are relevant to researchers who do not have access to powerful GPU workstations, especially junior early career researchers.

The experimental dataset was randomly divided in 3 subsets for training, evaluation and testing. This approach enables cross-validation and early stopping to prevent model overfitting. The relative sizes of the subsets were 85% (training), 10% (evaluation) and 5% (testing), respectively. The data was rotated during one training session using *k*-fold cross validation ($k = 10$). The cross-validation was stratified. However, the classes were imbalanced since 3.184 images were labelled "with mask", 716 as "without mask", and only 121 "mask worn incorrectly". In the worst case this could result in the neural network labelling all samples as "with mask" to minimize loss.

Table 2. Duration of training of YOLOv5 models in hours on the experimental Face Mask Detection dataset.

Model	Training duration [h]
YOLOv5n	1.17
YOLOv5s	0.83
YOLOv5m	1.83
YOLOv5l	2.83
YOLOv5x	8.67

Each model was trained for 300 epochs. The mini groups sizes were determined with *autobatch* function in *ultralytics* library. Patience property was set to 100 epochs to interrupt the training in the event of no progress. As could be seen in Table 2, training of the YOLOv5s model was the shortest. under an hour, while training of YOLOv5x was the longest and lasted almost nine hours. Generally speaking, the duration of training increases proportionally to the number of model parameters. However, due to the enabled early stopping feature, it is possible that training of more complex models will be shorter than for simpler models, as has happened with models *n* and *s*. The Google Colab has implemented a limit on free use (i.e. the stopping feature). In this regard, all Google Colab notebooks have an idle timeout of 90 minutes and an absolute timeout of 12 hours. If the user does not

1 https://github.com/ultralytics

interact with their Google Colab notebook for more than 90 minutes, the instance will automatically stop.

4 Results and discussion

The performance of the neural network is first determined by testing with a subset of training data that can indicate whether a particular model has sufficient learning ability to be used for this purpose, and with a subset of test data to determine the generalization ability of the network and how well it performs with the unknown input data. The classification experiment was performed independently on five YOLOv5 architectures (*n*, *s*, *m*, *l*, *x*) first on the training and then on the testing image subsets. An example of models' output is shown in Fig. 2.



Figure 2. Example of image classifications as outputs from the YOLOv5x model. Outputs from other YOLOv5 architectures in this study have the same format.

The results of models' performance assessment on training and testing subsets are given in Fig. 3. The most complex architecture YOLOv5x achieved the best results in the training subset, while models m and l performed slightly worse. This is quite expected, as is the fact that YOLOv5n performed the worst due to the correlation between the number of parameters and the ability to learn the network. Results on the training and testing subsets are presented to maintain the completeness of the report.

When testing with a set of unseen data (i.e. the testing subset), simpler neural networks can produce better results than more complex networks because of the ability to retrain the more complex networks. A smaller number of parameters alone can act as a form of regularization, giving a simpler network a greater ability to generalize.

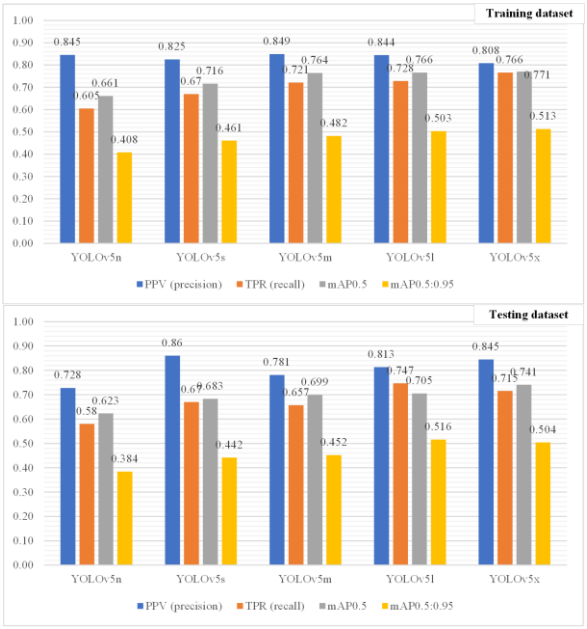


Figure 3. YOLOv5 (*n*, *s*, *m*, *l*, *x*) models' performance on the training subset (above) and testing data subset (below).

We see that more complex models are certainly more accurate, both in seen and unseen examples, but still, we cannot say that they are therefore better, especially if we take into account the everyday applications. In addition to accuracy itself, it is important, especially in the context of everyday applications, to consider the time it takes each model to process an unseen input example. The time it takes each model to process an image is given in Table 3.

A can be seen in Table 3, when choosing an optimal network architecture, it is necessary to make a trade-off between quality and speed. For more complex models to be competitive in terms of speed, they need stronger hardware, which may not always be available. For example, for a single-use device that automatically opens a door depending on whether a person is wearing a mask, it is not necessary to use a workstation when a device like the Raspberry Pi microcontroller with a camera could be used that can deliver nearly equal performance at much smaller cost.

Table 3. Duration of processing of a single random image from the experimental dataset. All values are expressed in milliseconds (ms). The shortest total processing time is in green, and the longest in red.

Model	Pre-processing	Processing	NMS	Total
YOLOv5n	0.2	5.1	2.5	7.8
YOLOv5s	0.5	5.0	2.0	7.5
YOLOv5m	0.3	11.2	1.8	13.3
YOLOv5l	0.6	17.8	2.0	20.4
YOLOv5x	0.8	32.7	1.8	35.3

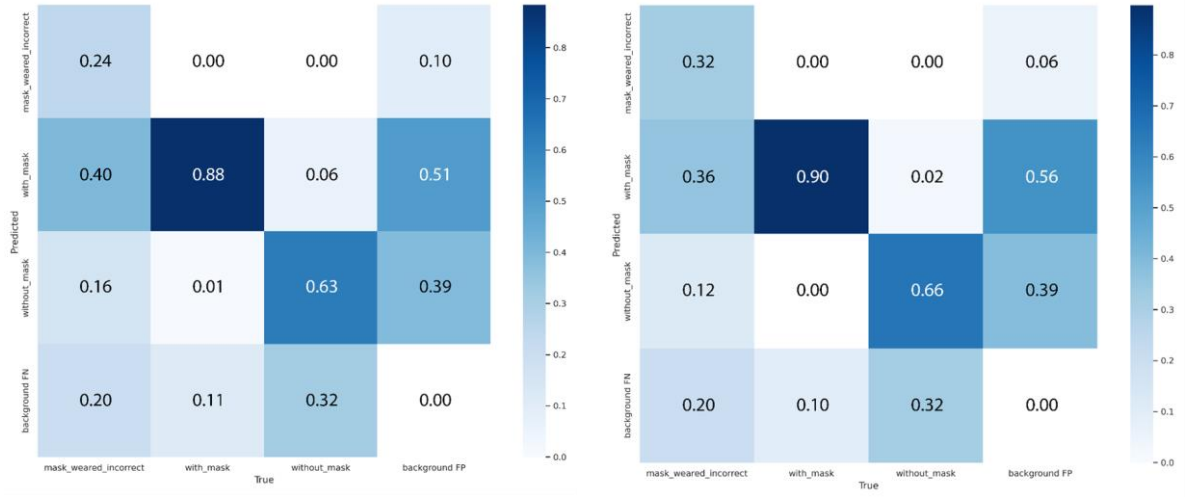


Figure 4. Confusion matrices for image classification with the most complex model YOLOv5x on the train (left) and test (right) subsets.

The YOLOv5x architecture achieved 88% TP (True Positive) for the class "with_mask" with the most examples in the training subset, and also 63% and 24% for the “without_mask” and "mask_weared_incorrect" classes with fewer data instances, respectively. In the testing, the YOLOv5x correctly classified 90% of “with_mask” instances, 66% “without_mask”, and 32% “mask_weared_incorrect” which is slightly better overall than in training. Nevertheless, such results on the testing subset may be explained by overfitting of the neural network during the training. Also, the results could be at least partially attributed to the previously mentioned problem of severely unbalanced class distribution. Class "mask_weared_incorrect" has the least images in the data set and examples belonging to this class in 52% of cases are falsely marked as "with_mask" which has the most instances in the dataset.

The results reported here and the YOLOv5 models implemented in a Jupyter notebook in the Google Colab development environment are freely available for educational and research purposes using the link provided above. Please cite this document when using the experimental results, software code, or conclusions in your work.

5 Conclusion

In the broader sense, the main advantage of YOLOv5 is twofold. First, it is the integration of advances from other research areas in computer vision theory that improve YOLOv5 object recognition performance. Second, they are made available to ML and data scientists as a collection in the PyTorch framework, which can be easily integrated and used in modern cloud-based application development environments and other software packages such as NumPy.

The latest YOLO versions v6 and v7 show even better performance than v5, especially in detecting small objects in densely packed environments. However, these versions have only recently been released and should be tested more thoroughly in various settings and with different datasets.

The YOLOv5 algorithm has significant improvements over its predecessor, YOLOv3, making it easier to train with a range of written scripts and simplifying the later application of a trained image processing network. In addition, it offers a full range of architectures of varying complexity enabling a range of new everyday applications and indicating that accuracy is not the only item to pay attention to when choosing a network. The YOLOv5 improves and facilitates the application of trained neural networks in real-time video signal processing applications such as crowd people counters (Doljanin et al., 2021; Sukkar, Kumar & Sindha, 2021), or as a part of an intelligent decision support system for monitoring student behavior that can assist teachers in optimizing personalized education during online courses (Horvat & Jagušt, 2020), or for example, in an automated system that selectively opens doors to public facilities only to persons properly wearing a medical mask.

In a nutshell, the versatile YOLOv5 algorithm can train a neural network for various practical applications, one of which is demonstrated in this paper. Moreover, unlike its predecessors, YOLOv5 improves the application of trained neural networks in video processing applications. It allows such applications to run on weaker devices, which has been challenging to do so far.

The YOLO architecture is ready to use out of the box, which is especially useful to junior engineers and researchers without much experience in machine learning and computer vision. But to improve object recognition performance without changing the algorithm, it would be advisable first to improve the

understanding of the scene. For example, classification performance could be improved by disregarding unnecessary parts of an image and selecting only genuine regions of interest before importing the image into the YOLO pipeline. In this regard, it is highly advisable to use ontologies, formal knowledge representation methods, and automated reasoning services to achieve a semantically rich description of the scene and an understanding of the concepts in the image and their mutual functional relationships (Horvat, 2013; Horvat, Grbin & Gledec, 2013a; Horvat, Grbin & Gledec, 2013b).

Finally, another important general point in practical applications of AI and DL algorithms such as the YOLO must be taken into consideration, and that is reliability or trust. Existing AI methods, if adequately trained, are very effective. Still, their output is very often hard to interpret and understand by the end-user, depending on the AI system's decisions. Therefore, in an Explainable AI (XAI) approach, an algorithm needs to provide the user with an explainable decision (Nott, 2017). Explainable AI output aims to provide the user with additional information that supports the decision made by an AI system. This is usually a set of key features that lead to the system's decision. Providing XAI explanations to the end user increases confidence and understanding of the system. To take full advantage of the YOLO family of algorithms, including future updated versions, it is vital to use them to provide a comprehensive explanation of why a particular classification or decision was made.

References

Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.

Chiang, D. (2020). AIZOOTech. *FaceMaskDetection*. Retrieved from https://github.com/AIZOOTech/FaceMaskDetection

Doljanin, D., Pranjić, L., Jelećević, L., & Horvat, M. (2021). Adaptive intelligent agent for e-learning: First report on enabling technology solutions. In *2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)* (pp. 1690-1694). IEEE.

Du, Y., Pan, N., Xu, Z., Deng, F., Shen, Y., & Kang, H. (2021). Pavement distress detection and classification based on YOLO network. *International Journal of Pavement Engineering*, *22*(13), 1659-1672.

Fang, W., Wang, L., & Ren, P. (2019). Tinier-YOLO: A real-time object detection method for constrained environments. *IEEE Access*, *8*, 1935-1944.

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).

Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, *37*(9), 1904-1916.

Horvat, M. (2013). *Generation of multimedia stimuli based on ontological affective and semantic annotation* (Doctoral dissertation, Zagreb: University of Zagreb Faculty of Electrical Engineering and Computing).

Horvat, M., Grbin, A., & Gledec, G. (2013). *WNtags: A web-based tool for image labeling and retrieval with lexical ontologies*. Frontiers in artificial intelligence and applications, 243, 585-594.

Horvat, M., Grbin, A., & Gledec, G. (2013). Labeling and retrieval of emotionally-annotated images using WordNet. *International Journal of Knowledge-based and Intelligent Engineering Systems*, *17*(2), 157-166.

Horvat, M., & Jagušt, T. (2020). Emerging Opportunities for Education in the Time of COVID-19-Adaptive e-Learning Intelligent Agent Based on Assessment of Emotion and Attention. In *Central European Conference on Information and Intelligent Systems* (pp. 203-210). Faculty of Organization and Informatics Varazdin.

Ieamsaard, J., Charoensook, S. N., & Yammen, S. (2021, March). Deep learning-based face mask detection using yolov5. In *2021 9th International Electrical Engineering Congress (iEECON)* (pp. 428-431). IEEE.

Jiang, X., Gao, T., Zhu, Z., & Zhao, Y. (2021). Real-time face mask detection method based on YOLOv3. *Electronics*, *10*(7), 837.

Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A Review of Yolo algorithm developments. *Procedia Computer Science*, *199*, 1066-1073.

Jocher, G. (2020). YOLOv5. Retrieved from https://github.com/ultralytics/yolov5

Jocher, G. (2020). YOLOv5: Train Custom Data. Retrieved from https://github.com/ultralytics/yolov5/wiki/Train-Custom-Data/

Kim, D., Park, S., Kang, D., & Paik, J. (2019, September). Improved center and scale prediction-based pedestrian detection using convolutional block. In *2019 IEEE 9th International Conference on Consumer Electronics (ICCE-Berlin)* (pp. 418-419). IEEE.

Krišto, M., Ivasic-Kos, M., & Pobar, M. (2020). Thermal object detection in difficult weather conditions using YOLO. *IEEE access*, *8*, 125459-125476.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*. .

Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ... & Zitnick, C. L. (2014, September). Microsoft coco: Common objects in context. In *European conference on computer vision* (pp. 740-755). Springer, Cham.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In *European conference on computer vision* (pp. 21-37). Springer, Cham.

Maranhão, A. (2020). Face Mask Detection. Retrieved from https://www.kaggle.com/andrewmvd/face-mask-detection

YOLOv6: a single-stage object detection framework dedicated to industrial applications. (2022). Retrieved from https://github.com/meituan/YOLOv6

Nelson, J., & Solawetz, J. (2020). Responding to the Controversy about YOLOv5. *Roboflow Blog*. Retrieved from https://blog.roboflow.com/yolov4-versus-yolov5/

Nott, G. (2017). Explainable artificial intelligence: Cracking open the black box of AI. *Computer world*, *4*.

Padilla, R., Netto, S. L., & Da Silva, E. A. (2020, July). A survey on performance metrics for object-detection algorithms. In *2020 international conference on systems, signals and image processing (IWSSIP)* (pp. 237-242). IEEE.

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).

Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7263-7271).

Redmon, J., & Farhadi, A. (2018). Yolo: Real-time object detection. *Pjreddie. com*. Retrieved from https://github.com/pjreddie/darknet/wiki/YOLO:-Real-Time-Object-Detection

Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.

Sharma, V. (2020). *Face mask detection using yolov5 for COVID-19* (Doctoral dissertation, California State University San Marcos).

Sukkar, M., Kumar, D., & Sindha, J. (2021, July). Real-Time Pedestrians Detection by YOLOv5. In *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 01-06). IEEE.

Solawetz, J. (2020). Yolov5 new version-improvements and evaluation. *Roboflow. Seach date*. Retrieved from https://blog.roboflow.com/yolov5-improvements-and-evaluation/

Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv preprint arXiv:2207.02696*.

Wang, W., Xie, E., Song, X., Zang, Y., Wang, W., Lu, T., ... & Shen, C. (2019). Efficient and accurate arbitrary-shaped text detection with pixel aggregation network. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 8440-8449).

Yang, G., Feng, W., Jin, J., Lei, Q., Li, X., Gui, G., & Wang, W. (2020, December). Face mask recognition system with YOLOV5 based on image recognition. In *2020 IEEE 6th International Conference on Computer and Communications (ICCC)* (pp. 1398-1404). IEEE.

Enhancing manual revision in manufacturing with AI-based defect hints

Jože M. Rožanec, Patrik Zajec
Jožef Stefan International Postgraduate School
Jamova cesta 39, 1000 Ljubljana, Slovenia
{joze.rozanec, patrik.zajec}@ijs.si

Jelle Keizer
Philips Consumer Lifestyle BV
Faculty of Electrical Engineering and Computer Science
Oliemolenstraat 5, Drachten, The Netherlands
jelle.keizer@philips.com

Elena Trajkova
University of Ljubljana
Faculty of Electrical Engineering
Tržaška 25, Ljubljana, Slovenija
trajkova.elena.00@gmail.com

Blaž Fortuna
Qlector d.o.o.
Rovšnikova 7, Ljubljana, Slovenija
blaz.fortuna@qlector.com

Bor Brecelj, Beno Šircelj, Dunja Mladenić
Jožef Stefan Institute
Jamova cesta 39, 1000 Ljubljana, Slovenia
bor.brecelj@gmail.com, {beno.sircelj, dunja.mladenic}@ijs.si

Abstract. *Quality control allows companies to verify whether the products conform to requirements and specifications. However, while Artificial Intelligence is increasingly used to automate the visual inspection process, a manual revision can be required when the model cannot determine whether a piece is defective or not with enough confidence. Therefore, means must be devised to optimize the manual revision of such products, to increase the speed and quality of labeling. In this paper, we perform experiments to determine whether different defect hinting techniques and data imbalance mitigation techniques can enhance the manual revision process. Furthermore, we contrast the performance of two groups of persons with different skills and education levels and their perceptions when executing the experiments. We performed the experiments on real-world data provided by Philips Consumer Lifestyle BV.*

Keywords. Intelligent Manufacturing Systems; Artificial Intelligence; Quality Assurance and Maintenance; Fault Detection; Human Centred Automation

1 Introduction

Quality control is one of the critical activities performed in the manufacturing industry to detect product defects to ensure quality standards are met, avoid rework, and supply chain disruptions, and avoid potential damage to the brands' reputation (Wuest et al. (2014); Yang et al. (2020)). In addition, information regarding the defective products enables precise and timely root cause detection and therefore considers mitigation actions to improve the manufacturing processes and overall product quality.

The advent of the Industry 4.0 paradigm has fostered the use and integration of information and communication technologies to increase the flexibility and efficiency of the manufacturing process while also leading to greater value over the whole product lifecycle (Frank et al. (2019)). The decreased cost of sensors has accelerated the digitalization of manufacturing (Benbarrad et al. (2021)), enabling the use of artificial intelligence for defect detection in industrial settings (Carvajal Soto et al. (2019); Chouchene et al. (2020)). Automated visual inspection provides greater scalability (does not require training and is not subject to fatigue, absenteeism, or inspection inefficiencies as humans are) and ensures all products are inspected following the same criteria. Given the abovementioned benefits, machine learning automated visual inspection has been applied in many scenarios (Gobert et al. (2018); Iglesias et al. (2018)). While supervised classification is frequently used to discriminate between known defects, unsupervised machine learning approaches enable discovering non-labeled defects or performing defect detection where such labeling is not feasible.

Regardless of the machine learning approach used to perform the automated visual inspection, there are usually certain products for which the machine learning model cannot determine with a high level of certainty whether they are defective or not. In such cases, manual inspection is required to perform a final evaluation of the product's quality. Therefore, we consider

it essential to develop means to enhance the manual revision process to achieve an increased speed and quality of labeling. This research compares balanced and imbalanced data streams with defect hinting strategies to understand which yield better outputs and how the users perceive them.

The main contribution of this research is the assessment of multiple scenarios regarding manual revision involving balanced and imbalanced data streams and different defect hinting strategies. The scenarios are assessed based on data collected during the experiments to determine the time and quality of labeling and the participants' perceptions evaluated through a set of surveys. The experiments were performed considering a real-world use case from the *Philips Consumer Lifestyle BV* corporation.

We evaluate the labeling quality through the precision, recall, and F1 metrics, considering it a binary classification problem (whether a defect exists or not, regardless the defect). Furthermore, we compare the mean and median labeling times and compute the labeling time standard deviation. Finally, we quantify the number of unidentified defects. To understand the participants' perception, we summarize the survey results and compare the outcomes obtained for two different groups: (a) students and researchers of an artificial intelligence laboratory and (b) *Philips Consumer Lifestyle BV* operators.

The rest of this paper is structured as follows: Section 2 presents related work, and the Section 3 describes the *Philips Consumer Lifestyle BV* use case. Section 4 introduces the methodology we followed in executing the experiments later. Section 5 describes the experiments we performed, while Section 6 describes and analyzes the results we obtained. Finally, Section 7 offers our conclusions and provides an outline for future work.

2 Related Work

Automated visual inspection is implemented with image processing techniques. Among them, machine learning models had been widely adopted, and state-of-the-art (SOTA) performance was achieved with deep learning models (Beltran-Gonzalez et al. (2020); Pouyanfar et al. (2018)). While such models achieve great precision, when they cannot classify a product with high confidence, the product must be manually inspected or thrown away. EXplainable Artificial Intelligence (XAI) can provide additional insights on such cases to understand gaps in the model's knowledge and improve it over time. Nevertheless, little research was devoted to the problem of manual revision when such products are manually inspected. Furthermore, there is a research void on how cues provided by XAI techniques or unsupervised machine learning models can assist the operators in manual revision.

The heatmap is a data visualization technique that aims to characterize a phenomenon through a color scale extended over two dimensions. As such, it has been widely adopted by XAI techniques related to image classification (Samek and Müller (2019)), and unsupervised machine learning models (Defard et al. (2021); Zavrtanik et al. (2021)), to convey information regarding what is being considered by the algorithm to make a prediction or where does the model expect the defect to be, frequently overlaying the heatmap with a certain degree of transparency to the original image.

3 Use Case

The research was performed based on a real-world use case concerning the visual inspection of manufactured products produced by *Philips Consumer Lifestyle BV* in Drachten, The Netherlands. The manufacturing plant is considered one of the largest Philips development and production centers in Europe. Printing machine setups for various products and logos exist, and many products produced on these machines are manually handled and inspected to determine their visual quality. If some defect is observed, they are removed from the manufacturing line, leaving only those manufactured products that comply with the existing quality standards. Operators spend several seconds handling, inspecting, and labeling the products. While such a process could be automated to a certain extent, there are certain cases whose quality cannot be automatically determined and therefore require manual inspection. Therefore, we explore means to increase the operators' labeling speed and labeling quality through artificial intelligence. We do so through the creation of synthetic images to achieve class balance and by providing defect hints using heatmaps obtained either from XAI techniques (GradCAM (Selvaraju et al. (2017))) or unsupervised defect detection models (DRAEM (Zavrtanik et al. (2021))), or showing the nearest labeled images. The research was performed with a labeled dataset of 3.518 images, with three possible categories (see Fig. 1): good printing, double printing, and interrupted printing.



Figure 1: The images considered for this research correspond to well-printed logos (good) or defective prints (double or interrupted prints).

4 Methodology

For this research, we developed a web application that enabled the participants to log in with an assigned user-

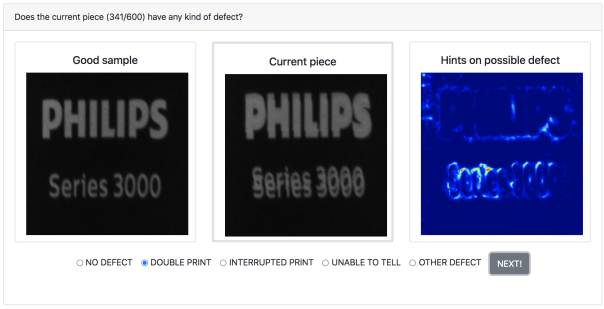


Figure 2: Screenshot of an application we developed to test multiple labeling scenarios. We always display a *good* image on the left side. The image to be inspected is shown in the center. On the right, we eventually provide some defect hinting. In the sample screenshot, we find an image of a *double-print* defect and the hint image obtained from a DRAEM anomaly map.

name and password and sequentially execute a series of experiments. When a given experiment was completed, the next one was unlocked to ensure all participants executed the experiments in a particular order. While the information collected during the experiments was associated with the username of the corresponding participant, no means were provided to associate the usernames back to a particular person. For each experiment, we collected the experiments' start and end time, the time required to label each image, and the labels the participants assigned to each image. For each image, we had prior information regarding the ground-truth label and whether the image was obtained from the original dataset provided by *Philips* or if it was synthetically generated.

We asked the participants to execute one experiment daily to ensure the labeling fatigue would not affect their performance. In addition, we asked them to fill out two kinds of surveys: (a) surveys to understand their perception regarding a particular experiment setup (Experiments 3-6), and (b) an overall survey to obtain basic demographic data, evaluate the overall experience, and compare the perceived usefulness of the setups suggested by each experiment, once all of them were completed. In the first survey, we asked (i) if the defect hinting component is useful, (ii) if the defect hinting component is understandable, (iii) whether it helped to gain new insights into the defect identification process, (iv) it improved the decision-making process, (v) and if the participant would use the defect hinting component in their working environment. For each question, an answer could be provided on a Likert scale between one (strongly disagree) and four (strongly agree). In the second survey, we asked for demographic data, technical background information, and overall feedback regarding the experiments performed. The demographic data we collected were the participants' gender, age range, and education level. Regarding the technical background information, we asked them whether they (a) have basic skills related to manipulation of standard software applications, (b) did some programming in the past but not regularly (on a monthly basis), (c) regularly work on programming/software development, (d) can identify, locate, retrieve, store, organize and analyze digital information and evaluate relevance and purpose, (e) can solve digital problems and explore new ways to take advantage of technology, (f) are aware of basic principles in the area of Machine learning and Artificial Intelligence, and (g) can solve problems using Machine learning and Artificial Intelligence techniques. Finally, we asked them to evaluate the defect hinting components from different experiments, considering whether the particular experiment setups are (a) useful, (b) understandable, (c) informative, (d) help them to be more efficient and (e) learn fast. The last survey question was devoted to understanding if they would adopt a specific setup in their working environment.

We considered two groups of participants: (A) four researchers and students from an artificial intelligence laboratory and (B) three operators tasked with visual quality inspection at the *Philips* manufacturing plant. The goal of having two distinct groups of participants was to understand whether the different backgrounds and domain knowledge influence the task execution and perception of defect hinting. We provide their demographic and technical literacy details in Table 1. All the participants signed consent forms regarding the experiments performed.

Participant	Demographic data			Technical literacy						
	Gender	Age	Education	a	b	c	d	e	f	g
A1	Female	18-30	University Degree							x
A2	Male	18-30	University Degree	x	x	x	x	x	x	x
A3	Male	18-30	Post-graduate degree	x		x	x	x	x	x
A4	Female	31-45	Post-graduate degree			x	x	x	x	x
B1	Female	45-60	Higher education							
B2	Female	45-60	Higher education							
B3	Male	31-45	Higher education				x			

Table 1: Demographic and educational information regarding the participants. Participants' ID is prefixed with the letter of their corresponding group (A for researchers and students of the artificial intelligence laboratory, and B for *Philips* operators.

5 Experiments

In this research, we were interested in exploring means of increasing the velocity and quality of the labeling process. To that end, we performed six experiments, analyzing how the class imbalance and inclusion of a defect hinting component affect the manual revision process. We detail them below:

- **Experiment 1**: stream of images with raw data imbalance without defect hinting. The experiment's goal was to provide a baseline regarding the current manual visual inspection setup.

- **Experiment 2**: balanced stream of images, mitigating the class imbalance with GAN-generated images of defective products (identical images were generated) and no defect hinting.
- **Experiment 3**: stream of images with natural data imbalance and DRAEM anomaly heatmap for defect hinting.
- **Experiment 4**: balanced stream of images (same as Experiment 2) with DRAEM anomaly heatmap for defect hinting.
- **Experiment 5**: balanced stream of images (same as Experiment 2) with a GradCAM heatmap for defect hinting.
- **Experiment 6**: balanced stream of images (same as Experiment 2) where the nearest labeled image (considering the structural similarity index measure) and its label were provided for defect hinting.

Above we describe six experiments. In Experiments 1-4 we aim to determine whether the class imbalance affects the quality of labeling, regardless the existence of some defect hinting. We considered that a higher labeling quality among balanced streams could signal enhanced participants' attention. On the other hand, in Experiments 4-6 we remove the class imbalance and aim to determine which defect hinting technique issues the best results in terms of time and labeling quality. We avoided experiments contrasting balanced and imbalanced settings across all of the defect hinting techniques, given the difficulty to gather participants to collaborate on them, and the additional time required to successfully complete all of them.

The heatmaps are data visualizations that aim to provide an overview of information of interest on a 2D scale, using a color scale to direct the users' attention towards the most interesting regions. In particular, the DRAEM anomaly heatmap highlights the regions that do not conform to what is considered a good sample and therefore can be potentially considered a defect. On the other hand, the GradCAM heatmap plots the values obtained from the gradient of the output of a particular layer, to highlight which image regions are being considered by the model for the task at hand. Ideally, a GradCAM heatmap should highlight image regions provide information regarding whether a defect is present at the image or not.

6 Results and Analysis

The results we obtained from the experiments are twofold. First, the data collected in the web application was analyzed to understand whether differences in labeling quality and times have given different experiment setups. Second, the responses obtained from the surveys completed by the participants were analyzed to get a better understanding of their perception and experience when labeling the data.

6.1 Application measurement results

Experiment	Labeling quality			Labeling time (s)			Unidentified defects	
	Precision↑	Recall↑	F1↑	Median↓	Mean↓	Std. Dev.↓	Other↓	Unable to tell↓
1	0,2857	0,3576	0,3090	*3*	4,27	5,85	36,00	34,33
2	*0,9166*	*0,6304*	*0,7444*	**2**	3,08	5,67	*9,75*	21,00
3	0,6774	0,4983	0,5660	**2**	**2,67**	*5,38*	16,67	11,50
4	0,7339	0,4691	0,5693	**2**	3,00	5,71	10,67	8,00
5	0,8238	0,5212	0,6374	**2**	*2,83*	**5,17**	13,00	*7,75*
6	**0,9393**	**0,8360**	**0,8831**	**2**	3,22	5,82	**1,50**	**5,25**

Table 2: The table summarizes the measurements obtained for the researchers and students of an artificial intelligence laboratory regarding labeling quality and time, along with the number of cases where the user could not tell whether there was a defect or considered the defect did not correspond to one of the existing classes. The best results are bolded, and the second-best results are highlighted in italics.

Experiment	Labeling quality			Labeling time (s)			Unidentified defects	
	Precision↑	Recall↑	F1↑	Median↓	Mean↓	Std. Dev.↓	Other↓	Unable to tell↓
1	0,2468	0,7361	0,3691	**2**	*4,67*	11,34	35,00	**1,00**
2	*0,7915*	*0,8196*	*0,8014*	**2**	5,23	**9,42**	56,33	*2,67*
3	0,4156	0,7845	0,5159	**2**	5,45	12,78	**14,33**	**1,00**
4	0,7103	0,7056	0,7042	*3*	8,67	14,81	109,00	**1,00**
5	0,7721	0,7132	0,7312	**2**	**4,16**	*9,78*	42,33	**1,00**
6	**0,8170**	**0,8598**	**0,8303**	**2**	6,14	13,85	*18,67*	**1,00**

Table 3: The table summarizes the measurements obtained for the *Philips* operators regarding labeling quality and time, along with the number of cases where the user could not tell whether there was a defect or considered the defect did not correspond to one of the existing classes. The best results are bolded, and the second-best results are highlighted in italics.

We summarize the results obtained from the application measurements in Table 2 and Table 3. From the results we observed that the *Philips* operators consistently achieved a high recall, while the recall measured for researchers was low for all of the experiments, except for Experiment 6. We consider the most probable reason for such responses is a different understanding of the quality process. While the researchers are not concerned for the mislabeling a defect, the *Philips* operators avoid a true defect going outside of the manufacturing process. Therefore, when uncertain, they most likely label a manufactured piece as defective and sacrifice good products to avoid leaking defective ones. While the labeling quality differed between both groups, we observed that for both, the best performance was achieved for Experiment 6 and the second-best for Experiment 2 for Precision, Recall, and F1. Given that both experiments correspond to a balanced data stream, we consider the balanced stream increased the participants' attention and that the defect hinting from Experiment 6 was beneficial towards achieving better labeling quality. Experiment 6 also had the lowest or second-lowest number of unidentified defects or cases where the participants could not tell whether a defect was present. Finally, both groups showed the best and

second-best mean labeling speed and low standard deviation in Experiment 5, which had defect hints created with the GradCAM XAI technique. Nevertheless, the median remained the same for most of the experiments, and the quality of labeling was notably lower than the one obtained for Experiment 6.

6.2 Survey results

Exp.	Group	Useful?	Understandable?	New insights?	Improves decision-making?	Would try in working environment?
3	A	**2**	**2**	**2**	**2**	**2**
	B	2	2	1	2	1
4	A	**2**	**2**	**3**	**2**	**2**
	B	1	2	1	1	1
5	A	**3**	**3**	**2**	**2**	**3**
	B	1	2	1	1	2
6	A	**4**	**4**	**4**	**4**	**4**
	B	1	2	2	2	2

Table 4: The table shows the summarized evaluation outcomes obtained from the surveys taken at the end of each experiment (Exp.) for both groups of participants. The results correspond to a Likert scale between one (strongly disagree) and four (strongly agree). With no exception, group A (participants from a laboratory of artificial intelligence) ranked the application higher than group B (*Philips* operators).

Among the survey results, we first analyzed the surveys targeting the participants' perceptions regarding Experiments 3-6 (see Table 4). We found that the perceptions of both groups were considerably different. While the operators mostly responded with values from the lower end of the scale, the researchers responded with higher rates on the Likert scale. The researchers considered the best defect hinting strategy was to provide the image and label of the closest labeled image (Experiment 6). Therefore, the researchers assigned the highest Likert scale value and enforced their decision in the final evaluation survey. They responded that they would prefer the Experiment 6 setup in a working environment. On the other hand, while the highest-ranking experiment among the operators was Experiment 6, their maximum rating was two on the Likert scale of four, tending towards disapproval. Furthermore, when asked which experiment set up they would prefer in a working configuration, they unanimously responded for Experiment 1 or No experiment, which are equivalent responses: to avoid introducing changes to the current working flow.

When comparing DRAEM defect hinting under balanced and imbalanced streams of data and the GradCAM explanations, we observed conflicting responses and a consensus that there is no clear value on such kinds of defect hints for the labelers. Nevertheless, the picture was different when we asked the participants to evaluate all the experiments once all experiments had concluded. The researchers from the artificial intelligence laboratory considered that defect hinting was useful, informative, helped to be more efficient, and learn fast if the defect hints were created based on DRAEM heatmaps or nearest labeled images (see Table 5). The responses provided by the operators were not informative (see Table 6): while a participant considered all of the experiments were informative, another one considered all experiments were understandable. A third participant provided an ambiguous response, which was excluded from the results to avoid interpretive bias.

Exp.	Useful?	Understandable?	Informative?	Helps to be efficient?	Helps to learn fast?
1	0,50	0,50	0,25	0,50	0,25
2	0,25	0,75	0,25	0,50	0,25
3	0,25	0,75	0,50	0,50	0,50
4	0,25	0,75	0,50	0,25	0,25
5	0,25	0,75	0,50	0,25	0,25
6	1,00	1,00	1,00	0,75	1,00

Table 5: The table presents the outcomes of the overall evaluation of the experiments (Exp. 1 through Exp. 6) by the participants from an artificial intelligence laboratory. Responses were provided as a Boolean value, and the table summarizes the proportion of participants answering positively to the questions presented above. The evaluation was based on a survey once all the experiments were completed.

Exp.	Useful?	Understandable?	Informative?	Helps to be efficient?	Helps to learn fast?
1	0,00	0,33	0,33	0,00	0,00
2	0,00	0,33	0,33	0,00	0,00
3	0,00	0,33	0,33	0,00	0,00
4	0,00	0,33	0,33	0,00	0,00
5	0,00	0,33	0,33	0,00	0,00
6	0,00	0,33	0,33	0,00	0,00

Table 6: The table presents the outcomes of the overall evaluation of the experiments (Exp. 1 through Exp. 6) by the participants from *Philips*. Responses were provided as a Boolean value, and the table summarizes the proportion of participants answering positively to the questions presented above. The evaluation was based on a survey once all the experiments were completed.

6.3 Perceptions vs. data

From the results presented in the subsections above, we found that the group of students and researchers in the artificial intelligence laboratory perceived the usefulness of the approach we developed to assist in the manual inspection and that the best-perceived setup matches the best-performing experiment (Experiment 6). However, on the other side, the experiments were not perceived as helpful by the *Philips* operators. Nevertheless, the data gathered from the application shows that while the median labeling time remained the same as in their current setup, and the mean labeling time was 30% higher, the quality of labeling was notably superior, resulting in more than three times greater precision, and more than two times the original F1 score. Furthermore, Experiment 6 helped the *Philips* operators reduce the number of unidentified defects by more than 80%. Therefore, we conclude that the setup used for Experiment 6 has shown benefits regardless of the participants' perceptions. Furthermore, further work needs to be done to determine how such changes could be communicated to the operators to change their perceptions.

7 Conclusion

This research presents a set of experiments designed to understand which machine learning technologies and approaches could be leveraged to enhance the manual revision process for those products whose quality cannot be determined with certainty with a machine learning model. To that end, we conducted six experiments with six hundred images each, based on a real-world dataset provided by *Philips Consumer Lifestyle BV* and evaluated two different groups of participants: (a) researchers and students from an artificial intelligence laboratory, and (b) *Philips* operators tasked with the visual inspection of products. We then compared data collected from the application during the experiments' execution with surveys. We analyzed which experiments yielded the best performance and whether the perception of each group matched the objective measures. We found the best performance was achieved with a balanced stream of images (an equal amount of images considering good and defective products) and hinting the user on the expected class by showing the most similar labeled image and the corresponding class. Gains in labeling precision were at least three times higher than in the existing setup, while the median labeling time remained the same or slightly decreased. Furthermore, the number of unidentified defects was reduced by more than 80% in the worst case. Future work will explore new setups that could enhance the manual revision process by introducing slight imbalances that could reduce the synthetic images overhead and monitor the operators' wellbeing, ensuring tasks are changed, or different strategies adopted when their attention decreases.

Acknowledgments

This work was supported by the Slovenian Research Agency and the European Union's Horizon 2020 program project STAR under grant agreement number H2020-956573.

The authors acknowledge the valuable input and help of Paulien Dam and Yvo van Vegten from *Philips Consumer Lifestyle BV*.

References

Beltran-Gonzalez, C., Bustreo, M., and Del Bue, A. (2020). External and internal quality inspection of aerospace components. In *2020 IEEE 7th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, pages 351–355. IEEE.

Benbarrad, T., Salhaoui, M., Kenitar, S. B., and Arioua, M. (2021). Intelligent machine vision model for defective product inspection based on machine learning. *Journal of Sensor and Actuator Networks*, 10(1):7.

Carvajal Soto, J., Tavakolizadeh, F., and Gyulai, D. (2019). An online machine learning framework for early detection of product failures in an industry 4.0 context. *International Journal of Computer Integrated Manufacturing*, 32(4-5):452–465.

Chouchene, A., Carvalho, A., Lima, T. M., Charrua-Santos, F., Osorio, G. J., and Barhoumi, W. (2020). Artificial intelligence for product quality inspection toward smart industries: quality control of vehicle non-conformities. In *2020 9th international conference on industrial technology and management (ICITM)*, pages 127–131. IEEE.

Defard, T., Setkov, A., Loesch, A., and Audigier, R. (2021). Padim: a patch distribution modeling framework for anomaly detection and localization. In *International Conference on Pattern Recognition*, pages 475–489. Springer.

Frank, A. G., Dalenogare, L. S., and Ayala, N. F. (2019). Industry 4.0 technologies: Implementation patterns in manufacturing companies. *International Journal of Production Economics*, 210:15–26.

Gobert, C., Reutzel, E. W., Petrich, J., Nassar, A. R., and Phoha, S. (2018). Application of supervised machine learning for defect detection during metallic powder bed fusion additive manufacturing using high resolution imaging. *Additive Manufacturing*, 21:517–528.

Iglesias, C., Martinez, J., and Taboada, J. (2018). Automated vision system for quality inspection of slate slabs. *Computers in Industry*, 99:119–129.

Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M. P., Shyu, M.-L., Chen, S.-C., and Iyengar, S. S. (2018). A survey on deep learning: Algorithms, techniques, and applications. *ACM Computing Surveys (CSUR)*, 51(5):1–36.

Samek, W. and Müller, K.-R. (2019). Towards explainable artificial intelligence. In *Explainable AI: interpreting, explaining and visualizing deep learning*, pages 5–22. Springer.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*, pages 618–626.

Wuest, T., Irgens, C., and Thoben, K.-D. (2014). An approach to monitoring quality in manufacturing using supervised machine learning on product state data. *Journal of Intelligent Manufacturing*, 25(5):1167–1180.

Yang, J., Li, S., Wang, Z., Dong, H., Wang, J., and Tang, S. (2020). Using deep learning to detect

defects in manufacturing: a comprehensive survey and current challenges. *Materials*, 13(24):5755.

Zavrtanik, V., Kristan, M., and Skočaj, D. (2021). Draem-a discriminatively trained reconstruction embedding for surface anomaly detection. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 8330–8339.

Bicycle Sharing Systems meet AI: forecasting bicycles availability and decision-making.

Jože M. Rožanec, Tadej Krivec
Jožef Stefan International Postgraduate School
Jamova cesta 39, 1000 Ljubljana, Slovenia
{joze.rozanec, tadej.krivec}@ijs.si

Vid Keršič
University of Maribor
Faculty of Electrical Engineering and Computer Science
UM FERI, Koroška cesta 46, 2000 Maribor, Slovenia
vid.kersic@um.si

Larsen Cundrič
Longevize BV
l.cundric@gmail.com

Blaž Stojanovič
blaz.sto@gmail.com

Marko Zeman, Ivan Bratko
University of Ljubljana
Faculty of Computer and Information Science
Večna pot 113, 1000 Ljubljana, Slovenia
{marko.zeman, bratko}@fri.uni-lj.si

Abstract. *With the ubiquitous increase in the number of people in cities, there is a growing need for sustainable transport possibilities. Smart cities should provide environment-friendly ways to travel inside the city. One of the most nature-preserving ways to travel is using bicycles, which is often encouraged by public bicycle sharing systems, which are present in many cities around the globe. While these systems are usually easily accessible, they still lack optimization regarding bicycle availability across stations. This paper contributes towards the optimization of the bicycle sharing system in Ljubljana, Slovenia. We developed classification and regression machine learning models for predicting the emptiness and occupancy across bicycle stations in near future. These predictions allow for the caretakers of the system to intervene on time and provide enough bicycles across all stations.*

Keywords. Smart city, sustainable transport, bicycle sharing, optimization

1 Introduction

The world is facing an environmental crisis, where the ecosystems are changing rapidly. In 2018, the intergovernmental body reported that we are facing irreversible consequences due to the global greenhouse gas (GHG) emissions (Masson-Delmotte et al. (2018)). For that reason, the goal is to address climate change by preparing for unavoidable consequences or reducing emissions. One of the areas where the benefits of mitigating the GHG emissions would be significant is transport.

Transportation systems account for a quarter of energy-related CO_2 emissions (Masson-Delmotte et al. (2018)). One of the strategies to reduce GHG emissions in transport is a modal shift to lower-carbon options (Rolnick et al. (2022)). One of the best ways to substitute the carbon emission intensive modes of transport in densely populated areas is with bicycles.

The first bicycle-sharing system (BSS) was introduced in 1965 in Amsterdam with the *White Bike Plan* bicycle-sharing program. Since then, Shaheen et al. (2010) reported that more than 130.000 bicycles have been made available through such programs worldwide. In all such programs, the users can borrow a bicycle from one of the stations and return it to some station within a specific time frame.

Among the biggest challenges of the BSS is the (re)balancing of the bicycles at the parking stations (Shaheen et al. (2010)). For example, in the morning and afternoon hours, the direction of travel by bicycle is usually heavily biased in one direction. This bias, in turn, causes an imbalance at the stations, where some stations are full and others are empty. The employed workers then redistribute bicycles from occupied stations to less occupied ones. However, the open question remains of how and when the employees should take action. Machine learning (ML) provides means to solve such a problem.

ML algorithms have already achieved great success in various forecasting tasks. There is a big motivation to use ML for mitigating climate change. For example, ML is used to forecast electricity demand, photovoltaic production, extreme events, and behavior change (Rolnick et al. (2022)). For the task at hand, ML algorithms can assist decision-making in the bicycle redistribution problem.

In this paper, we present an approach to classifying decision-making options (adding/removing bicycles from stations) and an approach to forecasting the number of bicycles at the parking stations. We do so considering multiple forecasting horizons. Our contributions involve developing and comparing classification models for decision-making options and regres-

sion models for bicycle availability, which can be used in real-time by the BicikeLJ BSS to assist the rebalancing process.

The rest of this paper is structured as follows: Section 2 presents related work, and Section 3 describes the use-case of a BSS used in Ljubljana, Slovenia. Section 4 introduces the methodology we followed to analyze the data, create the dataset, the corresponding features, and how we trained and evaluated the forecasting models. Section 5 presents the results, while Section 6 offers our conclusions and provides an outline for future work.

2 Related Work

Shaheen et al. (2010) provide an overview of BSS and related problems and experiences. Bicycle imbalance in a network of bicycle stations is a critical problem. Therefore bicycle redistribution is a necessary functionality of such systems. Usually, it is done by trucks or other vehicles to move bicycles to high-demand stations. Efficient maintenance of bicycle balance requires online information support, automated prediction of the number of bicycles and demand for a bicycle at individual stations, and planning optimal interventions to restore the balance.

Many works (Yang et al. (2016); Almannaa et al. (2020); Ashqar et al. (2017); Cagliero et al. (2017); Ashqar et al. (2021)) report on experiments with the automated prediction of the number of bicycles at different stations, using various statistical and ML techniques for learning to predict from past BSS data. Yang et al. (2016) proposed a prediction method that consists of (1) predicting the check-in numbers at individual stations, given the current data about check-out bicycles and their locations, taking into account the probabilities of bicycles currently traveling between pairs of stations and travel times; and (2) estimating the probability distribution of checking out a bicycle at given station depending on the time of day, kind of day, type of station location, weather conditions, among other factors. Using historical data, this is done by the Random Forest ML method (Breiman (2001)). They conducted an experimental evaluation of their approach regarding prediction accuracy using past data observed in the BSS of the Chinese city of Hangzhou, supposedly the world-largest BSS. They compared Random Forest results with three more traditional techniques for time-series prediction and found that Random Forest outperformed these baseline methods.

Almannaa et al. (2020); Ashqar et al. (2017, 2021) report on experiments in the prediction of the number of bicycles at bicycle stations using historical data in the BSS of the San Francisco Bay area. Prediction methods used in Almannaa et al. (2020) were dynamic linear models and Random Forest. Both methods produced similar results in terms of prediction accuracy. For the fifteen-minute predictions horizon of bicycle counts at a station, the average prediction error was about 2% of the station's total bicycle capacity, and the average error for the two-hour prediction horizon was 6%. Cagliero et al. (2017) developed a method for predicting critical conditions in a BSS regarding bicycle counts, integrating the Bayesian learning and Association learning techniques. The method was evaluated on historical bicycle numbers data for the New York City BSS.

The above papers are mainly concerned with predicting bicycle counts but not explicitly with planning optimal bicycle redistribution interventions. Seo et al. (2022) formulated the bicycle redistribution problem as a Markov decision problem. Then the choice of best actions (bicycle moving operations) according to a suitable cost function in a stochastic dynamic environment can be addressed by techniques of reinforcement learning (Sutton and Barto (2018)). The proposed solution suggests the best action every ten minutes, which depends on the observed current state of the BSS system and predicted future demands. Predictions are made by the Random Forest method. This approach was experimentally evaluated on the BSS of Yeouido Island District in South Korea (thirty-one bicycle stations). The experiment used historical data covering about one year from this BSS. The proposed solution to bicycle rebalancing has a clear advantage over previously proposed benchmark strategies in responding to dynamic changes in bicycle demand and thus reducing the chances of unmet demand.

Related work extends to studying various other problems and questions related to BSS. For example, Yoon et al. (2012) described a smartphone application for the Dublin BSS, which recommends the most suitable pair of bicycle stations concerning the user's current needs and the availability of bicycles, and the possibility of bicycle return at different stations. Shang et al. (2021) analyses the impact of the COVID-19 pandemic on the behavior of users in BSS. It is reported that notable effects of the pandemic on users' behavior are mostly environment friendly.

3 Use Case

We performed our research based on a real-world use case based on data from BicikeLJ, a BSS available in Ljubljana, the capital city of Slovenia. BicikeLJ was introduced in 2011. Users pay a small fee to register to the platform and benefit from the BSS. Bicycle usage under an hour is granted for free, while additional charges apply per hour if used for a more extended period. BicikeLJ provides near-real-time information on the number of bicycles available at each station and makes them available to the public through a website[1].

[1] A map with overlayed information regarding the number of bicycles available at each station is provided at the following web page: `https://www.bicikelj.si/en/mapping`

While this is not always the case, the optimal scenario of a BSS would be that there is always (a) a bicycle available for anyone willing to use it and (b) some parking slot at the station so that the bicycle can be parked. Therefore, in this work, we explore two kinds of prediction models: (i) classification models to predict decision-making options (whether we need to add bicycles, remove them, or no action is required), and (ii) regression models to predict the number of bicycles at the parking station. We envision such models could be used to:

- proactively alert decision-makers and ensure all stations have an adequate number of bicycles and parking slots. Such information could be used as inputs to vehicle routing optimization problems if some truck is used to re-balance the bicycles among the stations.

- provide forecasts to the users and ensure they can better understand if some bicycle or parking slot will be available upon arrival at the station. Such forecasts could be complemented with (a) soft reservations to gather demand data and proactively handle mismatches between supply and demand, and (b) some gamification to incentivize people moving bicycles between stations for certain benefit (Wu (2020)).

For the purpose of this research, we considered multiple forecasting horizons. Such forecasts enable us to consider bicycle-sharing dynamics over time and understand how the quality of the forecasts degrades over time.

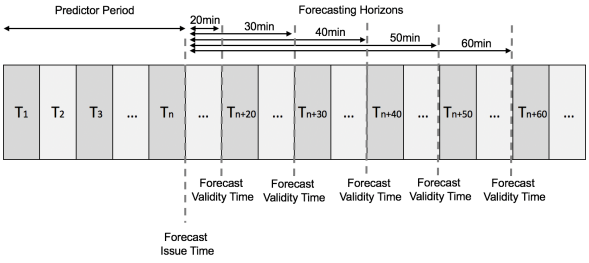


Figure 1: Diagram showing the predictor period taken into account to build the forecasting models, the forecast issue time, and the multiple forecasting horizons considered for this use case. Each model forecasts for a specific horizon and has a particular forecast validity time.

4 Methodology

The subsections below explain the methodology we followed to build and evaluate our forecasting models. We describe how data was collected and analyzed, the features we created, how we defined the target labels and baseline models, and how we trained and evaluated the models.

4.1 Data collection and analysis

The experiment was conducted based on data collected from public APIs provided by the BicikeLJ BSS. Since there are no publicly available historical datasets, we obtained the data from the developer's API and public endpoints of the system. We collected data between March 9th 2022 to May 8th 2022. The final dataset had minor gaps due to networking errors during the web scraping. The data we collected contained two types of records: (i) station states and (ii) bicycle states. Each data record for (i) contains station ID, station name, station address, latitude, longitude, capacity, and the number of free and occupied spaces, while the data from (ii) provide information regarding the bicycle ID, and status, timestamp, and station ID. We collected data for a total of 82 stations and found that there are nearly 800 bicycles offered in the BSS. We then created a single dataset informing the number of bicycles at each station for each reading, which is available at a five-minute frequency. We analyzed the data to find and understand patterns when visualizing the (i) time series (e.g., see Fig. 2), (ii) histograms describing the number of bicycles available at each station (e.g., see Fig. 3), and (iii) heatmaps to understand usual movements of bicycles between stations.

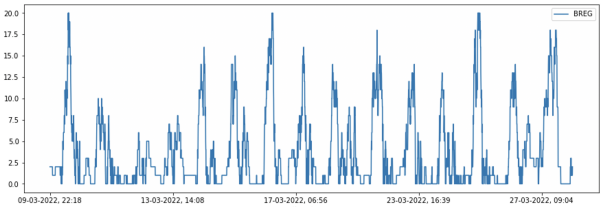


Figure 2: Time series obtained from the bicycle count over the time at the *Breg* station.

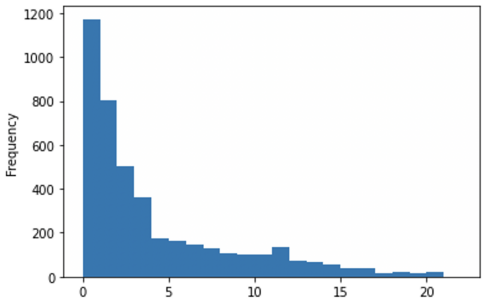


Figure 3: Histogram showing the frequency of bicycle availability for the *Breg* station.

4.2 Feature extraction

While we considered two problems (forecasting decision-making and the number of expected bicycles per station for a given point in time), we considered the same time series features for both. In particular, we

consider that the future number of bicycles available at a particular station is closely related to the current demand. Therefore, we computed two types of features (see Table 1): (a) features indicating how close we are to reaching the full stations' bicycle parking capacity (features 1-4) and (b) features indicating whether the number of bicycles in the station is growing or not (features 5-6). The features were created considering relative values to ensure the same features would generalize well between local and global models and make them comparable.

4.3 Creating target labels for classification

One of the possible use cases related to BSS is to assist decision-making. We frame it as a classification problem, aiming to predict whether three possible states will be reached in the future: (a) no bicycles will be available (therefore need to be supplied), (b) no parking slots will be available (therefore bicycles should be removed), or (c) no action is required, given enough bicycles and parking slots will exist. To determine these three labels, we analyzed each station and the preceding number of bicycles before no bicycles (L_{NB}) or parking slots were found (L_{NPS}). Therefore, three classes were assigned: (a) add bicycles if a less or equal amount of bicycles in a given station are found than the reference amount L_{NB}, (b) do not perform any action, and (c) remove bicycles if an equal or greater amount of bicycles in a given station is found than the reference amount L_{NPS}. The dataset has shown a strong class imbalance.

4.4 Baseline models

We created two baseline models, one for the classification and one for the regression problem. We followed a similar procedure for both baselines to the one described in Section 4.3. However, instead of considering state transitions from bicycles and parking slots available to their deficit, we considered the corrective actions that overturned the deficit of bicycles and parking slots. Therefore, we computed the expected future state based on the last observed value. For classification purposes, we considered that if a deficit of bicycles or parking slots occurs, the following action should overturn such a deficit. If no deficit is observed, no action is taken. For regression purposes, we considered that if a deficit of parking slots arises, we remove bicycles to reach the minimum number of parked bicycles historically observed at that station after no parking slots were available. Analogously, suppose no bicycles are available at the station. In that case, we consider adding a number of bicycles equal to the maximum number of bicycles observed historically at that station after no bicycles were available.

4.5 Model training

To train and evaluate the models, we performed a ten-fold stratified cross-validation, considering the labels of the classification problem. We used the same folds to train the classification and regression models. For the local models, we considered only the data for the same station we predicted. In contrast, for the global models, we trained the model on the data of the train folds of the station we predicted in addition to all the data available from the remaining stations (see Fig. 4).

For the classification problem, we compared two models: logistic regression and gradient boosted trees[2], training the gradient boosted trees for ten iterations using a multiclass loss function. We also compared two models for the regression problem: linear regression and gradient boosted trees, training the gradient boosted trees for ten iterations and considering the Root Mean Square Error (RMSE) as the loss function.

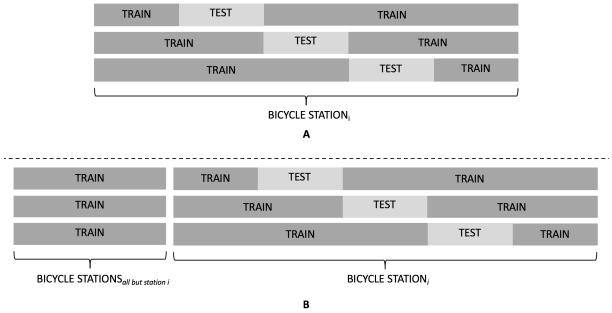


Figure 4: Cross-validation training procedure for (A) local and (B) global models. Given a station *i* for which a forecast is required, local models were trained only on the train folds of that particular station and evaluated on the test folds. Global models, on the other hand, followed the same cross-validation but enriched the train set with data from all the remaining stations.

5 Results and Analysis

We evaluated classification models' discriminative power with the Area Under the Receiver Operating Characteristic Curve (AUC ROC (Bradley (1997))) metric (see Table 2), considering that the metric is invariant to the *a priori* class probabilities, a relevant aspect given the class imbalance observed in the dataset. For the regression problem, we evaluated the models' performance with the Mean Absolute Error (MAE - see Table 3) and RMSE (see Table 4) metrics. While the RMSE metric penalizes large errors, the MAE metric is not sensitive to outliers and therefore provides a better estimate of the models' average performance. The results presented below were computed considering all but the *P+R Barje* bicycle stations.

[2]For the gradient boosted trees, we used the Catboost implementation Prokhorenkova et al. (2018).

#	Feature	Description
1	$mean_5$	The average number of parked bicycles over the last five readings and divided by the station's total capacity.
2	$mean_3$	The average number of parked bicycles over the last three readings and divided by the station's total capacity.
3	$mean_2$	The average number of parked bicycles over the last two readings and divided by the station's total capacity.
4	n-1	The last known number of parked bicycles, divided by the station's total capacity.
5	$mean_5/mean_3$	The average number of parked bicycles over the last five readings and divided by the average number of parked bicycles over the last three readings.
6	$mean_5/mean_2$	The average number of parked bicycles over the last five readings and divided by the average number of parked bicycles over the last two readings.
7	$mean_3/mean_2$	The average number of parked bicycles over the last three readings and divided by the average number of parked bicycles over the last two readings.

Table 1: The features considered to build the classification and regression models.

Model / Forecasting Horizon		20 min.	30 min.	40 min.	50 min.	60 min.
Baseline		0,6658	0,6591	0,6519	0,6443	0,6373
Local	**Logistic Regression**	*0,9901*	*0,9721*	*0,9538*	*0,9369*	*0,9199*
	Catboost Classifier	**0,9939**	**0,9758**	**0,9580**	**0,9418**	**0,9257**
Global	**Logistic Regression**	0,9287	0,9120	0,8959	0,8806	0,8665
	Catboost Classifier	0,9279	0,9114	0,8952	0,8801	0,8654

Table 2: Mean AUC ROC values (higher is better) were obtained across the stations for each forecasting horizon. The mean was computed over the mean AUC ROC obtained from the ten-fold cross-validation performed for each station. Best results are bolded, and second-best are displayed in italics.

Model / Forecasting Horizon		20 min.	30 min.	40 min.	50 min.	60 min.
Baseline		1,3395	1,5189	1,6694	1,8024	1,9247
Local	**Linear Regression**	**0,1607**	**0,6099**	**0,8979**	**1,1177**	**1,3016**
	Catboost Regressor	*0,2444*	*0,6600*	*0,9303*	*1,1370*	*1,3109*
Global	**Linear Regression**	0,8765	1,1722	1,3882	1,5622	1,7125
	Catboost Regressor	0,8642	1,1570	1,3713	1,5396	1,6882

Table 3: Mean MAE values (lower is better) were obtained across the stations for each forecasting horizon. The mean was computed over the mean MAE obtained from the ten-fold cross-validation performed for each station. Best results are bolded, and second-best are displayed in italics.

Model / Forecasting Horizon		20 min.	30 min.	40 min.	50 min.	60 min.
Baseline		2,0567	2,2601	2,4392	2,5998	2,7478
Local	**Linear Regression**	**0,3180**	**1,0519**	**1,4344**	**1,7131**	**1,9393**
	Catboost Regressor	*0,4220*	*1,0941*	*1,4601*	*1,7284*	*1,9460*
Global	**Linear Regression**	1,1319	1,6041	1,9115	2,1503	2,3508
	Catboost Regressor	1,1326	1,5897	1,8952	2,1294	2,3280

Table 4: Mean RMSE values (lower is better) were obtained across the stations for each forecasting horizon. The mean was computed over the mean RMSE obtained from the ten-fold cross-validation performed for each station. Best results are bolded, and second-best are displayed in italics.

We found that all models we created performed better than the baseline. On average, the local models performed best, with the Catboost Classifier leading over the Logistic Regression when predicting the decision-making actions and the Linear Regression leading over the Catboost Regressor when predicting the number of bicycles to be found at the station for a particular forecasting horizon. Nevertheless, when considering global models, we found that the Logistic Regression displayed slightly better performance against the Catboost Classifier in all cases, and the Catboost Regressor outperformed the Linear Regression when measuring MAE and RMSE in all cases, except for RMSE at the time horizon of twenty minutes.

When analyzing the impact of the time horizon on the models' performance, we found that the local models are not strongly affected, displaying an AUC ROC above 0,9 even when predicting an hour ahead. Regression models, on the other side, suffer more pronounced performance degradation. The best model increases the error eight and six times when comparing MAE and RMSE correspondingly, between time horizons of twenty and sixty minutes.

Finally, we were interested in which stations our forecasting models displayed the best and worst performance. We found that local models performed best when predicting decision-making for the *Črnuče* station and the *Zalog* station when forecasting the number of bicycles available. On the other hand, while the worst performance for the regression problem was observed for the *Nama* station, for classification different stations were found depending on the forecasting horizon (e.g., *P+R Dolgi Most* for twenty and thirty minutes, *Lidl Rudnik* for forty and sixty minutes, and *GH Šentpeter Njegoševa cesta* for fifty minutes). When considering global models, decision-making was best predicted for the *Bordarjev trg* station, while the worst performance was observed for the *Studenec* station. For regression, we found the best performance was displayed at the *BS4 Stožice* station. The worst performance for the MAE metric was found at the *Zalog* station. In contrast, for the RMSE metric, the worst performance was split among the *Zalog* (predicting twenty and thirty minutes ahead), *Kolodvorska ulica* (predicting forty minutes ahead), and *Nama* (predicting fifty and sixty minutes ahead) stations. Finally, when looking at the baseline performance for classification, we found that the best performance was observed at the *Živalski vrt* station and the worst one for the *Športni Center Stožice* in all forecasting horizons, except when forecasting sixty minutes ahead, where the worst performance was obtained for the *Mercator Center Šiška*. On the other hand, when considering the regression problem, we found that the best performance was obtained at the *Zalog* station, while the worst one at the *Parkirišče NUK 2-FF* station.

6 Conclusion

Optimizing BSS enables to amplify their use further and minimize the adverse effects on nature. In our work, we trained and compared ML models to predict when rebalancing of the bicycles across stations is re-

quired. We evaluated our approach on real-life data from BicikeLJ system in Ljubljana and achieved superior results compared to the baseline methods.

Given that our model can predict bicycle shortage or overflow in advance, this information is helpful for the caretakers of the system. Bicycles could be relocated on time by exploiting other means of transport, thus allowing for higher bicycle use. Another way we can exploit the model's insightful information is to increase the usability of the BSS by incentivizing users to move bicycles to the desired locations, i.e., from stations with no parking capacity to the empty ones.

The main limitation of our study is the real-life implementation, where the relocation of bicycles should consider at least three constraints: (a) must be performed timely, (b) in an environment-friendly way, and (c) be financially efficient. Therefore, our future work will aim to develop a holistic solution while considering the introduction and evaluation of our results in the practical operation of Ljubljana BSS, and the abovementioned constraints. Furthermore, we plan to test our approach on multiple BSS.

References

Almannaa, M. H., Elhenawy, M., and Rakha, H. A. (2020). Dynamic linear models to predict bike availability in a bike sharing system. *International journal of sustainable transportation*, 14(3):232–242.

Ashqar, H. I., Elhenawy, M., Almannaa, M. H., Ghanem, A., Rakha, H. A., and House, L. (2017). Modeling bike availability in a bike-sharing system using machine learning. In *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pages 374–378. IEEE.

Ashqar, H. I., Elhenawy, M., Rakha, H. A., Almannaa, M., and House, L. (2021). Network and station-level bike-sharing system prediction: A san francisco bay area case study. *Journal of Intelligent Transportation Systems*, pages 1–11.

Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145 – 1159.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Cagliero, L., Cerquitelli, T., Chiusano, S., Garza, P., and Xiao, X. (2017). Predicting critical conditions in bicycle sharing systems. *Computing*, 99(1):39–57.

Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., et al. (2018). Global warming of 1.5 c. *An IPCC Special Report on the impacts of global warming of*, 1(5).

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., and Gulin, A. (2018). Catboost: unbiased boosting with categorical features. *Advances in neural information processing systems*, 31.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., Creutzig, F., Chayes, J., and Bengio, Y. (2022). Tackling Climate Change with Machine Learning. *ACM Computing Surveys (CSUR)*, 55(2):96.

Seo, Y.-H., Kim, D.-K., Kang, S., Byon, Y.-J., and Kho, S.-Y. (2022). Rebalancing docked bicycle sharing system with approximate dynamic programming and reinforcement learning. *Journal of Advanced Transportation*, 2022.

Shaheen, S. A., Guzman, S., and Zhang, H. (2010). Bikesharing in europe, the americas, and asia: past, present, and future. *Transportation research record*, 2143(1):159–167.

Shang, W.-L., Chen, J., Bi, H., Sui, Y., Chen, Y., and Yu, H. (2021). Impacts of covid-19 pandemic on user behaviors and environmental benefits of bike sharing: A big-data analysis. *Applied Energy*, 285:116429.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Wu, J. (2020). Challenges and opportunities in algorithmic solutions for re-balancing in bike sharing systems. *Tsinghua Science and Technology*, 25(6):721–733.

Yang, Z., Hu, J., Shu, Y., Cheng, P., Chen, J., and Moscibroda, T. (2016). Mobility modeling and prediction in bike-sharing systems. In *Proceedings of the 14th annual international conference on mobile systems, applications, and services*, pages 165–178.

Yoon, J. W., Pinelli, F., and Calabrese, F. (2012). Cityride: a predictive bike sharing journey advisor. In *2012 IEEE 13th International Conference on Mobile Data Management*, pages 306–311. IEEE.

A Systematic Literature Review on Topic Modelling and Sentiment Analysis

Josip Kunsabo, Jasminka Dobša
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, Varaždin, Croatia
{jkunsabo, jasminka.dobsa}@foi.hr

Abstract. *The human need for organization has always existed, but due to the ever-increasing generation of digital records, it is becoming almost necessary. With the development of information technologies, the use and application of various machine learning methods in the processing of natural language is becoming more widespread. In this paper we will give an overview of methods of topic modelling and sentiment analysis as well as approaches in processing of documents needed for their application. These two tasks of natural language processing are naturally related in the problem of identification of sentiments in the topics automatically extracted from the collection of observed documents.*

Keywords. topic modelling, sentiment analysis, machine learning, deep learning, sentiment lexicon

1 Introduction

With the development and availability of the Internet, large amounts of information, i.e. digital records, are being created. Such records include various online publications, portals, blogs, forums, social networks, streaming sites, public administration data, company data, various databases, and the like. Timeliness and easy availability with the possibility of open and free use provide readers with access to a large and diverse amount of information. Metadata processing can help structure available records, but their large number and unstructuredness, often make this challenge impossible.

The availability of computer systems and information technologies with the application of natural language processing methods enable a significant shift in the speed and quality of processing text sets of data by various statistical methods, as well as machine and deep learning methods. Thus, methods of topic modelling and sentiment analysis can be used to identify and structure textual content.

Topic modelling is a data mining task of discovering topics that appear in the observed set of documents (Blei et al., 2003). The basic approach involves machine learning to group similar topics, often through various lexicons that define affiliation to a particular topic. Probabilistic methods of unsupervised machine learning have also been widely implemented to find hidden topics in electronic publications.

Computer sentiment analysis, also known as "opinion mining", is an area of text mining that analyzes human opinions, sentiments, attitudes, and emotions toward entities such as products, services, organizations, individuals, events, topics, and their attributes (Pang and Lee, 2008). There are several approaches to sentiment analysis but the most common is the approach in which the text is labeled as positive or negative. Sentiment classification can be either a binary or a multi-class problem. Binary sentiment analysis is the classification of texts into positive and negative classes, while multi-class sentiment analysis focuses on classifying data into fine-grained labels or multi-level intensities. (Minaee et al., 2021).

Topic modelling, as well as sentiment analysis, attracted researchers from academia and industry, and this paper will provide an overview of the methods used in this area.

2 Literature review

Topic modelling as well as sentiment analysis are well-studied methods in the last decade, as evidenced by the large increase in results in the past period. Based on searches of Ebsco, Scopus and Web of Science databases by keywords "topic modelling", "sentiment analysis" and combinations of words "topic modelling" and "sentiment analysis", a large increase in publication of scientific papers related to these topics can be observed in the period from 2012 and later compared to the period before 2012 (Table 1.). Particularly, huge increase can be observed for the topic of sentiment analysis and application of topic modelling methods for it. The reasons for this can be found in the greater availability of computer systems and information technologies, as well as in the fact that technological development has enabled complex or

extensive computer tasks to be processed in a shorter time. Furthermore, increasingly diverse areas of application and promising results are of growing interest to researchers and to the industry.

With the aim of analyzing methods for topic modelling and sentiment analysis of electronic publications, the results of the last six years of articles covering the above keywords were isolated and 313 articles in English were filtered for further analysis. A review of the results shows domination of health-related topics, mostly related to COVID-19, while topics before COVID-19 were mostly related to the oil and financial sector. Considering that COVID-19 was an integral part of the daily life of a large part of the world's population in the past period, it is understandable to observe its impact from all aspects (from health to economy).

Table 1. Statistics for database search

Query	Period	Database		
		Ebsco	Scopus	Web of Science
Topic modelling	<2012	8,091	27,891	192,212
	2012-2022	12,575	52,192	384,245
Sentiment analysis	<2012	2,239	2,186	1,175
	2012-2022	15,800	26,850	23,898
Topic modelling and sentiment analysis	<2012	183	38	28
	2012-2022	986	1,454	2,976

Looking at data sources, most of them are data from the social network Twitter and specialized databases or other electronic publication. The reasons for using Twitter as a source of data is in the fact that it is one of the most used social networks, which also allows easy and open access to data.

For further analysis of articles relevant to topic modelling and sentiment analysis, based on the reading of abstracts, 60 papers were selected, i.e. 45 of them remained after reading the whole paper. An analysis of the selected papers with an overview of type of textual documents and their amount used, text preprocessing methods and the methods used for topic modelling and sentiment analysis are shown in Table 2. The column that describes the data sets used contains two values, “article” and “post”. “Article” includes text documents with longer content such as publicly available articles, blogs, financial or media reports and reviews, while “post” means shorter messages such as posts on social networks or comments on articles.

3 Discussion

Topic modelling as well as sentiment analysis are confirmed as very current topics. Research shows that it attracts researchers from all over the world to understand the events around them. Research related to COVID-19 dominates thematically, but the observed aspects of the impact of COVID-19 are very diverse. Thus, topics related to crisis and health communication during the pandemic, as well as topics related to fears and views on vaccination are addressed. Also, comparison of communication in the news in relation to social networks and analysis of online education during a pandemic and the like is being made.

In order to select the appropriate research sample, when processing “post” data it is noticeable the dominance of using Twitter as a source of data with the use of filtering by keywords related to the research topic. In such processing, samples range from several thousand to several million records. When processing “article” data, only a few researchers included a very large number of publicly available news articles, which were not previously filtered on the basis of keywords selected according to the area of interest. The remaining research used specific reports and records of targeted exploration areas such as healthcare or oil.

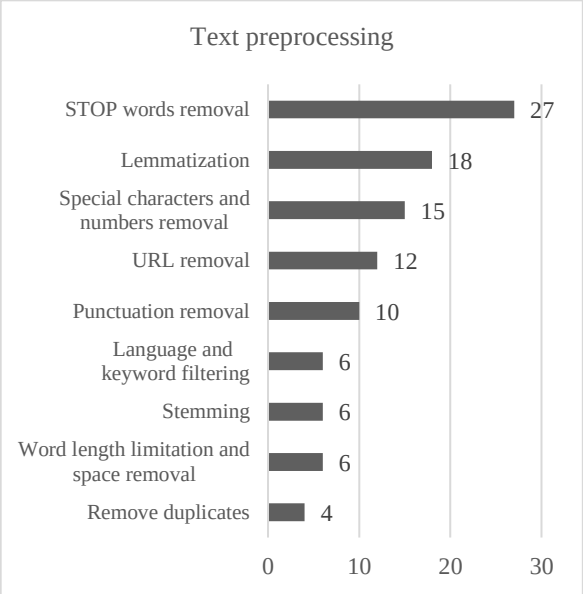


Figure 1. Distribution of text preprocesing methods

Textual records in most research go through a preprocessing phase of data. As shown on Figure 1, preprocessing often includes removing of stop words, URL links, punctuation and special characters, duplicate results, and changing uppercase letters to lowercase letters.

Lemmatization and tokenization are also often applied in the preprocessing phase. Lemmatization is the operation that maps words into their basic form and tokenization is a step that involves recognizing and

Table 2. Overview of the analyzed papers

Paper	Dataset	Preprocessing	Methods of topic modelling	Methods/lexicons for sentiment analysis
Kim et al. (2015)	16,189 articles 7,106,297 posts	stop words removal, Lemmatization	n-gram LDA	PKDE4J SO-CAL
Bandhakavi et al. (2016)	13,750 articles 280,000 posts	-	PMI, sLDA	DSEL
Liu et al. (2017)	1,455,947 posts	stop words removal, Stemming	DTSA	
Morimoto et al. (2017)	298,205 articles	stop words removal	MDTM	-
Cerchiello et al. (2018)	553,666 articles	stop words removal, Word length limitation	STM	-
Huang et al. (2018)	21,984 articles	stop words removal, Lemmatization	LDA	VADER
Li et al. (2018)	6,756 articles	stop words removal,	LDA DTM	Textblob
Carrera, Berny & Jung, Jae-Yoon. (2018)	143,876 posts	Language filtering, Punctuation and stop words removal, Lemmatization	LDA	Naive Bayes
Gambino et al. (2018)	17,743 articles 4,412 posts	Lemmatization, URL and numbers removal,	Binary Relevance (BR)	Naive Bayes, Random Forest, SVM
Tahmassebi et al. (2018)	Article stream	stop words removal, Lemmatization	LDA	SentiWordNet
Zolnoori et al. (2019)	3,763,737 articles	Special characters removal, Lemmatization, Stemming	TKM	VADER
Xu et al. (2020)	3,715 articles	Space removal, Punctuation, stop words removal, Lemmatization	MLbased	
Daou (2020)	1,216,343 posts	-	LDA, NMF, LSA	VADER SentiStrength
Liu et al. (2020)	7,791 articles	Remove duplicates, stop words removal, Stemming	LDA	-
Valdez et al. (2020)	86,581,237 posts	stop words removal	LDA	VADER
Abd-alrazaq et al. (2020)	167,073 posts	Language filtering, Punctuation and stop words removal, Lemmatization	LDA	Textbloob
Li et al. (2020)	9,715 articles	-	DP-Sent-LDA	Bi-LSTM
Muhamedyev, et al. (2020)	804,829 articles	-	BigARTM	BERT

Zhou et al. (2020)	2,211 articles	Numbers removal, Punctuation and stop words removal, Word length limitation	STM	DUST
Chu et al. (2020)	1,040 articles	Punctuation and stop words removal	LDA	NRC, Bing, AFINN,
Rudra et al. (2020)	600 articles	Special characters removal	LDA	Naive Bayes
Shah et al. (2021)	96,234 posts	URL removal, Special characters and stop words removal, Lemmatization	DTM	CrystalFeel
Yadav et al. (2021)	12,451,298 posts	Special characters and URL removal, stop words removal, Stemming, Lemmatization	LDA	SentiWordNet AFINN
Melo et al. (2021)	18,413 articles 1,597,934 posts	URL removal, Special characters, Punctuation and stop words removal, Lemmatization, Stemming	LDA	VADER
Monselise et al. (2021)	7,948,886 posts	Remove duplicates, URL removal and stop words removal, Lemmatization	NMF	VADER BERT
Lyu et al. (2021)	227,840 posts	Remove duplicates, URL and stop words removal, Lemmatization	LDA	NRC
Zhang (2021)	404 articles	stop words removal	LDA	LSD
Cheng et al. (2021)	800,000 posts	-	LDA	VADER, NER, RF
Febro and Catindig (2021)	9,842 articles	Punctuation, Special characters and stop words removal	TKM	-
Ghanem et al. (2021)	747,018 posts	Keyword filtering, Language filtering	MD-ULM	
Liew and Lee (2021)	672,133 posts	Space removal, Punctuation and stop words removal, Lemmatization	STM	VADER
Kuo et al. (2021)	125,570 articles	-	LDA	ANTUSD
Gerts D et al. (2021)	1,800,000 posts	URL removal, stop words removal	DTM	AFIN
Yakunin et al. (2021)	1,142,735 articles	-	LDA	Word-Net, ConceptNet,
Liu and Huang (2021)	30,789 articles	-	ODEE	VADER

Smith and Cipolli, (2021).	8,013 posts	Space removal, URL removal, Punctuation, Numbers and stop words removal	LDA	NRC
Kregel et al. (2021)	95,000 articles	URL removal, Lemmatization	LDA	SentiWordNet
Ghasiya et al. (2021)	102,278 articles	URL removal	top2vec	RoBERTa
van der Veen and Bleich (2021)	15,121 articles	-	NMF	lexicon
Dornick et al. (2021)	57,921 articles	Language filtering, Special characters and numbers removal, stop words removal, Lemmatization	BERT	TextBlob
Zheng et al. (2022)	315,136 posts	Remove duplicates, Special characters and URL removal, stop words removal, Lemmatization	LDA	LIWC
Oliveira et al. (2022).	1,700,000 posts	Special characters and numbers removal, Word length limitation, Lemmatization	LDA	RoBERTa
Waheeb et al. (2022)	853,000 posts	-	LDA	VADER, ULMS, TextBloob, SentiWordNet, TSA
Idler et al. (2022)	704 articles	Keyword filtering	STM	Sentimentr
Yin et al. (2022).	75,665 posts	URL removal, Special characters and stop words removal, Punctuation, Lemmatization, Stemming	LDA	VADER

separating individual phrases. Stemming is the process of reducing words to their word stem or root form. Already created packages in R or Python are most often used when performing preprocessing.

The Latent Dirichlet Allocation (LDA) method (Blei et al., 2003) was used in topic modelling in 23/45 (51%) of the research papers. LDA is a probabilistic model with a hierarchical structure for its components, which include documents, topics, and words. It assumes that a given document is generated from a mixture of topics and these topics produce the words in the documents according to their probabilistic distributions. LDA backtracks and derives the hidden topics that create those documents on the basis of the statistics of the included words.

In addition to the LDA, research also uses methods based on LDA, such as Dynamic Topic Modelling (DTM) (Blei and Lafferty, 2007), supervised Latent Dirichlet Allocation (sLDA) (Blei and McAuliffe, 2007), Dependency Parsing Sentence Latent Dirichlet Allocation (DP-Sent-LDA) (Li, et al., 2018) and Big Additive Regularization of Topic Models (BigARTM) (Muhamedyev, et al., 2020). DTM divides data into time slices based on the time they were generated. The set of topics at each time slice is then assumed to evolve from the set of topics at the previous time slice using a state space model. The result is an evolving probability distribution of words for each topic that shows how certain words become more or less important over time for the same topic (Gerts D et al., 2021). In Multiscale Dynamic Topic Model (MDTM),

current topic-specific distributions over words are assumed to be generated based on the multiscale word distributions of the previous epoch. (Iwata T et al., 2010). In sLDA, a response variable associated with each document is added to LDA. sLDA jointly models the documents and the responses, in order to find latent topics that will best predict the response variables for future unlabeled documents (Blei and McAuliffe, 2007). The LDA is a kind of bag-of-words model, which assumes that the order of words within a document does not matter, and different words in one sentence are likely to come from different topics. On the other hand, Sent-LDA model relaxes the assumption of bag-of-words and makes the assumption that all words in one sentence are sampled from the same topic (Bao and Datta 2014). Redundant words in sentences can disturb the effectiveness of discovering when using Sent-LDA, so Dependency Parsing (DP) is introduced (Li, et al., 2018). DP is a natural language processing technology that can reveal the syntactical structure of sentences and improves the effectiveness of Sent-LDA by removing the redundant words within the short text. BigARTM is an open-source library for parallel construction of thematic models on large text cases whose implementation is based on the additive regularization approach (ARTM) (Muhamedyev, et al. 2020).

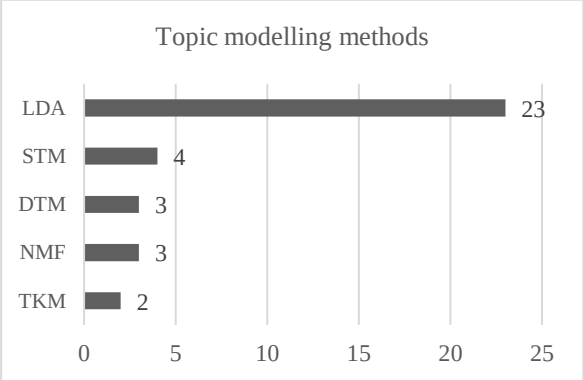


Figure 2. Frequency of usage of topic modelling methods used in at least two analysed papers

As shown on Figure 2., the usage of Structural Topic Modelling (STM) (Roberts et al. 2016), Topic Keyword Model (TKM) (Schneider and Vlachos, 2018). and Non-negative Matrix Factorization (NMF) (Lee and Seung, 1999) is noticeable. STM uses document-level metadata to estimate how the prevalence of topics might vary (Roberts et al. 2016). TKM considers the word order in a text document for topic modelling. TKM links a word to a topic if it or its adjacent words have a high association score with the topic (Schneider and Vlachos, 2018). NMF is a partitioning method that uses nonnegative factorization of the term-document matrix of a given corpus (Lee and Seung, 1999).

When analyzing sentiment, special lexicons are often used. All the lexicons and data sets in analyzed papers are in English except a few data sets in Portuguese (Melo et al., 2021), Arabic (Ghanem et al., 2021) or Chinese (Liu et al., 2020). Beside unsupervised approach of sentiment analysis using lexicons, a supervised approach using standard classification algorithms such as Naive Bayes and Random Forest classifiers is also noticeable.

The Valence Aware Dictionary and Sentiment Reasoner (VADER) lexicon was used in 11/45 (24%) sentiment analyses (Liew, Lee, 2021). To identify key sentiment (positive, neutral and negative), VADER uses a human-curated lexicon of 7500 emotion-related words as well as five human-interpretable rules that identify sentiment intensity.

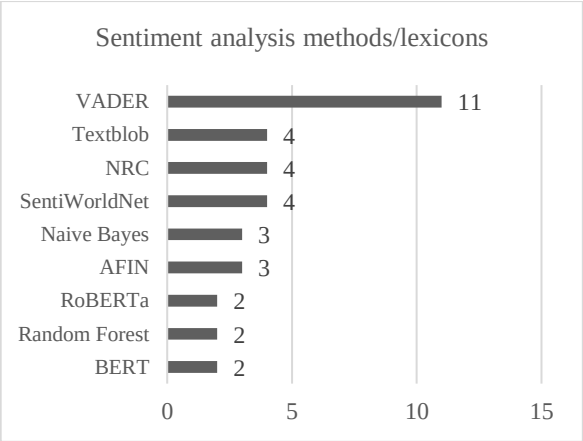


Figure 3. Frequency of usage of sentiment analysis methods/lexicons used in at least two analysed papers

In addition to VADER, National Research Council of Canada Emotion (NRC) (Mohammad and Turney, 2013), TextBlob, and SentiWordNet lexicons as well as BERT (Bidirectional Encoder Representations from Transformers) (Devlin et al., 2019) and RoBERTa (Robustly Optimized BERT Pretraining Approach) (Y. Liu et al., 2019) models were used for sentiment analysis (Figure 3). NRC is a sentiment and emotion lexicon that contains over 14,000 English words and their associations with two sentiments (positive, negative) and eight basic emotions: anger, fear, anticipation, trust, surprise, sadness, joy, and disgust (Mohammad and Turney, 2013). Textblob is a Python library which provides a simple API for diving into common natural language processing tasks, such as sentiment analysis, classification, translation, etc (Li et al., 2018), while SentiWordNet was designed to support sentiment analysis applications by providing an annotation based on three numerical sentiment values of positivity, negativity, and neutrality (Kregel et al., 2021). BERT is a word representation model that uses unannotated text to perform various natural language processing tasks such as classification and question answering (Devlin et al., 2019). RoBERTa is an extension of the original BERT model. Compared to BERT, RoBERTa is trained on more data, with longer sequences, bigger batches, for a longer time and does

not include the next sentence prediction objective during pretraining (Y. Liu et al., 2019).

Some research that have chosen to approach a common model for topic modelling and sentiment analysis such as Moroccan Dialect Universal Language Model (MD-ULM) (Ghanem et al., 2021), Manifold Learning-based model (ML-based) (Oikawa et al. 2016) and Dynamic Topic-based Sentiment Analysis model (DTSA) (Liu et al., 2017) models can be highlighted. MD-ULM is pre-trained model on large unlabeled corpora, which is fine-tuned using COVID-19 dataset (Ghanem et al., 2021). Manifold learning (ML) is an effective method for nonlinear dimensionality reduction (NLDR), whose main objective is to discover meaningful low dimensional structures hidden in high dimensional data (Oikawa et al. 2016). DTSA assumes that topic-based sentimental evolution of the current time epoch is influenced by its historical sentiment of multiple time slides (Liu et al., 2017).

When evaluating results in the case of topic modelling, the most common is the application of coherence score (Syed and Spruit, 2017) or cosine similarity (Rahutomo et al., 2012). When observing the evaluation of sentiment analysis results, it is common to use metrics of precision, recall, and F1-score measures. Data visualization in most research includes tabular and temporal representations. Some research uses a “text cloud” to describe topics, which clearly shows what the topic covers, while some researchs stand out using Social Network Analysis (SNA) methods and present their results via a network graph.

4 Conclusion

Topic modelling is a challenging area that provides insight into structure of data in electronic media and enables identification of sentiments present in the data. Although models such as the LDA have long been present in the scientific community and have been well studied, their dominant use confirms that they serve their purpose in determining the topics of text documents. Other models, although not necessarily new, are used to a lesser extent.

While the topic modelling task is independent of the language, language is important when analyzing sentiment. Most of the observed research determined sentiment with the help of one of the lexicons, i.e. words associated to the sentiment. As the words themselves have a different spelling in different languages, to determine sentiment, it is necessary to have a lexicon. This is the reason why English, as the most common language, was covered in most of the research. The frequency of using VADER lexicons is noticeable, but the diversity of lexicons tells us that there are still many different options that researchers are experimenting with. In addition to the lexicon approach, there are also models based on machine learning with the “gold standard” as well as pre-trained BERT models, which will be proven in future research. The most often used methods of machine and deep learning, as well as the amount and types of dictionaries that exist on different natural languages, will be part of future research.

References

Abd-alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M., & Shah, Zubair. (2020). Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study. Journal of Medical Internet Research. 22. 10.2196/19016.

Bandhakavi, A. S., Wiratunga, N., Deepak, P., & Massie, S. (2016). Lexicon based Feature Extraction for Emotion Text Classification. Pattern Recognition Letters. 93. 10.1016/j.patrec.2016.12.009.

Bao, Y., & Datta, A. (2014). Simultaneously discovering and quantifying risk types from textual risk disclosures. Management Science, 60, 1371–1391.

Blei, D., Ng, A., & Jordan, M. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research. 3. 993-1022. 10.1162/jmlr.2003.3.4-5.993.

Blei D, & Lafferty J. Dynamic topic models. In: ICML '06: Proceedings of the 23rd international Conference on Machine learning. 2006 Jun Presented at: 23rd International Conference on Machine Learning; June 25-29, 2006; Pittsburgh, PA p. 113-120.

Blei, D., & McAuliffe, J. (2007), Supervised topic models, in: Advances in Neural Information Processing Systems 20 (NIPS 2007).

Carrera, B., & Jung, J.Y. (2018). SentiFlow: An Information Diffusion Process Discovery Based on Topic and Sentiment from Online Social Networks. Sustainability. 10. 2731. 10.3390/su10082731.

Cerchiello, P., & Nicola, G. (2018). Assessing News Contagion in Finance. Econometrics. 6. 5. 10.3390/econometrics6010005.

Cheng, I., Heyl, J., Lad, N., Facini, G., & Grout, Z. (2021). Evaluation of Twitter data for an emerging crisis: an application to the first wave of COVID-19 in the UK. Scientific Reports. 11. 10.1038/s41598-021-98396-9.

Chu, C.Y., Park, K., & Kremer, G. (2020). A global supply chain risk management framework: An application of text-mining to identify region-specific supply chain risks. Advanced Engineering Informatics. 45. 10.1016/j.aei.2020.101053.

Daou, H. (2020). Sentiment of the public: the role of social media in revealing important events. Online Information Review. ahead-of-print. 10.1108/OIR-12-2019-0373.

Devlin J, Chang M, Lee K, & Toutanova K. BERT: pretraining of deep bidirectional transformers for language understanding. 2019 Presented at: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics; June 2019; Minneapolis, Minnesota. [doi: 10.18653/v1/n18-2]

Dornick, C., Kumar, A., Seidenberger, S., Seidle, E., & Mukherjee, P. (2021). Analysis of Patterns and Trends in COVID-19 Research. Procedia Computer Science. 185. 302-310. 10.1016/j.procs.2021.05.032.

Febro, J., & Catindig, M. (2021). Exploring cyber violence against women and girls in the Philippines through Mining Online News [Explorando la ciberviolencia contra mujeres y niñas en Filipinas a través de Mining Online News]. Comunicar. 30. 10.3916/C70-2022-10.

Gambino, O., & Calvo, H. (2018). Modelling distribution of emotional reactions in social media using a multi-target strategy. Journal of Intelligent & Fuzzy Systems. 34. 2837-2847. 10.3233/JIFS-169471.

Gerts, D., Shelley, C., Parikh, N., Pitts, T., Watson, R.C., Fairchild, G., Vaquera, C. N., & Daughton, A. “Thought I’d Share First” and Other Conspiracy Theory Tweets from the COVID-19 Infodemic: Exploratory Study JMIR Public Health Surveill 2021;7(4):e26527 URL: https://publichealth.jmir.org/2021/4/e26527 DOI: 10.2196/26527

Ghanem, A., Asaad, C., Hafidi, H., Moukafih, Y., Guermah, B., Sbihi, N., Zakroum, M., Ghogho, M., Dairi, M., Cherqaoui, M., & Baïna, K. (2021). Real-Time Infoveillance of Moroccan Social Media Users’ Sentiments towards the COVID-19 Pandemic and Its Management. International Journal of Environmental Research and Public Health. 18. 12172. 10.3390/ijerph182212172.

Ghasiya, P., & Okamura, K. (2021). Investigating COVID-19 News across Four Nations: A Topic Modelling and Sentiment Analysis Approach. IEEE Access. 9. 36645-36656. 10.1109/ACCESS.2021.3062875.

Huang M., ElTayeby O., Zolnoori M & Yao L. Public Opinions Toward Diseases: Infodemiological Study on News Media Data. J Med Internet Res. 2018 May 8;20(5):e10047. doi: 10.2196/10047. PMID: 29739741; PMCID: PMC5964307.

Idler, E., Bernau, J., & Zaras, D, (2022). Narratives and counter-narratives in religious responses to COVID-19: A computational text analysis. PloS one. 17. e0262905.; 0.1371/journal.pone.0262905.

Iwata, T., Yamada, T., Sakurai, Y., & Ueda, N. (2010). Online Multiscale Dynamic Topic Models. 663-672. 10.1145/1835804.1835889.

Kim, H.J., Jeong, Y.K., Kim, Y., Kang, K., & Song, M. (2015). Topic-based content and sentiment analysis of Ebola virus on Twitter and in the news. Journal of Information Science. 42. 10.1177/0165551515608733.

Kregel, I., Koch, J., & Plattfaut, R. (2021). Beyond the Hype: Robotic Process Automation's Public Perception Over Time. Journal of Organizational Computing and Electronic Commerce. 31. 1-21. 10.1080/10919392.2021.1911586.

Kuo, H.Y., Chen, S.Y., & Lai, Y.T. (2021). Investigating COVID-19 News before and after the Soft Lockdown: An Example from Taiwan. Sustainability. 13. 11474. 10.3390/su132011474.

Lee DD, & Seung HS. Learning the parts of objects by non-negative matrix factorization. Nature 1999 Oct 21;401(6755):788-791. [doi: 10.1038/44565] [Medline: 10548103]

Li, J., Li, G., Yao, Y., & Zhu, X. (2018). Identifying the influence factors of commodity futures market through a new text mining approach. In The 7th international conference on futures and other derivatives. Shanghai, China.

Li, X., Shang, W., & Wang, S. (2018). Text-based crude oil price forecasting: A deep learning approach. International Journal of Forecasting. 35. 10.1016/j.ijforecast.2018.07.006.

Li, J., Li, G., Liu, M., Zhu, X., & Wei, L. (2020). A novel text-based framework for forecasting agricultural futures using massive online news headlines. International Journal of Forecasting. 10.1016/j.ijforecast.2020.02.002.

Liew, T., & Lee, C. (2021). Examining the Utility of Social Media in COVID-19 Vaccination: Unsupervised Learning of 672,133 Twitter Posts (Preprint). JMIR Public Health and Surveillance. 7. 10.2196/29789.

Liu, J. & Huang, X. (2021). Forecasting Crude Oil Price Using Event Extraction. Papers 2111.09111, arXiv.org.

Liu, P., Gulla, J. & Zhang, L. (2017). A joint model for analyzing topic and sentiment dynamics from large-scale online news. World Wide Web. 21. 1-23. 10.1007/s11280-017-0474-9.

Liu, Q., Zheng, Z., Zheng, J., Chen, Q., Liu, G., Chen, S., Chu, B., Zhu, H., Akinwunmi, B., Huang, J., Zhang, C., & Ming, W.K. (2020). Health Communication Through News Media During the Early Stage of the COVID-19 Outbreak in China:

A Digital Topic Modelling Approach. 10.1101/2020.03.29.20043547.

Lyu, J., Han, E., & Luli, G. (2021). COVID-19 Vaccine–Related Discussion on Twitter: Topic Modelling and Sentiment Analysis (Preprint). 10.2196/preprints.24435.

Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad K.N., Asgari-Chenaghlu, M., & Gao, J. (2021). Deep Learning--based Text Classification: A Comprehensive Review. ACM Computing Surveys. 54. 1-40. 10.1145/3439726.

Melo, T., & Figueiredo, C. (2021). Comparing News Articles and Tweets About COVID-19 in Brazil: Sentiment Analysis and Topic Modelling Approach. JMIR Public Health and Surveillance. 7. 10.2196/24585.

Mohammad, S. & Turney, P. (2013). Crowdsourcing a Word-Emotion Association Lexicon. Computational Intelligence. 29. 10.1111/j.1467-8640.2012.00460.x.

Monselise, M., Chang, C.H., Ferreira, G., Yang, R., & Yang, C. (2021). Topics and Sentiments of Public Concerns Regarding COVID-19 Vaccines: Social Media Trend Analysis. Journal of Medical Internet Research. 23. 10.2196/30765.

Morimoto, T., & Kawasaki, Y. (2017). Forecasting Financial Market Volatility Using a Dynamic Topic Model. Asia-Pacific Financial Markets. 24. 10.1007/s10690-017-9228-z.

Muhamedyev, R., Yakunin, K., Mussabayev, R., Buldybayev, T., Kuchin, Y., Murzakhmetov, S., & Yelis, M. (2020). Classification of Negative Information on Socially Significant Topics in Mass Media. Symmetry. 12. 1945. 10.3390/sym12121945.

Oikawa, M.A., Dias, Z., de Rezende Rocha, A. (2016). Manifold learning and spectral clustering for image phylogeny forests. IEEE Transactions on Information Forenstcs and Security, 11(1), 5–19.

Oliveira, F., Haque, A., Mougouei, D., Evans, S., Sichman, J., & Singh, M. (2022). Investigating the Emotional Response to COVID-19 News on Twitter: A Topic Modelling and Emotion Classification Approach. IEEE Access. 10. 16883-16897. 10.1109/ACCESS.2022.3150329.

Pang, B. & Lee, L. (2008). Opinion Mining and Sentiment Analysis. Foundations and Trends in Information Retrieval. 2. 1-135. 10.1561/1500000011.

Rahutomo, F. & Kitasuka, T. & Aritsugi, M. (2012). Semantic Cosine Similarity.

Roberts, M.E., Brandon M. S., & Dustin T. 2016a. Navigating the Local Modes of Big Data: The Case of Topic Models. In Data Analytics in Social Science, Government, and Industry. New York: Cambridge University Press.

Rudra B.P., Kokatnoor, S., Yadav, C., & Mounika, M. (2020). Polarity Detection on Real-Time News Data Using Opinion Mining. 10.3233/APC200124.

Schneider J & Vlachos M. Topic modelling based on keywords and context. In: Proceedings of the 2018 SIAM International Conference on Data Mining. SIAM; 2018.

Shah, A., Yan, X., Tariq, S., & Shah, A. (2021). Tracking patients healthcare experiences during the COVID-19 outbreak: Topic modelling and sentiment analysis of doctor reviews. Journal of Engineering Research. 9. 10.36909/jer.v9i3A.8703.

Smith, H., & Cipolli, W. (2021). The Instagram/Facebook ban on graphic self-harm imagery: A sentiment analysis and topic modelling approach. Policy & Internet. 10.1002/poi3.272.

Syed, S. & Spruit, M. (2017). Full-Text or Abstract? Examining Topic Coherence Scores Using Latent Dirichlet Allocation. 165-174. 10.1109/DSAA.2017.61.

Tahmassebi, A., McCann, I., Meyer-Base, A., Erlebacher, G. & Foo, S. (2018). NewsAnalyticalToolkit: an online natural language processing platform to analyze news. 23. 10.1117/12.2304646.

Valdez D., ten Thij M., Bathina K., Rutter L., & Bollen J., Social Media Insights Into US Mental Health During the COVID-19 Pandemic: Longitudinal Analysis of Twitter Data J Med Internet Res 2020;22(12):e21418 URL: https://www.jmir.org/2020/12/e21418 DOI: 10.2196/21418

van der Veen, M., & Bleich, E. (2021). Atheism in US and UK Newspapers: Negativity about Non-Belief and Non-Believers. Religions. 12. 291. 10.3390/rel12050291.

Waheeb, S., Khan, N., & Shang, X. (2022). Topic Modelling and Sentiment Analysis of Online Education in the COVID-19 Era Using Social Networks Based Datasets. Electronics. 11. 715. 10.3390/electronics11050715.

Xu, Y., Li, Y., Liang, Y., & Cai, L. (2020). Topic-sentiment evolution over time: a manifold learning-based model for online news. Journal of Intelligent Information Systems. 55. 10.1007/s10844-019-00586-5.

Yadav, A., & Vishwakarma, D. (2021). A Language-independent Network to Analyze the Impact of COVID-19 on the World via Sentiment Analysis.

ACM Transactions on Internet Technology. 22. 1–30. 10.1145/3475867.

Yakunin, K., Kalimoldayev, M., Muhamedyev, R., Mussabayev, R., Barakhnin, V., Kuchin, Y. Murzakhmetov, S., Buldybayev, T., Ospanova, U., Yelis, M., Zhumabayev, A., Gopejenko, V., Meirambekkyzy, Z., & Abdurazakov, A. (2021). KazNewsDataset: Single Country Overall Digital Mass Media Publication Corpus. Data. 6. 31. 10.3390/data6030031.

Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, & V. Stoyanov, ``RoBERTa: A robustly optimized BERT pretraining approach," 2019, arXiv:1907.11692.

Yin, H., Song, X., Yang, S., Li, J. (2022). Sentiment analysis and topic modelling for COVID-19 vaccine discussions. World Wide Web. 25. 1-17. 10.1007/s11280-022-01029-y.

Zhang, C. (2021). Media Framing of Color-Blind Racism: A Content Analysis of the Charlottesville Rally*. Race and Social Problems. 13. 10.1007/s12552-021-09321-8.

Zheng, H., Goh, D., Lee, E., & Lee, C. S., & Theng, Y. (2022). Understanding the effects of message cues on COVID-19 information sharing on Twitter. Journal of the Association for Information Science and Technology. 73. 847-862. 10.1002/asi.24587.

Zhou, Pi., He, Y., Lyu, C., & Yang, X. (2020). Characterizing News Report of the Substandard Vaccine Case of Changchun Changsheng in China: A Text Mining Approach. Vaccines. 8. 691. 10.3390/vaccines8040691.

Zolnoori, M., Huang, M., Patten, C., Balls-Berry, J., Goudarzvand, S., Brockman, T., Pour, E., & Yao, L. (2019). Mining News Media for Understanding Public Health Concerns. Journal of Clinical and Translational Science. 5. 1-29. 10.1017/cts.2019.434.

Facilitating Smart Tourism Destination Development: Towards a Model of Information System for Smart Tourism

Ljiljana Zekanović-Korona, Jurica Grzunov
Department of Tourism and Communication Studies
University of Zadar
Street of dr. Franje Tudjmana 24i, Zadar
{ljkorona, jgrzunov}@unizd.hr

Abstract. *With the intensive development of ICT and rapidly growing outbound tourism markets, the concept of smart tourism has been mentioned increasingly over the past several years. However, this concept is still not well grounded in theory, which makes it difficult to comprehend and even more complicated to implement in practice. To better understand the concept and provide tangible mechanisms for implementing it, this paper addresses some of its key aspects and provides a theoretical framework and guidelines for the development of a general model of information system that would serve the development and management of a smart tourism destination.*

Keywords. Smart tourism, Smart tourism destination, Information system.

1 Introduction

It is unquestionable that the development of new information and communication technologies over the last 50 years has had a strong impact on leisure travel. With the development of the first computer reservation systems in the 1970s, and especially with the development of the Internet, modern ICTs were slowly becoming an integral part of each phase of the leisure travel cycle (Buhalis & Law, 2008, p. 609-610.). However, with the intensive development of mobile technologies in the 21st century, tourism is experiencing an unprecedented growth on a global scale. Namely, according to UNWTO statistics, from 1995 to 2019 and the outbreak of the SARS-CoV-2 coronavirus pandemic that brought the travel industry to a halt, the number of leisure travels in the world increased by more than 250% on an annual basis (UNWTO, 2019).

Some of the key roles of ICT in the context of leisure travel are providing fast and easy access to information regardless of the location and increasing the level of interactivity among all stakeholders in tourism. In addition, with respect to the aforementioned coronavirus pandemic, smart ICT solutions should play a vital role in reviving business processes once tourism enters recovery phase, especially in understanding the new patterns caused by the pandemics itself (Thong et al., 2022).

Tourism represents a very complex phenomenon, which is evident from one of the commonly used definitions of tourism set by AIEST in 1981, describing it as: "… the sum of the phenomena and relationships arising from the travel …" (Vanhove, 2012, p. 2). A massive increase in the number of travels resulted in some very negative consequences for many popular destinations. One of the most serious problems caused by this increase is overtourism, affecting mostly big urban destinations and often resulting in degradation of public spaces and cultural heritage, creation of monofunctional spaces, pollution caused by tourism, decline in local values and customs, conflicts between residents and tourists, rising prices and cost of living and over dependence on tourism (Nunkoo & Ramkissoon, 2010, p. 41-42). Most of these consequences are the result of poor destination management, most often caused by lack of knowledge, expertise and motivation among entities in charge of tourist destination development and management, but also lack of information about visitors and resources in the destination. Due to these and similar problems caused by the large increase in the number of travels, the future of tourism has been increasingly considered in the context of sustainable tourism development, enabed by sustainable and smart solutions. Sustainable tourism development can be defined as: "…the development that meets the needs of the present without compromising the ability of future generations to meet their own needs." (WCED, 1987, p. 43). The tools that can support such a development are smart technological solutions and innovations fueled by modern ICT, aimed at better information and resources management in order to increase the attractiveness and competitiveness of tourist

destinations. This is the context in which the phrase "smart tourism" has been increasingly mentioned over the last 15 years, as a reflection of the impact of technology on tourism. The use of modern technologies in the context of smart tourism destinations presupposes the networking of all stakeholders, which is usually realized through sensorization and the use of big data, as well as new ways of connecting and exchanging information (such as social media, Internet of Things, Near Field Communication, Radio Frequency Identification, etc.) (Gretzel et al., 2015a).

2 Smart Tourism Destinations

Although the basic idea on which smart tourism is based is relatively simple and solid, the very concept of smart tourism remains poorly grounded in theory and practice. This is evident from the rather small number of scientific papers addressing this concept. Some authors put special emphasis on the use of technology, while others are more focused on destination management, but all of them emphasize the use of (big) data, advanced tec1hnologies and smart systems as a *condicio sine qua non* when it comes to smart tourism. Many authors argue that papers on smart tourism still lack conceptual and empirical development and that researchers have not fully analyzed how destinations apply this concept, what initiatives are being implemented and what are the results of these initiatives.

According to Gretzel et al., smart tourism can be defined as: "…tourism supported by integrated efforts at a destination to collect and aggregate/harness data derived from physical infrastructure, social connections, government/organizational sources and human bodies/minds in combination with the use of advanced technologies to transform that data into on-site experiences and business value-propositions with a clear focus on efficiency, sustainability and experience enrichment" (Gretzel et al., 2015b, p. 181). This definition clearly emphasizes the importance of technologies, but brings to the fore data and how they are used. Also, Gretzel et al. state that smart tourism consists of three basic components: smart tourism destinations, smart business ecosystem and smart tourism experience, and that all of them reside on smart data collection, processing, and exchange (Gretzel et al., 2015b, p. 180.).

Smart tourism destination can be considered the most important component, as it represents a fundamental precondition for the development of other components. There are numerous definitions of a smart tourism destination, one of which was offered by López de Ávila Muñoz et al. in a report entitled "Smart Destinations Report: building the future" prepared by SEGITTUR (Spanish: Sociedad Estatal para la Gestión de la Innovación y las Tecnologías Turísticas), a state-owned company for tourism innovation and technologies management, defining a smart tourism destination as: "…an innovative tourist destination, built on an infrastructure of state-of-the-art technology guaranteeing the sustainable development of tourist areas, accessible to everyone, which facilitates the visitor's interaction with and integration into his or her surroundings, increases the quality of the experience at the destination, and improves residents' quality of life." (López de Ávila Muñoz et al., 2015, p. 32). The authors of this definition point out that the creation of smart tourism destinations requires the development of a strategic plan that considers the specifics of the destinations. Therefore, such a strategic plan requires an individualized approach for each destination and cannot be generalized, but there are still a number of general measures that such a plan should encompass, emphasizing in particular the encouragement of public-private partnerships, defining goals aimed at improving expected results, promoting innovations in business models, intensifying the use of advanced ICTs and implementation of systems for data collection and analysis and information distribution (i.e. communication) among all tourist destination stakeholders. (López de Ávila Muñoz et al., 2015)

This paper focuses on smart systems for data collection and analysis, as well as communication and distribution of information among all destination stakeholders. Given that a review of recent literature in the field of smart tourism shows that there are no scientific papers or related research approaching the analysis or implementation of such systems in a systematic and comprehensive manner, this paper proposes a framework for a general model of information system to support the development and management of smart tourism destinations.

3 Information Systems

Information systems are a complex combination of technology and human activity and, as such, can be defined in several ways. For example, one definition of information systems states that they represent "…a set of interrelated components that collect (or retrieve), process, store, and distribute information to support decision making and control in an organization" (Laudon, Laudon, & Elragal, 2012). In general, it could be argued that their primary purpose is to help people working with data in order to improve a particular collective work activity (Korpela et al., 2004). Olivé points out three main functions of the information system: memory function, informative function and active function. In other words, an information system records general system state representations, provides information about the system state and supports performing actions that change its state (Olivé, 2007, p. 2).

Many authors think of information systems as socio-technological systems. For instance, Watson

defines an information system as a socio-technological system consisting of a social and technological subsystem, where the technological subsystem includes the technological and process component, and the social subsystem includes the structural and human component. The technological component mainly refers to information technology used to collect, process, store and distribute data, while the process component refers to a set of steps/actions taken when performing a particular business or organizational activity. On the other hand, the social subsystem consists of human and structural components, where the human component includes all individuals directly related to the system and all their skills, interests, attitudes, prejudices, and the like, while the structural component refers to organizational structure, i.e., to the relations between individuals within the organization. (Watson, 2007)

Given this concept, the information system in this paper is also viewed primarily as a socio-technological system, so for the purpose of modeling the information system, a theoretical framework called Cultural-Historical Activity Theory was used. The Cultural-Historical Activity Theory (shorter: Activity Theory) serves as a framework to better describe and understand the structure, development and the context of human activity (Engeström, 2009). Activity theory has been used repeatedly in various research related to information system development, as it provides a well-built framework for analyzing the complex dynamic settings of the systems under study, which typically involve interaction with technical and human elements (Ditsa, 2003). One of the main advantages of using Activity Theory in information systems research is that it fuses technology (i.e. tools) and context into a unit of analysis, which is - activity. To be more specific, neither society nor technology is put in the foreground, but rather the activity, which is viewed from a socio-technological perspective. (Allen et al., 2013)

4 Methodology

The specific method that was used for the purposes of this paper is called the *Activity-Oriented Development Method* (abbr. AODM). It is a method based on the Activity Theory. The author of this method, Daisy Mwanza, designed AODM as a practical methodological framework to support system design processes from the perspective of Activity Theory. The method consists of four methodological tools: Eight-Step-Model, Activity Notation, Technique of Generating Research Questions and Technique of Mapping Operational Processes. The main purpose of AODM is facilitating data collection and analysis, as well as the transfer of acquired knowledge. The mentioned methodological tools can be used iteratively, through six phases of research: 1.Examining and interpreting activity systems in the context of activity theory, 2.Developing a general model of activity system, 3.Disassembling systems into components, 4.Generating research questions, 5.Conducting detailed testing of all system components, 6.Interpreting and communicating the obtained results. (Mwanza, 2002)

All of the abovementioned tools are motivated by Engeström's general model of Activity Theory, which describes the activity system as an object-oriented interaction between objects and subjects performing the activities and actions, communities involved and the division of work within them, and tools and rules applied in the transformation process, which all together direct the activity system towards the desired outcome. (Engeström, 1999)

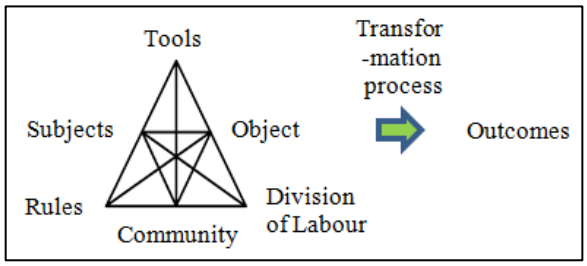


Figure 1. Activity Triangle Model (Engeström, 1987)

Specifically, the following objectives were derived using the AODM framework:

- To determine the different dimensions of a smart tourism destination and the possible indicators of their development (*Objects/Outcomes*).
- To describe communities involved in the development and management of smart tourism destinations and the division of labor within these communities in the context of smart tourism destination development (*Communities/Division of Labour*).
- To depict the important entities in the context of the development and management of a smart tourism destination, and the associated activities and actions (*Subjects/Division of Labour*).
- To identify rules and norms that affect the process of developing and managing a smart tourism destination (*Rules*).
- To identify tools (both material and abstract) which are used to carry out activities aimed at the development of smart tourism destinations, as well as their characteristics (*Tools*).
- To describe the central activity system of a smart tourism destination.

In order to meet all of the aforementioned objectives, an extensive literature review was conducted and the findings are described in the following section.

5 Findings

When it comes to determining the dimensions of a smart tourism destination, some of the related research has already cosidered possible ways of decomposing such complex ecosystems into simpler meaningful components (i.e. dimensions). In the year 2000, one of the most cited authors in the field, Dimitrios Buhalis, analyzed the components in the development of tourism destinations in general, and defined six dimensions that could describe successful tourism destinations. His framework was called the 6A, as it consisted of the following six dimensions: Attractions, Accesibility, Amenities, Available packages, Activities and Ancillary services (Buhalis, 2000, p. 98). In 2017, Tran, Huertas and Moreno expanded the 6A framework by adding smart componentes to all six dimensions, and they called it the (SA)6 framework. In addition to that, they also suggested a total of 57 indicators, divided into several subcategories, each of them assigned to a specific dimension. Namely, the dimensions they suggested were: Smart attractions (4 subcategories, a total of 7 indicators), Smart accessibility (3 subcategories, a total of 20 indicators), Smart amenities (3 subcategories, a total of 7 indicators), Smart ancillary services (8 subcategories, a total of 12 indicators), Smart activities (3 subcategories, a total of 5 indicators) and Smart packages (5 subcategories, a total of 6 indicators) (Tran, Huertas, & Moreno, 2017). In their later work, Tran, Huertas and Moreno further expanded their framework by assigning plausible scales and weigths to all of the indicators, making their framework more robust and applicable in real life (Tran, Huertas, & Moreno, 2019). Since all other models and frameworks suggested by different authors seem to be more abstract and less studious when compared to the (SA)6, this framework could serve as a good starting point in building an information system to service the development and management of different dimensions of a smart tourism destination.

The second objective was to describe communities involved in the development and management of smart tourism destinations and the division of labor within these communities in the context of smart tourism destination development. There are several ways in which communities, or different groups of stakeholders in tourism can be differentiated and categorized. For example, a number of authors have divided stakeholders into primary and secondary, with primary stakeholders defined as those who have some sort of a formal relationship with the destination and are vital in its functioning, and secondary are those that affect (or are affected by) the destination, but do not have to be formally related to it and are not necessary for its survival (Sheehan & Ritchie, 2005). In order to better describe our activity system, we will use the categorization of stakeholders proposed by Miller and Twining-Ward, who divided tourism stakeholders into four significant groups: Public sector (e.g local authorities, regional authorities, various levels of government responsible for tourism and its key assets, destination marketing/management organizations, etc.), Private sector (e.g. tour operators, travel agencies, accommodation, restaurants, guides, etc.), Non-governmental organizations (e.g. environmental protection groups, local communities, cultural groups, etc.) and finally, Tourists and organizations who represent them. (Miller & Twining-Ward, 2005, p. 183)

As for the entities involved in the activities associated to the development and management of a smart tourism destination (e.g. decision-makers), it can easily be argued that both primary and secondary stakeholders (or all groups) should be adequately represented in these activities in terms of subjects. Unfortunately, that is rarely the case in tourism, as not all stakeholders have equal power and influence in the decision-making process (Ruhanen, 2009). For example, politicians participating in the local or regional government (e.g. mayors, governors, ministers, etc.) typically have a significantly greater influence on decision-making than private sector stakeholders and various NGOs or tourists. Sigala and Marinidis emphasize the importance of cooperation between all stakeholders in the destination. They argue that by combining knowledge, expertise, capital resources and quality strategies, a synergy can be created among tourism stakeholders, which can lead to new opportunities, innovative solutions and a higher level of efficiency and competitiveness that would not be possible without such cooperation (Sigala & Marinidis, 2010).

The division of labour that corresponds with these groups of stakeholders and specific subjects within them can hardly be generalized in any way. In order to be able to address this problem, it is important to make a formal distinction (based on laws and policies) and a qualitative distinction (based on real-life scenarios) of all activities with regard to who is involved in their execution, and in relation to their impact on the outcome of the activity system (i.e. destination development and management). The formal distinction is conveyed in all applicable laws and policies that determine the jurisdiction and liabilities of all subjects that are relevant in the context of the development and management of a destination. As these regulations can vary significantly between two or more destinations, an individual approach is necessary to get the exact notion of who's who and who's responsible for what, where and when. On the other hand, the qualitative distinction (who actually does what and how does he do it) is possible to determine only through extensive qualitative research *in situ* (for example, interviews or focus groups with representatives of all stakeholder groups in a specific destination).

Correspondingly, every destination has rules and norms that define necessary interactions within the

activity system with respect to obligatory legal relationships between all important subjects. These rules and norms are also destination-specific and must be carefully studied and applied in the process of describing the activity system, and ultimately implemented in the information system model. Rules and norms encapsulate all state laws and regulations, acts, policies, development strategies, instructions, and guidelines concerning tourism and the development of tourism destinations in any way (for example, laws on tourist boards and the promotion of tourism, laws on tourist tax, environmental laws, GDPR, national/local tourism development strategies, etc.)

Finally, to complete the activity triangle model of the activity system present in a smart tourism destination, it is necessary to describe the tools used in and outside of the destination, not only by the visitors, but by all subjects in the destinations. Tools can be devided into material nad abstract. Material tools refer to all physical means used by subjects to perform different actions and operations (such as computers, smarthphones, sensors, etc.), whereas abstract tools include all intangible concepts (such as ideas or information) that affect the activity system. Furthermore, in doing so it is critical to keep in mind that the cycle of a tourist journey consists of three stages: pre-trip (inspiration, decision, purchase), during (purchase, experience, satisfaction) and post-trip (memory, satisfaction) (López de Ávila Muñoz et al., 2015), so it is very important to consider the tools that are used in other stages of travel outside of the destination (i.e. before and after). Last but not least, it is important not only to identify the tools used in the activity system, but also (if possible) to determine where and why they are used, as well as how and why they are developed.

Information and knowledge flows in the destination can be considered the most relevant mechanisms for the general behavior of the activity system. Productivity, innovation, and economic growth are strongly influenced by these flows, and the ways in which information and knowledge are disseminated can determine the speed with which individual stakeholders perform and plan their future activities within the destination. (Argote & Ingram, 2000) Technologies in general, with a special accent on information and communication technology, is probably what made the biggest impact on turism industry since the beginning of leisure travel. Obviously, a major paradigm shift in tourism occurred with the advent of the Internet, which enabled an unprecedented flow of information worldwide. In addition to the Internet, a variety of information systems also proved to be very important for the global tourism growth (such as Central Reservation Systems (CRS), Global Distribution Systems (GDS), Geographic Information Systems (GIS), Global Positioning Systems (GPS), Decision Support Systems (DSS), Management Information Systems (MIS), Recommendation Systems (RS), etc.). In the past 20 years technological advancements intensified significantly, with a special emphasis on mobile technology (such as smartphones, tablets, laptops, etc.), which enabled the travelers to be even more informed and perform transactions regardless of where they are. The advancements in mobile technology also influenced the development of some new intelligent systems (such as Virtual and Augmented Reality Systems, Location Based Services, Intelligent Transport Systems, etc.). Out of all tools that are often mentioned in scientific papers related to smart tourism, the Internet of Things (IoT) and big data are the most frequent ones. The International Telecommunication Union (ITU) defines IoT as "…a global infrastructure for the information society, enabling advanced services by connecting physical and virtual things based on interoperable information and communication technologies." (ITU, 2012, p. 1) Big data generally refers to the huge amount of data generated by people and various technologies, which is increasing exponentially on a daily basis, causing problems in the functioning of resources intended for storage, processing and analysis of such data. The fundamental value of big data lays in enabling the formation and management of knowledge. Bernabeu et al. identify the major dimensions that caraterize big data: volume (large amounts of data with a high update frequency), diversity (different types and structures of data and multiple channels from which they are collected) and speed (speed of data storage, access and analysis and the ability to display data in real time). (Bernabeu et al., 2016) López de Ávila Muñoz and Sánchez also point out other important characteristics of big data, such as their business value and authenticity, in a sense that they can be used to generates competitive advantages in different areas, and in terms of data reliability, i.e. the quality of information and its accuracy that enables better decision making. (López de Ávila Muñoz & Sánchez, 2013)

Furthermore, some of the most disruptive tools in tourism industry in the past decade have proven to be social media sites and apps (such as Facebook, YouTube, Instagram, TikTok, etc.). Using social media, every individual can be informed in a quick and easy way about the advantages and disadvantages of a tourist service or product before making a purchase decision, but can also inform others about their experience. It is noteworthy to point out that social media users generate huge amount of valuable data, which can be considered an asset for all stakeholders in terms of marketing and decision making.

Although there is an abundance of other tools that are used in tourism, it is beyond the scope of this paper to address all of them. However, it would be a major flop not to emphasize the importance of smartphones and mobile apps in the tourist industry.

This is especially true for the younger generations, who value the experience of getting to know the environment more than the quality of the accommodation and other traditional values in tourism, and such an experience is most accessible using smartphones and appropriate mobile apps. Grieve, Bendon and Hunsdon categorized apps with respect to the type of service they provide:

- apps for providing useful information
- apps related to culture
- transport-related apps
- specialized apps created for the needs of certain companies
- interactive electronic guides
- navigation apps
- accommodation booking apps
- social networks based on location services
- apps based on event information
- travel planning apps

Also, many applications are a combination of two or more of these types and provide several different services at the same time. (Grieve, Bendon, & Hunsdon, 2010)

All of the aforementioned technologies (and many more) combined with the laws, policies, development strategies and other legal acts, form an amalgam of tools that are used by different subjects in a tourism destination in order to achieve their objectives.

6 Towards a Model of a General Information System for Smart Tourism Destinations

One of the main goals of this paper was to set up a theoretical framework for the development of a general information system model that would enable collection, processing, storing and dissemination of tourism related information to support decision-making, coordination, control, analysis, and visualization of data in a smart tourism destination. Such an information system could greatly contribute to the main objectives of smart tourism destinations in general. In other words, it could support sustainable development of the destination, facilitate visitor's interaction and integration into the environment, improve the tourist experience and the quality of life of residents. The AODM theoretical framework based on the Activity Theory could prove to be adequate for analyzing such a complex system and its main components, as well as the way they interact with each other. Smart tourism destinations represent a complex ecosystem which can be devided into simpler components or dimensions. In relation to that, the aforementioned (SA)6 framework can be utilized to ease the development of such a system, as it provides a toolset and plausible metrics and indicators for the analysis of smart tourism destination development, along with all its components. The Activity Theory works well with such divisions, as it provides conceptual tools for analyzing and understanding a network of activity systems working together to form a more complex activity. (Engeström, 1999) It is also important to note that the Activity Theory recognizes the contradictions and tensions within the system of activities as a source of change in the activity itself, which corresponds well with tourism. (de Souza Bispo, 2016, p. 173)

As a part of authors previous research, it can be argued that the central activity system of a smart tourism destination can be represented as the sum of activity subsystems (one for each of the dimensions of the (SA)6 framework), with all the individual components (Tools, Subjects, Rules & Norms, Communities, Division of Labour) comprised as a union of corresponding components for each of the dimensions of the (SA)6, as shown on Figure 2. (Grzunov, 2020, p. 233-256)

The proposed outcome of this activity system can only be achieved through extensive cooperation of all stakeholders in the system, assuming the responsible actions of all communities with regard to the existent division of labor, in accordance with the applicable norms and rules, and providing appropriate investments in the development and ethical use of all necessary tools.

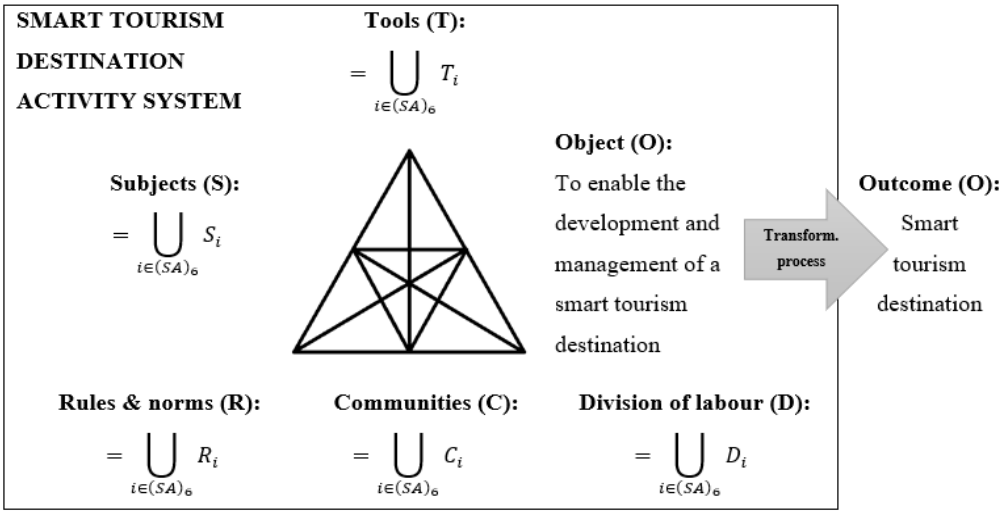


Figure 2. Central Activity System of a Smart Tourism Destination (source: author's work)

7 Conclusion

The proposed theoretical framework provides a good starting point for all further research associated to information systems that could contribute to the development and management of smart tourism destinations, and alleviate the tensions and problems caused by overtourism and poor management of the destinations, as well as the consequences of the recent pandemic. Due to the exceptional complexity of such systems, in terms of different user roles, communication methods, modalities of data collection, structuring and storage, this paper doesn't address any details regarding the entities and links between different system components, as well as specifications of user roles, interfaces, specific communication protocols and other specifics related to the implementation of such a system. Therefore, it is suggested that any future work on this subject should focus on the processes that take place within the system, as well as the exact data types and structures on which its work is based.

The suggested frameworks for analyzing smart tourism destination that were referenced in this paper (namely, the AODM and the (SA)6) can be seen as just one of the numerous possibilities, and by no means taken as a necessity or imperative in future research. Hopefully, the concept of smart tourism will strengthen over time and become a new paradigm in tourism.

References

Allen, D. K., et al. (2013). How Should Technology-Mediated Organizational Change Be Explained? A Comparison of the Contributions of Critical Realism and Activity Theory. *MIS Quarterly, 37 (3)*, p. 838.

Argote, L., & Ingram, P. (2000). Knowledge Transfer: A Basis for Competitive Advantage in Firms. *Organizational Behavior and Human Decision Processes 82 (1)*, p. 150–169.

Bernabeu, C., et al. (2016). Big Data and Smart Tourism Destinations: Challenges and opportunities from an industry perspective. *Proceedings of School of Hospitality and Tourism Management Conference*, (p. 1-17).

Buhalis, D. (2000). Marketing the Competitive Destination of the Future. *Tourism Management 21 (1)* , p. 97–116.

Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet - The state of eTourism research. *Tourism Management 29 (4)*, p. 609-610.

de Souza Bispo, M. (2016). Tourism as Practice. *Annals of Tourism Research 61*.

Ditsa, G. (2003). Activity Theory as a Theoretical Foundation for Information Systems Research. U G. Ditsa, *Information Management: Support Systems & Multimedia Technology* (p. 192-293).

Engeström, Y. (1987). Learning by Expanding–An Activity-Theoretical Approach to Developmental Research. *Doctoral thesis*. Helsinki: University of Helsinki.

Engeström, Y. (1999). Activity Theory and Individual and Social Transformation. U Y. Engeström, R. Miettinen, R.-L. Punamaki, Y. Engeström, R. Miettinen, & R.-L. Punamaki (Ed.), *Perspectives on Activity theory* (p. 19-38). Cambridge: Cambridge University Press.

Engeström, Y. (2009). The Future of Activity Theory: A Rough Draft. U A. Sannino, H. Daniels, & K. D. Gutierrez, *Learning and expanding with Activity theory* (p. 303–328). Cambridge: Cambridge University Press.

Gretzel, U., et al. (2015a). Smart Tourism Challenges. *Journal of Tourism 16 (1)*, p. 41-47.

Gretzel, U., et al. (2015b). Smart tourism: foundations and developments. *Electronic markets 25 (3)*, p. 179–188.

Grieve, C., Bendon, T., & Hunsdon, S. (2010). Our mobile future: how smartphones will transform visiting experiences. *Horizon Digital Economy Research*, p. 1-8.

Grzunov, J. (2020). A model of information system for smart tourism destinations support. Doctoral thesis. University of Zadar, Zadar.

ITU, International Telecommunication Union (2012). *Overview of the Internet of Things.* ITU.

Korpela, M., et al. (2004). Information Systems Research and Development by Activity Analysis and Development: Dead Horse or the Next Wave? (B. Kaplan, & e. al., Ed.) *Information Systems Research*, p. 453–471.

Laudon, K. C., Laudon, J. P., & Elragal, A. (2012). *Management Information Systems: Managing the Digital Firm (12th ed.).* Prentice Hall.

López de Ávila Muñoz, A., et al. (2015). *Smart Destinations Report: building the future.* Madrid: SEGITTUR.

López de Ávila Muñoz, A., & Sánchez, S. G. (2013). Destinos turísticos inteligentes. *Harvard Deusto business review 224*.

Miller, G., & Twining-Ward, L. (2005). *Monitoring for a Sustainable Tourism Transition: The Challenge of Developing and Using Indicators.* Wallingford, UK ; Cambridge, MA.

Mwanza, D. (2002). Towards an Activity-Oriented Design Method for HCI Research and Practice. The Open University.

Nunkoo, R., & Ramkissoon, H. (2010). Small Island Urban Tourism: A Residents' Perspective. *Current Issues in Tourism 13 (1)* , p. 41–42.

Olivé, A. (2007). *Conceptual Modeling of Information Systems.* Berlin; New York: Springer.

Ruhanen, L. (2009). Stakeholder Participation in Tourism Destination Planning Another Case of Missing the Point? *Tourism Recreation Research 34 (3)*, p. 283–294.

Sheehan, L. R., & Ritchie, J. B. (2005). Destination Stakeholders Exploring Identity and Salience. *Annals of Tourism Research 32 (3)*, p. 714.

Sigala, M., & Marinidis, D. (2010). DMOs, e-Democracy and Collaborative Destination Management: An Implementation Framework. (U. Gretzel, et al., Ed.) *Information and Communication Technologies in Tourism*.

Thong, C. L., et al. (2022). The Effectiveness of Smart Tourism in Malaysia in Covid-19 Post-pandemic Era: A Case Study. In: Salvendy, G., Wei, J. (eds) *Design, Operation and Evaluation of Mobile Communications.* HCII 2022. Lecture Notes in Computer Science, vol 13337. Springer, Cham., p. 60-73.

Tran, H. M., Huertas, A., & Moreno, A. (2017). (SA)6: A New Framework for the Analysis of Smart Tourism Destinations. A Comparative Case Study of Two Spanish Destinations. *Congresos - Seminario Destinos Turisticos Inteligentes 2017 - Libro de Actas*, p. 190-214.

Tran, H. M., Huertas, A., & Moreno, A. (2019). Which destination is smarter? application of the (Sa)6 framework to establish a ranking of smart tourist destinations. *International Journal of Information Systems and Tourism (IJIST) 4 (1)*, p. 19-28.

UNWTO, World Tourism Organization (01 2019). *International Tourist Arrivals Reach 1.4 Billion Two Years Ahead of Forecasts*. accessed on 21. 05 2019 from www2.unwto.org: https://www2.unwto.org/press-release/2019-01-21/international-tourist-arrivals-reach-14-billyon-two-years-ahead-forecasts

Vanhove, N. (2012). *The Economics of Tourism Destinations, 2nd ed.* Elsevier.

Watson, R. (2007). *Information Systems.* Georgia: University of Georgia.

WCED. (1987). *Our Common Future.* WCED.

Quality in ICT

Renata Mekovec, Ruben Picek, Sandra Lovrenčić, Marina Klačmer Čalopa
Creating jobs for future computer professionals

Creating jobs for future computer professionals

Renata Mekovec, Ruben Picek, Sandra Lovrenčić, Marina Klačmer Čalopa
Faculty of organization and informatics
University of Zagreb
Pavlinska 2, Varaždin
{renata.mekovec, ruben.picek, sandra.lovrencic, marina.klacmer}@foi.unizg.hr

Abstract. *The constant need for harmonization and improvement of competencies in changing labour market conditions leads to an increased need for standard definition/description for particular occupation. An occupational standard is a document that outlines the competencies required for a specific job. It is generated using a well-defined methodology and data collection. This article describes the process of developing occupational standards in four areas: 1) management of information security and privacy systems, 2) distributed and interactive systems, 3) management of business system transformation and innovation, and 4) data technology and artificial intelligence.*

Keywords. Occupational standards, future jobs, ICT

1 Introduction

Computer science trends like cloud computing, information security and big data collection and storage are creating new jobs and better career opportunities. Innovation, support for high-value industries, and a major reliance on continual technological improvement characterize the ICT sector. In order to create new employment, ICT experts will need to combine new digital talents with managerial skills and trends such as digital leadership, big data, cloud computing, the Internet of Things and smart systems.

Data science has become one of the essential components of the information science (Sonavane, 2021). It is a key area for future careers and the role of data science is emerging across the globe. People in today's world require a broader set of skills to work, interact, access data, goods, and services, and participate in social and civic activities (European Commission, Directorate-General for Employment, 2019). Digital skills on higher level during pandemic of corona virus become a prerequisite for many jobs. But still, there is a lack of digitally skilled people. The U.S. Bureau of Labor Statistics projects 11% growth for computer and information technology occupations from 2019 to 2029, a faster-than-average growth rate (Simmons, 2022). The forecast for 2020 was a deficit of more than 500.000 information and communication technology professionals in Europe (The Future of Work - Trends, Challenges and Potential Initiatives, 2021). IBM research data shows that data science accounts for 28% of all digital jobs by 2020, but also, they revealed that on average, each of these places remains unfilled for up to 45 days due to a lack of talent (Hopping & Curtis, 2018).

The McKinsey Global Institute research in 2020 shows that many firms in Europe are competing for well-educated and digitally fluent professionals (Smit et al., 2017). Still, a lot of research papers show that there is a lack of professionals with computer science degree in Europe comparing to USA. A report by PwC indicates that demand for data scientists is clearly growing, especially in less developed markets, and that there will be an additional 2.7 million jobs created in the data science and analytics industry in the period from 2020 to 2023 ("PwC's Global Artificial Intelligence Study: Sizing the Prize," 2022).

This is a reason why in some research papers practitioners and scientists strongly recommended data science curriculum also in high school. The importance of computer science to the national curriculum has been addressed by some states in USA with the aim that teachers who work with students become more aware of the professional opportunities available to their students upon graduation (Morreale & Joiner, 2011).

These data show that new skills and expertise will be required for future occupations, as well as a rise in demand for ICT specialists. As a result, there is a need for education and training that will produce adequate employees, meet labour market demand, and enhance ICT skill levels.

The goal of the study described in the article is to evaluate the labour market in order to determine the skills that employees in the field of information technology require. The following questions guided the research:

- Will workers dealing with information security and privacy systems management, distributed and interactive systems management, business system transformation and innovation, and data technology and artificial intelligence be required?

- What are the most important tasks that persons in these jobs should perform?
- What knowledge and skills should these professionals have?

This article describes the process of developing occupational standards according to the Methodology for developing occupational standards proposed by Croatian Ministry of science, education and sports.

2 European Qualifications Framework

The European Qualifications Framework (EQF) is a common European reference framework aimed at making qualifications easier to read and understand across countries and systems. The EQF's core are its eight reference levels, which are defined in terms of learning outcomes, namely: knowledge, skills, and autonomy-responsibility. Learning outcomes describe what people know, understand and can do at the end of a learning process (European Centre for the Development of Vocational Training, 2022).

All types and levels of qualifications are encompassed by the EQF. The level increases in accordance with the level of proficiency, with level 1 being the lowest and level 8 being the highest. The EQF was implemented in accordance with the Recommendation on the European Qualifications Framework for Lifelong Learning, which was adopted by the European Parliament and the Council in 2008 and revised in 2017.

The EQF has been a driving force behind the creation of comprehensive national qualification frameworks based on learning outcomes. All countries that have accepted the EQF presume that such national frameworks are required to make their qualifications comparable across sectors and countries. National qualification frameworks of 35 countries have been formally linked ('referenced') to the EQF by September 2021 (European Centre for the Development of Vocational Training, 2022).

Also, the EQF aims to improve qualification transparency, comparability, and portability. The framework acts as a translator between various qualification systems and their levels. It is meant to help students, workers, job applicants, companies, labour unions, education and training providers, qualification recognition organizations, government agencies, and international institutions (European Commission, Directorate-General for Employment, 2019). Qualifications serve a number of functions. They indicate to employers what their holders know and can do in theory ('learning outcomes'). They may be required for entry into certain regulated professions. They assist educational and training authorities and providers in determining an individual's level and content of learning. They are also significant for an individual as a symbol of personal achievement. As a result, qualifications play an important role in improving employability, mobility, and access to higher education (Council recommendation of 22 May 2017 on the European Qualifications Framework for lifelong learning and repealing the recommendation of the European Parliament and of the Council of 23 April 2008 on the establishment of the European Qualifications Framework for lifelong learning, 2017).

2.1 Croatian Qualifications Framework

The Croatian Qualifications Framework (CROQF) is a tool for regulating the qualification system in the Republic of Croatia (Ministry of Science and Education, 2022). The placement of Croatian qualifications at various levels not only determines their relationship, but it also allows for the linking of Croatian qualifications levels to the levels of the EQF and the levels of the European Higher Education Area Qualifications Framework. The CROQF lays the groundwork for the development of Recognition of Prior Learning (RPL), or the recognition and validation of non-formal (organized activity, but no public document) and informal (unorganized activity, everyday experience) learning (Zakon o Hrvatskom Kvalifikacijskom Okviru, 2013), as well as the mandatory implementation of a quality assurance system and clear quality assessment procedures.

CROQF creates the necessary quality assurance elements and instruments to serve as a long-term base for creating a suitable system at all levels, as well as for all types and forms of learning and education (Nove Boje Znanja - Strategija Obrazovanja, Znanosti i Tehnologije, 2015).

The Croatian Qualifications Framework Act establishes and governs the CROQF, as well as the qualifications system in the Republic of Croatia, by introducing standards qualifications based on occupational standards and consisting of sets of learning outcomes (Zakon o Hrvatskom Kvalifikacijskom Okviru, 2013). The Ordinance on CROQF Register specifies the process for proposing and evaluating occupational standards, qualifications, and sets of learning outcomes via sectoral councils (Ordinance on CROQF Register, 2014).

3 Methodology

Development of occupational standards according to CROQF is prescribed with Methodology for developing occupational standards and includes Occupational standard survey. In Methodology are described guidelines for the following activities related to the organization and implementation of the survey (Ministarstvo rada i mirovinskog sustava obitelji i socijalne politike, 2021):

- Adjustment and optimization of the survey
- Conducting the survey

- Systematization of collected information and data
- Formalization of occupational standard proposal

All activities are described in general in (Čalopa et al., 2020) and adjustment and optimization of the survey with a case study is also described in detail.

3.1. Conducting the survey

The survey that is adjusted to a specific occupation(s) should be used in guided survey and structured interview with future employers and current employees. The purpose is to obtain all relevant information for development of occupational standard(s) from their perspective. The following tasks will ensure the desired outcome:

- Selection of employers and employees for guided survey and structured interview
- Preparation of survey leader and assistant for guided survey and structured interview
- Conducting the guided survey and structured interview
- Systematization of information and data through preparation of report about conducting the survey
- Selection of employers and employees

The goal of the first task is to select the sample of employers and employees that will reflect majority opinion about knowledge, skills and competencies needed for occupational standard(s). For selection of employers are used principles of representativeness (number, size, business sector, nature of business activity and spatial distribution of employers), accessibility and willingness to cooperate. Principle of representativeness should be always used first, and other two only in rare cases at the expense of representativeness. Selection of the representative of each employer organization is also very important, and this should be a leader that has knowledge about all jobs of an occupation. When selecting employees, principle of representativeness should also be taken into consideration, as well as expertise of employee, and a sample should represent a population of employees in occupation(s). A number of employers and employees should be similar. It is expected that a total number of interviewees will be eight to twenty, depending on occupation.

3.1.1. Preparation of survey leader and assistant

Guided survey and structured interviews should be conducted in person and therefore education for survey leader and assistant is needed. They should have detailed knowledge about the structure and the contents of the survey, as well as about the goals of the process. Members of the group that adjusted and optimized the survey should participate in conducting, possibly as leaders. Leaders and assistants must understand the importance of confirmation and supplementation of information as well as detailed description of occupation(s).

3.1.2. Conducting the guided survey and structured interview

This task includes introducing employers and employees to tasks in the survey so that they can answer properly. Through guided survey are gathered opinions about elements developed by expert group and through structured interview is gathered new information. All terms in the survey (e.g. key task) must be explained and understood by employers and employees and they must estimate for each key task (a) the frequency of performance and (b) level of expertise needed. Information about revision of key tasks is also welcomed. All questions should be addressed during the survey. Some elements can be accepted, others rejected or modified. Survey leader and assistant should guide the survey to achieve its goals.

3.2. Systematization of collected information and data

The process of systematization begins already during conducting the survey with gathering notes. The goal is to develop lists of key tasks, knowledge and skills for occupational standard(s). Along with information obtained from employers and employees during the survey, for achieving this goal experts that adjusted and optimized the survey can be included. Two main tasks of systematization are:

- Development of key tasks list
- Development of knowledge and skills list

3.2.1. Key tasks list

After the analysis of gathered data, it can be established that some of the suggested key tasks are not performed at all in a specific occupation(s) – they can be removed from the list. For the tasks that are performed very rarely or rarely can also be suggested to be removed, and those that are performed often or very often are definitely the key tasks.

Considering the level of expertise, for often and very often performed tasks it should be analysed if those tasks for which basic level of expertise is needed are in fact those that are performed with the leadership or supervision of others. Tasks for which medium level of expertise is needed should be those mostly performed independently, and tasks for which advanced expertise of level is needed, are those that should show leadership and management of others. If those relationships are not visible, descriptions of key tasks should be adjusted. The same procedure is valid for key tasks that are not suggested in survey but are added during guided survey and structured interview.

3.2.2. Knowledge and skills list

Similar to key tasks, it should be firstly established if some suggested knowledge and skills are not needed for specific occupation(s) so that they can be removed from the list. Knowledge and skills that are necessary for specific occupation(s) are mainly those that are on

the list, but some optional knowledge and skills can also be included if they are really important for successful performance of a certain key task.

The level of expertise needed in knowledge and skills is also analysed. It is expected that basic level of expertise will be mostly needed for optional tasks. As with key tasks, adjustment of description of knowledge and skills may be needed and also the same procedure is valid for knowledge and skills that ae not suggested in survey but are added during guided survey and structured interview.

4 Results

The survey procedure was made for four areas for which the following names of occupational standards were proposed:

- Management of information security and privacy systems - Information Security and Privacy Architect
- Distributed and interactive systems - Distributed and / or Interactive Systems Architect
- Management of business system transformation and innovation - Business Systems Transformation Architect
- Data technology and artificial intelligence - Intelligent Information Systems Architect

After the survey, the analysis of data began, where the first step was to define the key tasks of a particular standard of occupation. As described, during the survey, respondents were asked for each key job to state the frequency of performance of each task and the level of expertise required. Based on the total answers of the respondents, each key task could be divided into three categories: candidate for the standard of occupation, candidate for expulsion and the need to discuss. After the analyses and consultations with experts in the field, key tasks were defined:

1. Information Security and Privacy Architect - 17 key tasks
2. Distributed and / or Interactive Systems Architect - 13 key tasks
3. Business Systems Transformation Architect - 16 key tasks
4. Intelligent Information Systems Architect - 14 key tasks

Here is example of defined key tasks for Information Security and Privacy Architect:

1. Planning the deployment and monitoring of information security and privacy systems, as well as the resources required to implement and monitor them
2. Creating an organizational structure for integrating data security and privacy systems into the business system
3. Organizing teamwork of information security and privacy
4. Planning and developing competencies in the domain of data security and privacy
5. Analysing the current situation's (in) compliance with the information security and privacy system's requirements
6. Identifying and assessing the risks associated with data assets
7. Risk management planning for the information system's operation (in crisis conditions)
8. Examining the information system's integrity in terms of security and privacy
9. Managing roles and authorities to ensure that information resources are used safely
10. Reporting to the Management Board on the business system's security and privacy status
11. Participation in sales and/or procurement activities with security and privacy concerns
12. Involvement of stakeholders in the development of information security and privacy systems
13. Collaboration with the company's personnel
14. Communicating with regulatory bodies and exchanging information with businesses and professional organizations
15. In the development, improvement, or innovation of products or services, collaborating to ensure that security and privacy standards are met
16. In quality assurance, defining security and privacy indicators
17. Raising awareness of the importance of data security and privacy in terms of health and the environment

There were two approaches to summarizing knowledge and skills. The first method refers to a situation in which a framework (international or national Body of Knowledge (BoK), occupational standard) with specified categories existed for a specific field. The card sorting method was used in this situation. The second approach was to define 3 types of knowledge/skills in the absence of comparative standards or frameworks (presented in Figure 1):

a. Specific – innovative knowledge/skills in relation to the new occupational standard – clustering method was used in order to define this type of knowledge/skills
b. Informatics - cards sorting method was used with categories and subcategories based on the European e-Competence Framework (*The E-Competence Framework - ITPE,* 2022)
c. Generic (Key/Long Learning) - cards sorting method was used with categories based on the European key competency framework for lifelong learning (Council Recommendation of 22 May 2018 on Key Competences for Lifelong Learning, 2018).

Occupational standard/Sets of competencies	Specific	Informatics	Key /Long Learning
Information Security and Privacy Architect	Information Systems and Business Strategy Alignment Service Level Management Product/ Service Planning Incident and problem management Information Security Strategy Development Information security risk management Data security management	Modeling, reengineering and improving business processes Personal and professional development in the IT profession Monitoring technological trends and innovations in informatics Professional and socially responsible behavior in the IT profession Cooperation and communication in informatics	Citizenship competence Entrepreneurial competence
Distributed and/or Interactive Systems Architect	Integration of distributed and/or interactive system components Distributed and/or interactive system engineering Architectural design of distributed and/or interactive system Distributed and/or interactive system design Distributed and/or interactive system support and upgrade Distributed and/or interactive system development Service developmnet in the field of distributed and / or interactive system for third parties	Design of technical and communication infrastructure and information system platform Professional and socially responsible behavior in the IT profession Monitoring technological trends and innovations in informatics Software development Cooperation and communication in informatics Development and project Management in informatics	Mathematical competence and competence in science, technology and engineering
Business Systems Transformation Architect	Information Systems and Business Strategy Alignment Business Plan Development Technology Trend Monitoring Documentation Production System Engineering User Support System upgrade support Purchasing Personnel Development Project and Portfolio Management Risk Management Business Change Management	Software development Monitoring technological trends and innovations in informatics	Entrepreneurial competence
Intelligent Information Systems Architect	Effective application of methods and techniques for the intelligent information systems development based on statistical approach to artificial intelligence Life cycle management of intelligent information systems with the application of the principles of information and software engineering Development, implementation and management of solutions for storage and processing of data, information and knowledge using distributed and cloud platforms Design, develop and manage databases and knowledge bases and provide efficient ETL, OLTP and OLAP processes for large data sets Integration of heterogeneous data from different sources and their preparation for further analysis and use Application of methods and techniques for the intelligent information systems development based on a symbolic approach to artificial intelligence Programming for artificial intelligence and data science Business application of artificial intelligence and data science methods	Software development Monitoring technological trends and innovations in informatics Professional and socially responsible behavior in the IT profession	Mathematical competence and competence in science, technology and engineering Personal and social competence and competence to learn how to learn

Figure 1. Defined knowledge/skills for occupational standards

5 Conclusion

Standardization of occupations and qualifications at national levels and linking them with international standards is important, because it improves their quality, allows mutual recognition and enhances world labour market. Those standards should follow the evolution of the areas for which they are developed so that they can ensure the education of professionals who can meet the requirements of the chosen occupation.

The proposed occupational standards for four areas listed above are an important step in monitoring trends such as:

(1) a growth in the number of security threats and attacks (e.g. stalking, cyberbullying, information warfare, industrial espionage, identity theft attacks),

(2) a rise in the number of monitoring sensors in the environment and intelligent networking of machines and other devices (with the help of the industrial Internet of Things, cloud computing),

(3) improvement of the sophistication of technology that supports business processes, autonomous decision making, and real-time asset and process monitoring (virtual and augmented reality, artificial intelligence, advanced robotics, autonomous robots, simulation, digitization and system integration), and

(4) an increasing number and types of data collected and analysed (machine learning methods and techniques, in-depth data analysis, technologies for integration, storage and processing of data, recognition of behavioural patterns)

The recognition of these specialists in the Republic of Croatia, as well as in Europe, has been ensured thanks to the Methodology of the Croatian Qualifications Framework. In other words, the duties, knowledge, and abilities that ICT professionals should possess were defined, independent of which faculty they attended.

Acknowledgments

This material is based on data from project DIP2Future: Development of educational programs, qualification standards and occupational standards in the field of ICT in accordance with the CROQF financed from the Operational Programme Efficient Human Resources (OPEHR), European Social Fund (ESF). Authors want to express gratitude to the DIP2Future project partners and their staff for the provided data.

The opinions, findings, and conclusion or recommendations expressed in the paper are those of the authors and do not necessarily reflect the view of the DIP2Future project of the University of Zagreb, Faculty of Organization and Informatics.

References

Council recommendation of 22 May 2018 on key competences for lifelong learning, (2018). https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32018H0604(01)&from=EN

Council recommendation of 22 May 2017 on the European Qualifications Framework for lifelong learning and repealing the recommendation of the European Parliament and of the Council of 23 April 2008 on the establishment of the European Qualifications Framework for lifelong learning, C 189 Official Journal of the European Union 189/15 (2017). https://ec.europa.eu/ploteus/sites/eac-eqf/files/en.pdf

Čalopa, M. K., Lovrenčić, S., Mekovec, R., & Picek, R. (2020). The Occupational Standards Development According to CROQF*. Central European Conference on Information and Intelligent Systems, 73–80. https://www.proquest.com/conference-papers-proceedings/occupational-standards-development-according/docview/2531366478/se-2?accountid=17242

European Centre for the Development of Vocational Training. (2022). European qualifications framework (EQF). https://www.cedefop.europa.eu/en/projects/european-qualifications-framework-eqf

European Commission, Directorate-General for Employment, S. A. and I. (2019). The European Qualifications Framework: supporting learning, work and cross-border mobility 10 th Anniversary. https://doi.org/10.2767/750617

Hopping, C., & Curtis, J. (2018). Data scientist jobs: Where does the big data talent gap lie? | IT PRO. ITPro.Co.Uk. https://www.itpro.co.uk/careers/28929/data-scientist-jobs-where-does-the-big-data-talent-gap-lie

Ministarstvo rada i mirovinskog sustava obitelji i socijalne politike. (2021). Metodologija za izradu standarda zanimanja i skupova kompetencija.

Ministry of Science and Education. (2022). About CROQF - CQF. http://www.kvalifikacije.hr/en/about-croqf

Morreale, P., & Joiner, D. (2011). Reaching Future Computer Scientists. Communications of the ACM, 54(4). https://doi.org/10.1145/1924421.1924448

Nove boje znanja - Strategija obrazovanja, znanosti i tehnologije, (2015).

Ordinance on CROQF register, (2014). https://www.azvo.hr/images/stories/novosti/Pravilnik_o_Registru_HKO-a.pdf

PwC's global Artificial Intelligence study: sizing the prize. (2022). In PwC. https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html

Simmons, L. (2022). Top 6 Computer Science Trends in 2021. https://www.computerscience.org/resources/computer-science-trends/

Smit, S., Tacke, T., Lund, S., Manyika, J., & Thiel, L. (2017). The future of work in Europe. In The Future of Work in Europe. https://doi.org/10.4324/9781351146609

Sonavane, A. K. K. (2021). Study of Emerging Role of Data Science in Business Intelligence. Design Engineering, 6.

The e-Competence Framework - ITPE. (2022). https://itprofessionalism.org/about-it-professionalism/competences/the-e-competence-framework/

The future of work - Trends, challenges and potential initiatives, 12 (2021).

Zakon o Hrvatskom kvalifikacijskom okviru, (2013). http://www.kvalifikacije.hr/sites/default/files/documents-publications/2021-03/ZHKO_2021_pročišćeni.pdf

Strategic Planning and Decision Making

Ivana Dvorski Lacković
Learning analytics implementation in higher education institutions: Enterprise risk management approach

Maja Gligora Marković, Tena Jagačić, Nikola Kadoić
Analyzing Decision-Making Styles: Army vs. Business students in Croatia

Merve Güler, Esin Mukul, Gülçin Büyüközkan
Digital Transformation Technology Evaluation in a post COVID-19 World with an Integrated Fuzzy SAW-ARAS Methodology

Esin Mukul, Merve Güler, Gülçin Büyüközkan
Evaluation of Supply Chain Tools with Fuzzy AHP-COPRAS Methodology for Intelligent Supply Chain Management

Jan Fauser
Business model development and business ecosystem evaluation in the energy domain

Learning analytics implementation in higher education institutions: Enterprise risk management approach

Ivana Dvorski Lacković
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, Varaždin, Croatia
idvorski@foi.unizg.hr

Abstract. *Higher education institutions (HEIs) that are implementing learning analytics (LA) are faced with many open questions related to strategic, operational, ethical, pedagogical and privacy issues. LA implementation in HEI requires strategic decision-making, but since LA is a relatively new and interdisciplinary field, the process of reaching quality strategic decisions may be somewhat difficult due to lack of clearly defined decision-making criteria. In this paper we present enterprise risk management (ERM) approach to LA implementation in HEI. We transpose COSO (2017) risk management framework on the case of LA implementation in HEI. The purpose of the paper is to present how risk management approach may be beneficial for HEI in setting clear aims related to LA implementation, defining strategy and measuring achievement of the set goals. The process is expected to lead to enhanced value for stakeholders related to LA implementation in HEI.*

Keywords. Learning analytics, higher education institutions, enterprise risk management, COSO (2017) framework

1 Introduction

The roles of higher education institutions (HEIs) have traditionally been defined through three core activities: research, education and service to the society (Oosterlinck, 2004). The societal change and technological progress led to the expansion of these traditional roles. Hayter and Cahoy (2016) draw from the existing literature that HEIs' roles also include contribution to the economic development and commercialization through incubators, science parks and spin-offs, as well as contribution to the societal sustainability. This is in line with Guerrero et al. (2016) who perceive HEIs as focal points in the development of three types of capital: human, knowledge and entrepreneurial.

Although digitalization, driven by external processes (policy) and internal processes (leadership and staff development), was already present in some HEIs (Tømte et al., 2019), the COVID-19 pandemics led to the widespread adoption of digitalisation in HEIs. One of the aspects related to digitalisation is the employment of learning analytics (LA). Interest in LA is growing both in researchers' and practitioners' community, but having on mind that this field is new and interdisciplinary, it requires cooperation among different experts in HEI that are included in its implementation.

Some of the most common issues related to LA implementation are related to strategic, operational, ethical, pedagogical and privacy issues. HEI's management is thus faced with the necessity to make many strategic decisions related to LA implementation, as well as its usage and continuous improvement, once LA is implemented. One of the main issues decision-makers in organizations are faced with are risks related to the introduction of a new processes. Thus, HEI's management is faced with the challenge how to approach risks connected to LA implementation, usage and improvement.

The main aim of this paper is to approach LA implementation in HEIs from enterprise risk management (ERM) perspective. We follow the guidelines provided by the standardized risk management framework COSO (2017), but we translate them in the context of LA implementation in HEIs. Having on mind that ERM benefits organizations in terms of enhanced value, derived from adequate risk management process for all the stakeholders, we argue that ERM approach to LA implementation in HEIs may be beneficial for all the involved parties and lead to improved strategic decision-making.

Although common in business practice, ERM approach is not widely discussed in the scientific literature in the context of HEI decision-making. Lundqvist (2015) argues that ERM adoption at the universities in the United States of America began at year 2002 and that ERM approach is very useful in managing all the risks universities are exposed to and reaching strategic decisions. On the other hand, there is no literature corpus related to exploration of ERM in

HEI outside of the USA geographical context. This presents impetus for our paper. Since LA implementation is a strategic process that requires significant resources from the HEIs point of view, we aim to present a conceptual paper how ERM may be used in the process of LA implementation in HEIs. The base for using this concept are its widespread usage in large companies, public companies and not-for-profit companies (AICPA, 2021).

The paper consists of five parts. Following the introductory notes, we present a short overview of LA in HEIs. Third section of the paper is related to ERM and COSO (2017) overview. In fourth section we present how COSO (2017) may be translated in HEIs' context on the case of LA implementation. Finally, we provide conclusion remarks and references.

2 Learning analytics in higher education institutions

HEIs around the globe are undergoing the process of digital transformation, accelerated by the COVID-19 pandemics. One of prominent areas of HEIs' digital transformation is related to LA implementation. LA may be defined as "*measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and environments in which it occurs*" (SOLAR, 2011).

The strategic relevance of LA in education has been recognized and accentuated by UNESCO (2021), OECD (2021) and European Commission (2018; 2021). According to OECD (2021), LA provides opportunities for improvements in educational organizations' strategic management. While EU Commission (2018) is dedicated to improving education by data analysis and prediction, the same governing body emphasizes the need to develop ethical guidelines on data usage in teaching and learning (EU Commission, 2021). UNESCO (2021) prioritizes the need to strengthen the capacity of big data usage in education, develop standards and ethical guidelines related to technology implementation in education and provide evidence how technology impacts teaching and learning processes.

Schumacher and Ifenthaler (2018) argue that HEIs implement LA to increase their understanding of how students learn and support them in the process. Namely, LA enables personalized learning environment that may enhance students' learning. But it is not only students who benefit LA implementation. LA provides benefits for other HEI stakeholders as well. These stakeholders may be divided in four levels: mega-level (governance), macro-level (institutional level), meso-level (curricular level or teacher level) and micro-level (learner) (Ifenthaler and Widanapathirana, 2014; in: Schumacher and Ifenthaler, 2018). In the context of our paper, this means that each stakeholder group has its own interest in pursuing LA implementation, thus resulting in higher risk exposures through the process. This view is supported by Shum and Luckin (2019) who state that when using LA and artificial intelligence for educational purposes, "*we need the educators to be talking with developers and builders, and the joint narrative must speak to public policy*".

According to Joksimović, Kovanović and Dawson (2019), LA roles include predictive analytics, social learning analytics, discourse analytics and learning design. Main focus of using LA for predictive purpose is understanding and optimising learning. Some examples include identification of students at risk of dropout, predicting students' academic performance and retention rates (Joksimović, Kovanović and Dawson, 2019). According to Ifenthaler and Widanapathirana (2014) benefits of LA for learners include understanding learning habits, analysing learning outcomes, tracking learner's progress towards set aims, comparing learning paths, receiving automated interventions, supporting collaboration, taking assessments that include just-in-time feedback, optimizing learning paths, increasing students' engagement and success rate. The social LA role is dedicated to understanding how students create social relationships with their peers and teachers. Discourse LA explores students' communication, i.e. it focuses on "using textual discourse data for supporting student learning" (Joksimović, Kovanović and Dawson, 2019). In relation to learning design, early predictive learning models did not consider specific learning contexts, while newest research tendencies go in direction of exploring how LA can be used in various learning settings to improve learning design and enhance learning (Joksimović, Kovanović and Dawson, 2019).

Guzmán-Valenzuela et al. (2021) argue that main challenges for LA arise from:

I) *non-participation of students and teachers in active LA development*: While it would be expected that students and teachers are involved in the process of LA implementation and discussion about its effects, there is evidence from the field that LA implementation, data gathering and analysis are under control of a centralised unit, without clear connection to students and teachers. This challenge implies lack of organizational culture related to LA implementation, thus leading to the risk of possible misunderstandings and underusage of LA potential.

II) *students' learning process*: Main risk identified in this field is related to the fact that LA may be used as a source of information on students' success and prediction of possible learning problems. Although valuable in its core, this LA role may lead to the exact opposite of its intention, i.e. labelling students with problems and leading to self-fulfilling prophecy regarding the expected success rates. We may perceive this challenge as a pedagogical and ethical risk related to LA implementation.

III) *LA' impact on students' outcome achievements*: there is a lack of evidence on true LA impact on students' learning achievements. In the concurrent literature there is evidence that LA benefits researchers and administrators, but the effects on teachers and students are not well documented. This means that there is a risk that LA aims are not fully defined, understood or possibly adequately measured, thus leading to potential misunderstandings and overestimated benefits of LA.
IV) *methodologies of LA data interpretation*: there is a need that protocols are developed how LA data can be used within HEI's different timeframes and contexts.
V) *privacy issues* present one of major concerns in relation to LA implementation because it may be perceived as a source of permanent observation and limitation of students' freedom. Thus, collection, management, storage and data usage present potentially risky situation from the perspective of the institution that manages LA connected data.

According to EDUCASE (2016), major challenges related to LA inclusion in educational practice is connected to data-quality, difficulties with system integration, lack of support from key leadership of the institution and possible resistance of organisational staff to LA implementation.

Among the existing research related to LA implementation we point out Ferguson et al. (2014), who present how a structured approach Rapid Outcome Mapping Approach, may be used in order to overcome barriers that exist to LA implementation. This model includes definition of a clear set of policy objectives, mapping the context of LA implementation, identification of key stakeholders, identification of LA purposes, development of implementation strategy, analysis of capacities and development of human resources and development of a monitoring and learning system. The model clearly defines main steps that are very useful prior to LA implementation and tracking of implementation, i.e. it is focused on answering strategic questions that precede LA implementation. What is missing in this model is a more detailed connection how these strategic issues will be translated into HEI practice.

On the other hand, according to Sheikh et al. (2022), despite of LA' popularity, many HEIs have failed in the achievement of their previously defined strategic goals through LA support. According to these authors LA research is predominantly focused on tactical issues, i.e. how to implement LA on operational level, while the effects of LA on HEI's value are unclear, thus providing rationale for research how LA implementation can be used in order to achieve strategic goals of HEI. Since primary aim of ERM is to increase the likelihood that strategic objectives are realized and stakeholders' value enhanced (Dvorski Lacković et al., 2021), ERM approach is in its core focused on complying risk management activities with achievement of strategic goals of the organization. Thus, we find an argument that using ERM approach to LA implementation may be beneficial for aligning risk management with strategic aims set for LA implementation.

We may conclude that although LA implementation is an emerging and interdisciplinary field (Blackmon and Moore, 2020; Phillips and Ozogul, 2020; Divjak, 2021) that raised interest among researchers and practitioners, it is in its infancy (Viberg et al., 2018) and requires thorough research that will clarify LA' interaction with different strategic, operational, ethical, pedagogical and privacy aspects. This implicitly means that there are many risks related to LA implementation. These risks have been tackled only partially and until now there are no all-encompassing solutions that would include strategic, operational and oversight aspects of LA implementation at the same time. Thus, in continuation of this paper we approach LA implementation and usage from enterprise risk management perspective.

3 Enterprise Risk Management

Enterprise Risk Management (ERM) is a strategic approach to managing risks in companies. It is based on the premise that all the risks organization is exposed to should be managed holistically. This requires the centralization of risk managers role, regular identification and assessment of all the risks organization is exposed to, assessment of interconnection of different risk types, communication on risks throughout the organization and using risk management process insights for strategic decision-making. The purpose of using this approach is increasing the value for organization's stakeholders. The corporate sector practices ERM by managing strategic, operational and oversight aspects of risks management (Dvorski Lacković et al., 2021).

When implementing ERM, organizations have the freedom to use a tailor-made approach that suits their resources. Standardized frameworks COSO (2017) and ISO 31000 (2018) may offer guidance on the path of ERM implementation by providing an all-encompassing and structured approach to ERM process. The application of ERM approach on the universities is quite common in the United States of America (Lundquist, 2015), but we are not aware of its application on HEIs outside of this geographical context.

The main aim of this paper is to apply ERM on a specific process in HEIs – LA implementation. In order to reach our aim, we will use COSO (2017) standardized framework. It will serve as a basis point in exploring how LA implementation in HEIs may be supported through a standardized structured approach that manages risks holistically.

COSO (2017) is a set of principles that may be used in organizations of different sizes and sectors. It consists of five components: (1) governance and culture, (2) strategy and objective-setting, (3)

performance, (4) review and revision and (5) information, communication and reporting (Figure I). COSO (2017) components consist of a total of 20 principles. Namely, *Governance and Culture* component encompasses: a) board risk oversight, b) establishment of operating structures, c) definition of desired culture, d) demonstration of commitment to core values and e) attraction, development and retention of capable individuals. *Strategy and Objective-setting* component includes: a) analysis of business context, b) definition of risk appetite, c) evaluation of alternative strategies and d) formulation of business objectives. Under the component *Performance* following principles are encouraged: a) risk identification, b) assessment of risk severity, c) risk prioritization, d) implementation of risk responses and e) development of portfolio view. Component *Review and Revision* includes: a) Assessment of substantial changes, b) reviews of risk and performance, c) continuous improvement in risk management. Finally, *Information, Communication and Reporting* component encompasses: a) leveraging information and communication technology, b) communicating risk information and c) reporting on risk, culture and performance.

The main characteristics of COSO (2017) are related to the fact that importance of strategy is clearly accentuated in all the phases of ERM process, strategic and operational risk are integrated and holistically assessed in order to measure performance and achieve strategic aims, and iterative approach to risk management is practiced, thus leading to continuous monitoring of the strategic aims (Pierce and Goldstein, 2018).



Figure I. COSO visual. *Source: COSO (2017)*

4 ERM approach to LA implementation in HEIs

In this section we explore how COSO (2017) may be used in order to enhance LA implementation in HEIs. As argued in the previous Section, COSO (2017) framework consists of five main components and 20 principles divided onto these components. In order to analyse how COSO (2017) framework may be transposed on the process of LA implementation in HEI, in this section we go through detail of every COSO (2017) component and underlying principles in terms of specific HEIs' context and the process of LA implementation.

4.1 Governance and Culture

The first COSO (2017) principle *Governance and Culture* is strongly related to organization's mission, vision and core values. This means that HEIs should define or re-examine its mission, vision and core values in general and analyse how LA implementation fits into these realities. According to Keefe (2020), mission in HEI may be regarded as "*the lens through which the organization views relationship with students, educators, academia and other stakeholders, including local and global community*". More precisely, HEI should analyse its relation to all of its key stakeholders in the context of LA implementation. Some questions of concern related to LA implementation in relation to HEI's mission include: What is the purpose of LA implementation in relation to our students/teachers/ other staff/local community/other key stakeholders? What are the benefits we aim to achieve by using LA in the learning process for each of the stakeholders? What is the value we wish to deliver to our key stakeholders?

HEI's vision should clearly define how the institution will achieve its mission, i.e. it should state desired future and strategic plan how to achieve it (MacLeod, 2016, in: Keefe, 2020). In the context of LA implementation, a vision should be a statement of where the institution sees itself in the context of LA in the upcoming period and main manners how it aims to achieve it. For instance, what are our main aims related to LA in the next five years? How do we plan to achieve these aims?

The definition of HEI's mission and vision is the responsibility of the HEI's Management. This requires an adequate level of Management's oversight of all the risks HEI is exposed to and establishment of clear operating procedures in order to manage these risks. Specifically, in the context of LA implementation, HEI

should have clear written procedures related to both strategic and operational aspects of LA. In order for the desired outcomes of LA implementation are achieved, the HEI Management should aim to define the desired culture that supports and promotes the desired behaviours related to LA, but also demonstrate the commitment to these values through written policies and behavioural guidelines. The organizational values are beliefs about socially or personally desirable end states or actions that are explicitly or implicitly shared by members of organization (Schwartz, 1992, in: Mueller and Straatmann, 2014).

In HEIs, it is crucial that each employee is introduced to HEI's values so that LA practitioners are well-aware of how LA connects to values of their institution. By knowing values related to LA, each employee is encouraged to practice acceptable and ethical behaviour, thus complying with the set aims and enhancing the process of LA implementation. As discussed in the second section of this paper, LA implementation carries a certain amount of concern related to the ethics and privacy (Guzmán-Valenzuela et al., 2021), thus it is highly recommendable that detailed mapping is conducted how LA implementation may contribute to practicing desirable behaviour.

Once HEI's top management achieved a consensus on its mission, vision and core values in relation to LA, it posed a fertile ground on which it can develop sound ERM governance and culture. This first step enables HEI to practice the principles stated by COSO (2017). If HEI's management analysed what LA implementation means in relation to its mission, vision and core values, it has a clear oversight of LA implementation. Further on, this means that HEI established the general structure how LA will be implemented and how it supports HEI's commitment to what matters – its core values. Also, HEI's management defines the desired organizational culture and analyses how LA supports this culture. It is a step that may be specifically related to ethical and pedagogical issues of LA. Main question for HEI's management in this step is: How will we develop such an organizational culture that promotes the ethical use of LA and supports the pedagogical outcomes? For instance, a debate has been raised in the United Kingdom related to the problem of an algorithmic bias in grading the students during the COVID-19 pandemic (Smith, 2020). Such issues pose ethical risk for students who are under impact of such practices, but also reputational risk for educational institutions that practice biased grading approach.

Finally, the step that is of crucial importance is that HEI's management identifies capable individuals in different fields, as LA implementation requires an interdisciplinary approach, and motivates them to work on LA implementation. It is exactly these individuals and their interdisciplinarity that may contribute to successful LA implementation that is in line with HEIs strategy, but that also supports pedagogical and ethical issues.

4.2 Strategy and Objective-setting

The second COSO (2017) principle *Strategy and Objective-setting* relates to the development of a specific LA strategy. The first step for HEI to develop a high-quality LA strategy is based in the analysis of the business context. This means analysing HEI's macro-environment, namely all the demographic, geographical, political, economic, technological and social factors that impact HEI's processes, by tackling how these factors interwind with LA. It is important that business context analysis encompasses both external and internal factors that are specific for LA in HEI. These factors may be analysed by using SWOT matrix in order to identify the strengths and weaknesses coming out of HEI, as well as opportunities and threats arising from HEI's surrounding.

In business settings, the usual step following the business context analysis is risk appetite definition. A thoroughly conducted analysis enhances the process of formulating specific business objectives related to LA implementation and defining risk appetite, i.e. answering the question: What is the amount of risk connected to LA implementation that HEI is willing to take over? Risk appetite represents the amount of risk an organization is willing to take over in order to achieve a certain result. In relation to LA this would mean answering the question which risks and in which amount the HEI is willing to take over in order to implement LA. For instance, a simple example may be that HEI may define that it is willing to bear a certain additional financial amount (cost) of LA implementation in order to manage privacy risks related to LA implementation, such as data leakage.

Once the risk appetite is defined, possible alternative strategies may be evaluated and analysed to see how they match business aims achievement. For example, HEI may want to formulate alternative solutions to data leakage protection and analyse the optimal solution. Based on the analysed scenarios, HEI may formulate its clear business objectives that match its mission and vision in terms of LA implementation. Based on clear formulation of business aims, HEI may plan its resources necessary for LA implementation and match the exact amount of resources to a specified aim, but also aggregating all the necessary resources and manage them holistically, not partially.

Further on, a consistent and clear definition of business aims enables performance tracking and measurement. In the context of LA implementation this means that HEI may have a clear picture what is expected of LA implementation and define the metrics to measure the performance of LA, thus answering to impetus that methodologies of LA data interpretation are missing in current research (Guzmán-Valenzuela et al., 2021).

4.3 Performance

The third COSO (2017) principle *Performance* is oriented on risks associated to LA. It is based on the

risk identification. This means that HEI should make a list of all the risks that may occur in the process of LA implementation, usage and improvement. An initial point for making a risk list is a literature analysis: conducting a search through the scientific databases to identify research that was focused on LA associated risks. Further on, experts may be consulted for a further risk list complementation. What is very important in this step is that not only risks associated to LA implementation are considered, but that various risks that may occur in the advanced stages of LA usage in HEI, are considered.

Once the risk list is set, risk assessment is expected to take place in HEI. One of the main problems related to risk assessments is who should be responsible for conducting it. In business practice, an employee in charge for risk management conducts risk assessment based on inputs received from various employees. It would be recommendable that HEI's management appoints an employee in charge for risk management. Also, HEI's management should consider very carefully and target the employees that have the most knowledge in various technical, pedagogical and ethical issues related to LA and consult them to provide an adequate input for risk assessment.

The risk assessment should be conducted based on two parameters: the probability that specific risk will occur and impact that it may have on HEI. It is important to mention that the corporate sector usually assesses impact that risk carries for the company in the financial terms. Having on mind particularities of HEIs, not only financial impact should be considered, but HEI should analyse the impact of LA associated risks according to different estimated parameters. These parameters should be comparable so that HEI's management can prioritize the assessed risks.

For each risk that has been identified and assessed, a specific risk response should be formulated. Risk response should answer the question how identified risk will be treated, a time framework should be set for specific measures that will be undertaken and each risk should be connected to a specific risk owner, i.e. the person in charge to implement identified risk response.

This approach enables HEI's management oversight over LA related risks and how they are being handled. A very important point to consider when analysing risks and defining responses, is connected to developing a portfolio view. This means that each identified risk is assessed in relation to the other risks that may occur and it is carefully analysed whether there is a spill-over of certain risk among different HEI's organizational units.

4.4 Review and Revision

The fourth COSO (2017) principle is *Review and Revision*. It is connected to the post-implementation and performance impact review. Namely, each substantial change that occurs related to LA process should be carefully assessed from the risk management perspective.

Further on, prior identified risks should be monitored through responsibilities and time framework set while implementing risk responses. Also, the impact of these risks on performance should be the subject of a continuous review. In case there is a possibility to improve LA in relation to the risk managerial aspects, they should be pursued.

4.5 Information, Communication and Reporting

The fifth COSO (2017) principle, *Information, Communication and Reporting*, encompasses principles that are oriented on using ICT in the process of risk management. In our case this would mean leveraging all the possibilities ICT offers in extracting information about LA implementation and its usage once it is present in the company. It also implies communicating information about risk.

In HEI a clear path should be defined and preferably described within internal procedures about risk communication, i.e. who is in charge for communicating regular and outstanding information about LA related risks and in what manner (orally, written) and to whom. Finally, reports on risk should be a common practice and HEI's management should receive regular reports about LA related risks from the person in charge for reporting.

Since LA implementation in HEIs is a quite interdisciplinary field, it requires cooperation of different stakeholders. Thus, a person in charge for coordination and systematization of the received information on risks and reporting should be assigned. Altogether, the described COSO (2017) principles lead to enhanced value, meaning that in the analysed context HEI and all of its stakeholders should receive more value from the process of LA implementation.

5 Conclusion

In this paper we argue that there are many open issues when it comes to LA implementation in HEIs, both on the strategic and operational level. In order for HEIs' decision-makers to reach adequate decisions related to LA implementation and usage, a structured approach is required. We propose that LA implementation in HEIs is supported by using ERM approach. The main expected outcome of this approach is delivering more value to all LA stakeholders.

We use a standardized risk management framework COSO (2017) as a base to translate different risk management aspects on LA implementation in HEIs. The main motivation for this paper is that HEIs implementing LA use this all-encompassing approach in order to identify, set and achieve its strategic aims related to LA implementation.

Main impetus for this paper stems from the fact that HEIs are exposed to growing number of risks. LA implementation in its core carries various risk factors that may expose HEI to more risky profile. Generally, the research related to ERM in HEI is not widespread outside of the USA practice. Since ERM proved its usefulness in both profitable and non-for-profit business sector, we transpose its principles to the process of LA implementation in HEI.

This paper is primarily conceptual in nature. Both ERM and COSO (2017) framework are validated and widely used in practice. But we are aware that using already validated approaches in new settings requires additional validation. Main limitation of our paper is related to the fact that the presented concept should be empirically tested. Therefore, we propose that the future research deepens the topic by exploring LA experts' opinions on risks HEIs face in the process of implementation. Also, concrete case studies with evidence how HEIs benefit from using the suggested approach in the process of LA implementation, are highly welcome.

Acknowledgments

This work has been fully supported by the Croatian Science Foundation under the project IP-2020-02-5071.

References

American Institute of Certified Public Accountants (2021). *The State of Risk Oversight. An Overview of Enterprise Risk Management Practices*. Enterprise Risk Management Intitiative.

Blackmon, S.J., Moore, R.L. (2020). A Framework to Support Interdisciplinary Engagement with Learning Analytics. In: Ifenthaler, D., Gibson, D. (eds). *Adoption of Data Analytics in Higher Education Learning and Teaching. Advances in Analytics for Learning and Teaching*. Springer, Cham.

Committee of Sponsoring Organizations of the Treadway Commission (COSO). (2017). *Enterprise Risk Management Integrating with Strategy and Performance. Executive Summary*. New York: American Institute of Certified Public Accountants.

Divjak, B. (2021). Learning analytics – multi-user and multi-level perspective. *12th International Conference on eLearning 2021*, Belgrade Metropolitan University, 17-21.

Dvorski Lacković, I., Kurnoga, N., Miloš Sprčić, D. (2021). Three-factor model of Enterprise Risk Management implementation: exploratory study of non-financial companies. *Risk Management*, Online first, 22

EDUCASE (2016). *Learning Analytics in Higher Education*. Research report.

European Commission (2021), Digital education action plan 2021-2027, https://education.ec.europa.eu/sites/default/files/document-library-docs/deap-communication-sept2020_en.pdf (February 10, 2022)

European Commission (2018), Digital education action plan, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:22:FIN (February 11, 2022)

Ferguson, R., MacFayden, L.P., Clow, D., Tynan, B., Alexander, S., Dawson, S. (2014). Setting Learning Analytics in Context: Overcoming the Barriers to Large Scale Adoption. *Journal of Learning Analytics*, 1 (3), pp. 120-144.

Guerrero, M., Urbano, D., Fayolle, A., Klofsten, M., Mian, S. (2016). Entrepreneurial universities: emerging models in in the new social and economic landscape. *Small Business Economy* (47), pp. 551-563.

Guzman-Valenzuela, C., Gomez-Gonzalez, C., Tagle, A., Lorca-Vyhmeister, A. (2021). Learning analytics in higher education: a preponderance of analytics but very little learning? *International Journal of Educational Technology in Higher Education*, 18:23

Hayter, C.S., Cahoy, D.R. (2016). Toward a strategic view of higher education social responsibilities: A dynamic capabilities approach. *Strategic Organization*, 16 (1), pp. 12-34.

Ifenthaler, D., & Widanapathirana, C. (2014). Development and Validation of a Learning Analytics Framework: Two Case Studies Using Support Vector Machines. *Technology, Knowledge and Learning*, 19(1-2), 221-240.

ISO - International Organisation of Standardisation (2018). *A Risk Practitioners Guide to ISO 31000:2018*. https://www.theirm.org/media/3513119/IRM-Report-ISO-31000-2018-v3.pdf.

Joksimović, S., Kovanović, V., Dawson, S. (2019). The Journey of Learning Analytics. *HERDSA Review of Higher Education*, Vol. 6

Keefe, T. (2020). *The Mission, Vision, and Values in Higher Education: A Study of 39 Colleges of Art and Design.* 10.13140/RG.2.2.32784.43528.

Lundquist, A. E. (2015). Enterprise Risk Management (ERM) at U.S. Colleges and Universities: Administration Processes Regarding the Adoption, Implementation and Integration of ERM. Doctoral dissertation. Western Michigan University.

MacLeod, L. (2016). Mission, vision and values statements: The physician leader's role. *Physician Leadership Journal*, 3 (5), pp. 18-25.

Mueller, K., Straatmann, T. (2014). Organizational Values. *Encyclopedia of Quality of Life and Well-Being Research.*

OECD (2021), *OECD Digital Education Outlook 2021. Pushing the frontiers with AI, blockchain and robots*. https://read.oecd-ilibrary.org/education/oecd-digital-education-outlook-2021_589b283f-en#page5 (February 10, 2022)

Oosterlinck, A. (2004). The modern University and its main activities. Chapter 9 in: Weber, L., Duderstadt, J. (Eds), (2004). *Reinventing the Research University*, London: Economica.

Phillips, T., Ozogul, G. (2020). Learning Analytics Research in Relation to Educational Technology: Capturing Learning Analytics Contribution with Bibliometric Analysis. *TechTrends*, 64, 878-886.

Pierce, E.M., Goldstein, J. (2018). ERM and strategic planning: a change in paradigm. *International Journal of Disclosure and Governance*, Vol. 15 (1), str. 51-59.

Schumacher, C, Ifenthaler, D. (2018). Features students really expect from learning analytics. *Computers in Human Behaviour*, 7, pp. 397-407.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology, 25*, 1–65.

Sheikh, R.A., Bhatia, S., Metre, S.G., Faqihi, A.Y.A. (2022). Strategic value realization framework from learning analytics: a practical approach. *Journal of Applied Research in Higher Education*, Vol. 14 (2), pp. 693-713.

Shum, S. J. B., & Luckin, R. (2019). Learning analytics and AI: Politics, Pedagogy and Practices. *British Journal of Educational Technology*, 50(6), 2785-2793.

Smith, H. (2020). Algorithmic bias: should students pay the price?. *AI & society*, 35(4), 1077-1078.

Society for Learning Analytics Research (2011), *What is Learning Analytics*, Available at: https://www.solaresearch.org/about/what-is-learning-analytics/ Approached: 10.05.2022.

Tømte, C.E., Fossland, T., Aamodt, P.O., Degn, L. (2019). Digitalisation in higher education: mapping institutional approaches for teaching and learning. *Quality in Higher Education*, Vol. 25 (1), pp. 98-114.

UNESCO (2021), *UNESCO Strategy on Technological Innovation in Education (2022-2025).* https://unesdoc.unesco.org/ark:/48223/pf0000378847 (February 13, 2022)

Viberg, O., Hatakka, M., Bälter, O., Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behaviour*, Vol. 89, pp. 98-110.

Analyzing Decision-Making Styles: Army vs. Business students in Croatia

Maja Gligora Marković
Faculty of Medicine
University of Rijeka
Braće Branchetta 20, Rijeka
majagm@medri.uniri.hr

Tena Jagačić, Nikola Kadoić
Faculty of organization and informatics
University of Zagreb
Pavlinska 2, Varaždin
{tjagacic, nkadoic}@foi.unizg.hr

Abstract. *This research is a continuation of previous research in which two instruments for analysing the individuals' decision-making styles are applied to the population of students. The first is the Rowe instrument (used to identify the dominant decision-making style among analytical, behavioral, conceptual, and directive). The second is Scott & Bruce instrument (used to identify the dominant decision-making style among rational, intuitive, spontaneous, dependent, and avoidant styles). Now again, both were applied to a new sample. However, in both cases, we talk about the army and business students from the same faculties. The first sample was collected in 2020 (before COVID-19), and the second was collected in 2022 (after the COVID-19). We used descriptive statistics, t-tests with one-way ANOVA, and the χ^2 test in the data analysis. We analyzed the two samples separately and made the analysis on the joint sample, and finally, we analyzed some subsets. Significant differences were not identified in many cases, but not all. The distribution of 2022 students per Rowe's dominant decision styles differs significantly from that of students in 2020. The distribution of 2022 army students per Bruce & Scott's dominant decision styles is significantly different from the distribution of 2020 army students. There are significant differences in dominant decision-making styles using Scott & Bruce's instrument in students with respect to high school education and type of students.*

Keywords. Decision-making styles, army students, business students, Rowe, Scott & Bruce, instrument

1 Introduction

How people think to make decisions is considered a decision-making style (Scott & Bruce, 1995). There are several approaches to analyzing the decision-making styles. Some approaches are related to the number of people included in the process of decision making as a criterion for differing different styles. Here we differ in basic styles: autocratic and democratic (Ali, 1993; Dyczkowska & Dyczkowski, 2018; Lührs et al., 2018; Sikavica et al., 2014). The second approach uses the way of thinking and/or tolerance of uncertainty as a criterion for differing different styles (Abdelsalam et al., 2013; Robbins et al., 2016; Scott & Bruce, 1995; Thunholm, 2004).

Analyzing decision-making styles has been an ongoing research topic for decades. In the situation characterized by the COVID-19 virus, analyzing the topic can have additional importance since the strict measures could affect people's thinking and the way they make decisions. This research is not related to analyzing if some populations changed their way of decision-making in the post-COVID period compared to the before-COVID period. However, we will analyze decision-making styles concerning the time component.

In our previous research conducted in 2020 (before the COVID-19), we analyzed the differences in decision-making styles in the population of students in the fields of informatics, business, army, and navy (Kostanjevac et al., 2021). Now, we bring the results of the analysis of the decision-making styles in the student population of students in the fields of business and the army. The sample is different, but still, we can compare the differences between those two samples. Additionally, we present the analysis of the decision-making styles results of those two samples separately and with respect to the demographic characteristics of students. Finally, we give the results of decision-making styles analysis on two subsamples, those with respect to the gender of the students.

The paper is organized as follows. In Section 2, we present two decision-making style approaches (Rowe/Robbins and Scot & Bruce) used in the research. In Section 3, we give a brief overview of other authors' research related to applying decision-making styles in different contexts, including student populations. Further, in Section 4, we present the methodology of the paper and research questions. In Section 5, we present the results of the research. Finally, in Section 6, we discuss future research.

2 Decision-making styles by Rowe and by Scott and Bruce

2.1 The Decision Style Inventory by Rowe and Boulgarides

According to Rowe and Boulgarides, The Decision Style Inventory is an instrument that distinguishes persons by their mode of decision-making and tolerance for ambiguity, grouping them into four styles - analytical, conceptual, behavioral, and directive styles. A pre-designed questionnaire should be filled in by persons in different positions and further analyzed to identify their dominant decision-making style. Based on the questionnaire, it is identifiable how the person approaches problem-solving, how he interacts with other organization members, and how the individual's approach is harmonized with the dominant decision-making style (A. J. Rowe & Boulgarides, 1983).

The analytical style is characteristic of those who deal with complex situations, and think logically, solve problems well because solutions are based on real and verified information (Al-Omari, 2013; Boulgarides & Cohen, 2001; Connor & Becker, 2003; A. J., Rowe & Davis, 1996; A. J., Rowe & Mason, 1987). They have a high tolerance for ambiguity and are kept in written reports (A. J. Rowe & Boulgarides, 1983). This style is positively correlated with Allport Theoretical Value because of the empirical, critical, and rational approach (Boulgarides, 1984).

Further, the conceptual style uses a creative person-oriented toward people and solving social problems. It is led by intuition, makes long-term decisions decentralized, and is ready to risk. They are often considered idealists who take care of values and ethics (Al-Omari, 2013; Boulgarides & Cohen, 2001; Connor & Becker, 2003; A. J., Rowe & Davis, 1996; A. J., Rowe & Mason, 1987). They have high cognitive complexity and negatively correlate with Economic and Aesthesic Values (Boulgarides, 1984).

The preference for behavioral style is shared by people who are focused on other people. When making decisions, they listen actively, accept suggestions, are not inclined to conflict, and involve other people in the decision-making process (Rowe & Boulgarides, 1983). They are oriented towards short-term decisions because making difficult decisions creates problems and discomfort for them (Al-Omari, 2013; Boulgarides & Cohen, 2001; Connor & Becker, 2003; A. J., Rowe & Davis, 1996; A. J., Rowe & Mason, 1987). The behavioral style is negatively correlated with theoretical and political values and positively with social and religious values (Boulgarides, 1984).

The directive style is used by a person who is inclined to direct other people and very focused on work to achieve results as quickly as possible (Boulgarides, 1984). Such persons often act authoritatively and uncompromisingly and decide quickly using a small amount of information. Therefore, they aim to make short-term decisions (A. J. Rowe & Boulgarides, 1983). They are characterized by low tolerance to ambiguity and low cognitive complexity and are oriented towards structure and information by verbal means (Al-Omari, 2013b; Boulgarides & Cohen, 2001; Connor & Becker, 2003; A. J. , Rowe & Davis, 1996; A. J. , Rowe & Mason, 1987). The Directive style positively correlates with Allport Economic and Political Values recognizable by materialism, striving for results, and power (Boulgarides, 1984).

2.2 The General Decision–Making Style by Scott and Bruce

The decision-making style can be specified based on the General Decision-Making Style by Scott and Bruce. The General Decision-Making Style was created as an upgrade of the proposed models by Harren (1979) and Phillips, Pazienza, and Ferrin (1984). It distinguishes between five decision styles: rational, intuitive, dependent, avoidant, and spontaneous (Scott & Bruce, 1995).

Rational style is used by persons who approach decision-making logically and structured. Decisions relate deliberately, and the decision-maker assumes personal responsibility.

Unlike a rational style, intuitive style relies on intuition, impressions, and feelings. Great attention is paid to detail in decision making, but it makes a decision relatively quickly (Gambetti et al., 2008; Scott & Bruce, 1995).

In the dependent style, before making a final decision, a person consults and seeks advice from other persons. It can be seen from the above that decision-makers do not have much confidence in themselves when solving problems and making important decisions (Scott & Bruce, 1995; Urieta et al., 2022).

The avoidant style is used by the person who tries to delay the decision by avoiding or withdrawing (Motvaseli & Lotfizadeh, 2016; Scott & Bruce, 1995).

A person with a dominant spontaneous style is characterized by making sudden and unplanned decisions (Scott & Bruce, 1995; Weerasekara & Bhanugopan, 2022).

3 Previous applications of Rowe and Scott & Bruce instruments

In addition to authors whose expertise was in business decision making and decision-making styles, various researchers used Rowe's and Scott & Bruce's instruments to recognize dominant decision-making styles. In most new research, researchers tried to find a correlation between decision-making styles and

other variables. The application of presented instruments was often conducted on the student population and in business, and some of the researches are listed below.

Palmiero and associates researched the relationship between verbal abd visual divergent thinking and students' dominant decision-making styles. The 186 students of psychology at L'Aquila University participated in the survey. The assessment of the decision-making styles was carried out based on the Italian version of the GDMS test. The results of the tests confirmed the hypothesis that the rational decision-making style plays the most important role in divergent reasoning, while the intuitive style has little significance (Palmiero et al., 2020).

The assessment of personality's influence on emotional intelligence (E.I.) and decision-making styles was carried out by researchers on 296 medical students at Lebanese University. The GDMS test was used to identify decision-making styles, and the results suggest that the E.I. has a positive effect on the intuitive style and a negative impact on the avoiding and dependent style (el Othman et al., 2020).

The relationship between decision-making styles identified by the GDMS test and cognitive styles determined by Motvaseli and Lotfizadeh was analyzed in a survey conducted on 162 Iranian students. It was found that cognitive style has a positive impact on decision-making styles and that there is a need to develop cognitive style and introduce it into the educational system (Motvaseli & Lotfizadeh, 2016).

Urieta and associates examined the relationship between personality, decision styles, and the frequency of problematic smartphone use (PSU). There were 1,562 participants, out of which 556 were students. The results showed that avoiding, dependent and spontaneous styles had a positive relationship with PSU. On the other hand, there is negative relationship between PSU and of rational and intuitive styles (Urieta et al., 2022).

The research from Indonesia with the goal of determining the relationship between the management accounting system (MAS) and the decision-making style of Cooperatives' managerial performance was conducted in Ponorogo on 54 cooperative managers. The conclusions of the study show that "MAS did not affect managerial performance, and the decision-making style had a positive effect on managerial performance" (Efendi & Kusuma, 2021).

The following study, which was applied to 194 participants, aimed to determine the impact of the default rules in making decisions on financial investments. The default rules can improve financial decisions in individuals whose dominance is rational, avoiding, and dependent styles (Gambetti et al., 2022).

Based on the Decision Style Questionnaire, another survey was conducted and found that organizational commitment and satisfaction with the job are linked to the "decision-making styles of deputy vice-chancellors in Malaysian public Universities" (Khairunneezam Mohd Noor, 2020).

The same instrument was applied in studying the relationship between leadership styles and decision-making styles with 108 school principals from the Russaif Education District in Jordan. The results revealed that there is "no significant correlation between decision making styles and leadership styles of school principals" (Al-Omari, 2013a).

4 The methodology

The research in this paper was conducted among undergraduate students from two fields (and cities in Croatia): army students from Zagreb and business students from Varaždin.

The data were collected using a survey that included questions from two instruments (Rowe instrument and Scott and Bruce instrument) that followed the several general questions related to the demographic data (gender, study year, institution, the type of high school education). The first dataset was collected in 2020, and the second was in 2022.

We differ seven datasets in this research:

- DS1: joint 2020&2022 dataset,
- DS2: 2020 dataset,
- DS3: 2022 dataset
- DS4: Male subset of DS1
- DS5: Female subset of DS1
- DS6: Army subset of DS1
- DS7: Business subset of DS1

The collected data were analyzed using the M.S. Excel and Medcalc, and methods that were applied included the descriptive statistics, t-test, ANOVA, and χ^2 test.

There are several research questions in this research:

1. Is there a difference in achieved results in decision-making style types by Scott & Bruce with respect to demographic data in the first five datasets?
2. Is there a difference in achieved results in decision-making style types by Rowe with respect to demographic data in the first five datasets?
3. Is there a difference in the distribution of dominant decision-making style types by Scott & Bruce with respect to demographic data in the first five datasets?
4. Is there a difference in the distribution of dominant decision-making style types by Rowe with respect to demographic data in the first five datasets?
5. Is there a difference in achieved results in decision-making style types by Scott & Bruce with respect to the year in DS1, DS6, and DS7?

6. Is there a difference in achieved results in decision-making style types by Rowe with respect to the year in DS1, DS6, and DS7?
7. Is there a difference in the distribution of dominant decision-making style types Scott & Bruce with respect to the year in DS1, DS6, and DS7?
8. Is there a difference in the distribution of dominant decision-making style types by Rowe with respect to the year in DS1, DS6, and DS7?

5 Results and discussion

Now we bring the answers to the research questions.

5.1 Demographic data about the respondents

The demographic data about the respondents are presented in Table 1.

Table 1. Demographic analysis about the students

Demographic characteristics	Year	Values	Nr. of students	%
Gender	2020	Male	47	34,30
		Female	90	65,69
	2022	Male	35	28,00
		Female	90	72,00
Age	2020	20	26	18,97
		21	67	48,90
		22	30	21,89
		23	9	6,56
		24	4	2,91
		29	1	0,72
	2022	20	5	4,00
		21	64	51,20
		22	39	31,20
		23	15	12,00
		24	2	1,60
Year of study	2020	2nd	30	21,89
		3rd	107	78,10
	2022	2nd	5	4,00
		3rd	120	96,00
High school education	2020	vocational	65	47,44
		grammar school	72	52,55
	2022	vocational	69	55,20
		grammar school	56	44,80
Type of student	2020	army	65	47,44
		business	72	52,55
	2022	army	40	32,00
		business	85	68,00

There are 262 students in total: in DS1, 137 in DS2, 125 in DS3, 82 in DS4, 180 in DS5, 105 in DS6, and 157 in DS7. Most of the students in DS1, DS2 and DS3 are female students and 22 years old from the third year of study. In the 2020 dataset, more students graduated from the grammar school program, and in 2022, more students graduated from the vocational high school program.

5.2 Research question 1

Several t-tests with one-way ANOVA were conducted to respond to the first research question: input variables were related to the demographic characteristics, and output variables were scores achieved by respondents in the Scott & Bruce test. The methods were applied to five datasets. The full results for DS1 are presented in Table 2. The columns R, I, A, D, and S are related to the information about achieved averages of scores in the Scott & Bruce test for respected decision-making style (R = rational, I = intuitive, A = avoidant, D = dependent, S= spondaneous). Rows *p* are t-test and one-way ANOVA test scores.

Table 2. Scott & Bruce styles analysis (DS1)

DC	Values	R	I	D	A	S
Gender	Male	19,36	18,86	16,21	12,32	14,94
	Female	19,98	19,51	18,12	13,10	14,67
	p	***0,04***	*0,07*	***0,00***	*0,23*	*0,46*
Age	20	19,45	18,77	16,65	11,03	14,48
	21	19,79	19,67	17,50	12,78	15,18
	22	20,10	18,72	17,83	13,17	13,74
	23	19,71	20,00	18,33	14,21	15,63
	24	20,67	18,17	16,00	13,50	13,50
	29	21,00	16,00	11,00	11,00	11,00
	p	*0,81*	*0,05*	*0,30*	*0,20*	*0,05*
Year of study	2nd	19,20	18,74	16,66	11,03	15,14
	3rd	19,87	19,38	17,63	13,13	14,70
	p	*0,20*	*0,19*	*0,19*	***0,01***	*0,48*
High school education	vocational	20,21	19,68	17,69	12,46	14,34
	grammar school	19,30	18,87	17,30	13,20	15,16
	p	***0,01***	***0,01***	*0,44*	*0,20*	*0,06*
Type of student	army	19,60	19,28	16,64	11,53	15,32
	business	19,90	19,31	18,08	13,73	14,38
	p	*0,41*	*0,92*	***0,01***	***0,00***	***0,03***

From the summary table for the first research question of t-tests and one-way ANOVA implemented, we can draw several conclusions for DS1:

- There is no significant difference in scores achieved in Scott & Bruce test between groups based on demographic characteristics in most cases.
- There are significant differences in rational and dependent styles between male and female

students. Female students more often apply both rational and avoidant styles.

- There is a significant difference in dependent style with respect to age.
- There are significant differences in rational and intuitive styles between students who graduated from vocational school and grammar school. Students who graduated from vocational school more often apply rational and intuitive styles.
- There are significant differences between army and business students in avoidant, dependent, and spontaneous styles. While army students more often apply the spontaneous style, business students often apply the avoidant and dependent style.

Due to the limitation in the number of pages, we will not present full results like Table 1 for other datasets. Here, we bring the most interesting conclusions:

- There is no significant difference in scores achieved in Scott & Bruce test between groups based on demographic characteristics in other datasets in most cases.
- DS2: Female students more often apply a dependent style than male students. Third-year students more often apply avoidant style than second-year students. Similarly, business students more often apply an avoidant style than army students.
- DS3: Female students more often apply a dependent style than male students. Business students more often apply a dependent style than army students. Students who finished vocational study programs more often apply rational and intuitive styles than students who finished grammar school.
- DS4: Male business students more often apply avoidant style than male army students.
- DS5: Female business students more often apply avoidant style than female business students, while female business students more often apply spontaneous style than female army students.

5.3 Research question 2

Several t-tests with one-way ANOVA were conducted to respond to the second research question. In Table 3, we bring a summary results table related to DS1.

Here are the conclusions:

- Female students more often apply behavioral style than male students.
- There are significant differences in analytical style with respect to gender, age, year of study, and type of students.

When analyzing other datasets, we brought the following conclusions (A = analytical, B = behavioural, C = conceptual, D = directive):

- DS2: No significant differences detected.
- DS3: Male students more often apply the analytic style, while female students apply the conceptual and behavioral styles.
- DS4: Male 3rd-year students more often apply directive style than male 2nd-year students.
- DS5: No significant differences detected.

Table 3. Rowe styles analysis (DS1)

DC	Values	D	A	C	B
Gender	Male	75,11	83,15	74,13	67,61
	Female	73,10	77,44	75,70	73,77
	p	*0,30*	***0,00***	*0,38*	***0,00***
Age	20	70,29	84,42	77,61	67,68
	21	74,07	80,03	74,38	71,52
	22	76,17	76,00	75,09	72,74
	23	70,63	75,92	76,42	77,04
	24	70,50	83,50	75,50	70,50
	29	66,00	106,00	82,00	46,00
	p	*0,39*	***0,02***	*0,86*	*0,21*
Year of study	2nd	69,31	84,80	78,40	67,49
	3rd	74,43	78,44	74,70	72,43
	p	*0,06*	***0,01***	*0,13*	*0,10*
High school education	vocational	73,03	78,00	75,79	73,18
	grammar school	74,50	80,64	74,56	70,30
	p	*0,42*	*0,14*	*0,46*	*0,15*
Type of student	army	73,46	82,03	73,63	70,89
	business	73,94	77,46	76,24	72,36
	p	*0,80*	***0,01***	*0,12*	*0,48*

5.4 Research question 3

Several χ^2 tests were conducted to respond to the third research question: input variables were related to the demographic characteristics, and output variables were the dominant decision-making style using the Scott & Bruce test. Columns I, D, R, S, and A contain information about the distribution of students' dominant styles in the Scott & Bruce test.

When analyzing Table 4, we can conclude that there are significant differences in dominant decision-making styles in DS1 with respect to high school education and type of students.

In DS2, there is a significant difference among students in dominant decision-making style with respect to the type of students. Further, in DS3, there is a significant difference in the dominant decision-making style with respect to gender. In DS4, there is a significant difference among students in dominant decision-making style with respect to the type of students.

Table 4. Dominant Scott & Bruce styles (DS1)

DC	Values	R	I	D	A	S	SUM
Gender	Male	39	29	9	5	3	85
	Female	86	49	32	7	3	177
	χ^2=4,255, cv=0,126, p=0,372						
	20	18	8	4	0	1	31
	21	57	43	22	4	5	131
	22	34	20	11	4	0	69
	23	10	7	4	3	0	24
	24	5	0	0	1	0	6
	29	1	0	0	0	0	1
	χ^2=24,057, cv=0,297 p=0, 239						
Year of study	2nd	18	12	4	1	0	35
	3rd	107	66	37	11	6	227
	χ^2=2,040, cv=0,087 p=0,728						
High school education	vocational	26	11	53	6	8	94
	grammar school	26	9	56	1	1	104
	χ^2=10,832, cv=0,199 p=0,028						
Type of student	army	67	51	32	4	3	157
	business	58	27	9	8	3	105
	χ^2=12,438, cv=0,213 p=0,014						

5.5 Research question 4

Several χ^2 tests were conducted to respond to the fourth research question: input variables were related to the demographic characteristics, and output variables were the dominant decision-making style using the Rowe test. Columns A, B, C, and D contain information about the distribution of students' dominant styles in the Rowe test.

Table 5. Dominant Rowe styles (DS1)

DC	Values	D	A	C	B	SUM
Gender	Male	14	37	5	12	68
	Female	34	57	19	46	156
	χ^2=7,231, cv=0,177 p=0,064					
	20	2	15	10	4	31
	21	22	50	27	32	131
	22	17	20	16	16	69
	23	6	6	6	6	24
	24	1	2	3	0	6
	29	0	1	0	0	1
	χ^2=15,916, cv=0,239 p=0,387					
Year of study	2nd	12	10	3	10	35
	3rd	36	84	59	48	227
	χ^2=10,865, cv=0,200 p=0,012					
High school education	vocational	28	43	28	35	134
	grammar school	20	51	34	23	128
	χ^2=4,943, cv=0,136 p=0,176					
Type of student	army	29	52	36	40	157
	business	19	42	26	18	105
	χ^2=2,143 cv=0,103 p=0,543					

When analyzing Table 5, we can conclude that there is a significant difference in the dominant decision-making style using the Rowe instrument in DS1 with respect to the year of study (however, we should be careful with this conclusion due to the high difference in a number of students in each set). There is no significant difference in the dominant decision-making style using the Rowe instrument in other datasets.

5.6 Research question 5

In research questions 5-8, we can see the differences between 2020 and 2022 datasets and possibly discuss the role of COVID-19.

Similar analysis as in RQ1 was applied to respond to the fifth research question, but now with respect to the year. The results are presented in Table 6.

Table 6. Scott & Bruce styles – analysis by year

Sample	Year	R	I	D	A	S
DS1	2020	19,43	19,06	17,21	12,58	14,75
	2022	20,15	19,55	17,80	13,13	14,75
	p	***0,04***	*0,14*	*0,24*	*0,33*	*0,98*
DS6	2020	19,70	18,98	17,38	13,80	14,36
	2022	20,05	19,58	18,65	13,65	14,38
	p	*0,46*	*0,17*	***0,03***	*0,84*	*0,96*
DS7	2020	19,13	19,15	17,03	11,23	15,20
	2022	20,35	19,47	16,00	12,02	15,52
	p	***0,03***	*0,55*	*0,25*	*0,34*	*0,64*

Business students in 2022 significantly more apply rational style than business students in 2020. Army students significantly more apply dependent style than army students in 2020.

5.7 Research question 6

Similar analysis as in RQ2 was applied to respond to the sixth research question, but now with respect to the year. The results are presented in Table 7.

Table 7. Rowe styles – analysis by year

Sample	Year	D	A	C	B
DS1	2020	71,62	80,78	76,03	71,56
	2022	76,08	77,65	74,26	72,00
	p	***0,01***	*0,07*	*0,28*	*0,82*
DS6	2020	72,95	81,06	76,62	69,34
	2022	76,17	76,23	74,42	73,16
	p	*0,14*	***0,03***	*0,28*	*0,14*
DS7	2020	70,13	80,46	75,38	74,01
	2022	75,87	80,67	73,92	69,52
	p	*0,07*	*0,94*	*0,61*	*0,16*

Students in 2022 significantly more often apply directive style than students in 2020. Army students in

2020 significantly less apply analytic style than army students in 2020.

5.8 Research question 7

Similar analysis as in RQ3 was applied to respond to the sixth research question, but now with respect to the year. The results are presented in Table 8.

The distribution of 2022 army students per Scott & Bruce's dominant decision styles is significantly different from the distribution of army students in 2020.

Table 8. Scott & Bruce dominant styles – analysis by year

Sample	Values	R	I	D	A	S	SUM
DS1	2020	70	39	15	7	6	137
	2022	55	39	26	5	0	125
	χ^2=10,557, cv=0,197, p=0,032						
DS6	2020	40	16	6	7	3	72
	2022	32	28	21	4	0	85
	χ^2=15,342, cv=0,298, p=0,004						
DS7	2020	30	23	9	0	3	65
	2022	23	11	5	1	0	40
	χ^2=4,612, cv=0,205, p=0,329						

5.9 Research question 8

Similar analysis as in RQ4 was applied to respond to the sixth research question, but now with respect to the year. The results are presented in Table 9.

Table 9. Rowe dominant styles – analysis by year

Sample	Year	D	A	C	B	SUM
DS1	2020	17	56	38	26	137
	2022	31	38	24	32	125
	χ^2=10,785, cv=0,199 p=0,012					
DS6	2020	13	31	0	11	55
	2022	21	21	19	24	85
	χ^2=22,226, cv=0,370 p=0,0001					
DS7	2020	4	25	0	15	44
	2022	10	17	5	8	40
	χ^2=11,060, cv=0,341 p=0,011					

The distribution of 2022 students per Rowe's dominant decision styles is significantly different from the distribution of students in 2020. This conclusion is the case in all three analyzed datasets.

6 Conclusion

In this paper, we were dealing with analyzing the decision-making styles of the population of business and army students from Croatia. The data were collected in 2020 and 2022. We analyzed the use of different decision-making styles by Scott & Bruce and Rowe instruments and identified differences among students with respect to demographic data.

Special attention is passed to analyzing the differences among students with respect to the year when the data were collected. The identified differences can be partly analyzed in the light of COVID-19 influence: COVID-19 changed the way how people reason while making decisions. However, we must not exclude other variables, especially if we know that respondents in 2020 and 2022 were not the same participants.

This paper continues the previous research and covers some of the analyses announced then as future research. Additionally, discussing the minimal decision-making style (less dominant, rarely used style) and introducing the intensity of dominance are to be done in future research. The intensity of dominance can be investigated by (1) analyzing the scores achieved by the dominant decision-making style and all scores achieved by all other styles; and (2) analyzing the scores achieved by the dominant style and the second-best dominant style. In addition, correlation analysis can be applied: (1) correlation between dominant styles per two instruments, (2) correlation between minimal styles per two instruments, and (3) correlation between dominant and minimal styles. The correlation analysis should also consider the intensity of dominance.

Acknowledgments

This work has been fully supported by the Croatian Science Foundation under the project IP-2020-02-5071.

References

Abdelsalam, H. M., Dawoud, R. H., & ElKadi, H. A. (2013). An examination of the decision making styles of Egyptian managers. Business Strategies and Approaches for Effective Engineering Management, January, 219–236. https://doi.org/10.4018/978-1-4666-3658-3.ch013

Ali, A. J. (1993). Decision-Making Style, Individualism, and Attitudes toward Risk of Arab Executives and Its Environment in the Middle East. International Studies of Management & Organization, 23(3), 53–73.

Al-Omari, A. A. (2013a). The relationship between decision making styles and leadership styles among public schools principals. International Education Studies, 6(7). https://doi.org/10.5539/ies.v6n7p100

Boulgarides, J. D. (1984). The Decision Style Inventory : O.D. Applications. Management

Research News, 7(4), 17–20. https://doi.org/10.1108/eb027850

Boulgarides, J. D., & Cohen, W. A. (2001). Leadership Styles vs Leadership Tactics. Journal of Applied Management and Entrepreneurship, 6.

Connor, P. E., & Becker, B. W. (2003). Personal value systems and decision-making styles of public managers. Public Personnel Management, 32(1). https://doi.org/10.1177/009102600303200109

Dyczkowska, J., & Dyczkowski, T. (2018). Democratic or Autocratic Leadership Style? Participative Management and its Links to rewarding Strategies and Job Satisfaction in SMEs. Athens Journal of Business & Business, 4(2), 193–218. https://doi.org/10.30958/ajbe.4.2.5

Efendi, D., & Kusuma, E. A. (2021). The Role of the Management Accounting System and Decision-Making Style on Managerial performance. Jurnal Keuangan Dan Perbankan, 25(1). https://doi.org/10.26905/jkdp.v25i1.5330

el Othman, R., el Othman, R., Hallit, R., Obeid, S., & Hallit, S. (2020). Personality traits, emotional intelligence and decision-making styles in Lebanese universities medical students. BMC Psychology, 8(1). https://doi.org/10.1186/s40359-020-00406-4

Gambetti, E., Fabbri, M., Bensi, L., & Tonetti, L. (2008). A contribution to the Italian validation of the General Decision-making Style Inventory. Personality and Individual Differences, 44(4). https://doi.org/10.1016/j.paid.2007.10.017

Gambetti, E., Zucchelli, M. M., Nori, R., & Giusberti, F. (2022). Default rules in investment decision-making: trait anxiety and decision-making styles. Financial Innovation, 8(1). https://doi.org/10.1186/s40854-021-00329-y

Khairunneezam Mohd Noor. (2020). The Decision Making Styles Of Deputy Vice Chancellors In Malaysian Public Universities And Its Effects Towards Academics' Organizational Commitment And Satisfaction. https://oarep.usim.edu.my/jspui/handle/123456789/6878

Kostanjevac, M., Gligora Marković, M., & Kadoić, N. (2021). The differences in decision-making styles in the Croatian student population of military and nonmilitary studies. In I. Tomičić & B. Kos (Eds.), Proceedings of the Central European Conference on Information and Intelligent Systems (pp. 315–322). University of Zagreb, Faculty of Organization and Informatics Varaždin.

Lührs, N., Jager, N. W., Challies, E., & Newig, J. (2018). How Participatory Should Environmental Governance Be? Testing the Applicability of the Vroom-Yetton-Jago Model in Public Environmental Decision-Making. Environmental Management, 61(2), 249–262. https://doi.org/10.1007/s00267-017-0984-3

Motvaseli, M., & Lotfizadeh, F. (2016). Entrepreneurs` Cognitive and Decision Making Styles. ASEAN Marketing Journal, 7(2). https://doi.org/10.21002/amj.v7i2.5264

Palmiero, M., Nori, R., Piccardi, L., & D'Amico, S. (2020). Divergent Thinking: The Role of Decision-Making Styles. Creativity Research Journal, 32(4). https://doi.org/10.1080/10400419.2020.1817700

Robbins, S., DeCenzo, D. A., & Wolter, R. . (2016). Supervision Today. In S. Wall (Ed.), Eight edition (Eight Edit). Pearson Education.

Rowe, A. J. , & Davis, S. A. (1996). Intelligent information systems: Meeting the challenge of the knowledge era. . CT: Quorum books.

Rowe, A. J. , & Mason, O. R. (1987). Managing with style: A guide to understanding, assessing, and improving decision making. CA: Jossey-Bass.

Rowe, A. J., & Boulgarides, J. D. (1983). Decision Styles — A Perspective. In Leadership & Organization Development Journal (Vol. 4, Issue 4, pp. 3–9). https://doi.org/10.1108/eb053534

Scott, S. G., & Bruce, R. A. (1995). Decision-Making Style: The Development and Assessment of a New Measure. Educational and Psychological Measurement, 55(5), 818–831. https://doi.org/10.1177/0013164495055005017

Sikavica, P., Hernaus, T., Begičević Ređep, N., & Hunjak, T. (2014). Poslovno odlučivanje. Školska knjiga Zagreb.

Thunholm, P. (2004). Decision-making style: Habit, style or both? Personality and Individual Differences, 36(4), 931–944. https://doi.org/10.1016/S0191-8869(03)00162-4

Urieta, P., Sorrel, M. A., Aluja, A., Balada, F., Lacomba, E., & García, L. F. (2022). Exploring the relationship between personality, decision-making styles, and problematic smartphone use. Current Psychology. https://doi.org/10.1007/s12144-022-02731-w

Weerasekara, S., & Bhanugopan, R. (2022). The impact of entrepreneurs' decision-making style on SMEs' financial performance. Journal of Entrepreneurship in Emerging Economies. https://doi.org/10.1108/JEEE-03-2021-0099

Digital Transformation Technology Evaluation in a post COVID-19 World with an Integrated Fuzzy SAW-ARAS Methodology

Merve Güler, Esin Mukul, Gülçin Büyüközkan
Industrial Engineering Department
Galatasaray University
Ortaköy-Istanbul, 34349, Turkey
{mguler, emukul, gbuyukozkan}@gsu.edu.tr

Abstract. *The COVID-19 pandemic has generated unfamiliar worldwide health and economic crisis and caused many results. In order to be successful in the changing world, organizations should efficiently integrate digital technologies into their digital working practices. This paper aims to present an analytical technique for evaluating the digital technologies in a post COVID-19 world. Fuzzy sets approach is used for reflecting uncertainty in the decision-making process. Fuzzy Simple Additive Weighting (SAW) technique is used to calculate critical success factors' weights and fuzzy Additive Ratio ASsessment (ARAS) technique is used to rank digital technologies. A case study is realized to show the potential of the research methodology. Finally, the conclusions and perspectives for future research are presented.*

Keywords. ARAS, Digital Transformation, MCDM, SAW, Technology Evaluation.

1 Introduction

The COVID-19 pandemic has drastically caused organizational changes and worked as a catalyst for digital transformation in many sectors (Hawash et al., 2020; Priyono et al., 2020). Companies have been forced to adopt new digital working practices and redesign their business models (Almeida et al., 2020). In this context, digital transformation can be defined as "… the journey of using digital technologies to develop new business models and strategies"(Büyüközkan & Güler, 2020). Digital technologies' capabilities and their application fields have been increasing day by day. For instance, higher education institutions are using cloud-based services allowing for on-demand access and improving their security (Insights & Innovators, 2022).

In the literature, the digital transformation subject in a post COVID-19 world has been examined in various studies. (Hawash et al., 2020) investigated the impact of the pandemic on oil and gas sector while (Agostino et al., 2020) examined the accelerator role of the pandemic for the public service delivery. (Priyono et al., 2020) aimed to determine paths of digital transformation for small and medium sized enterprises. (Almeida et al., 2020; Hai et al., 2021; Kutnjak, 2021) investigated the challenges and opportunities of digital transformation in a post COVID-19 world. (Kruszyńska-Fischbach et al., 2021) aimed to assess e-readiness of healthcare organizations during the pandemic in Poland. (Bartsch et al., 2021) examined the role of leadership on service providers' digital transformation process during the pandemic.

In this paper, the aim is providing an analytical technique for evaluating digital technologies in a post COVID-19 world. The digital transformation of an organization is affected by a number of critical success factors. The importance degrees of these factors can be taken into consideration by using Multi-Criteria Decision-Making (MCDM) techniques. The importance degrees of the critical success factors are determined by implementing the fuzzy Simple Additive Weighting (SAW) technique. The digital technologies are ranked by applying the fuzzy Additive Ratio ASsessment (ARAS) technique.

The fuzzy sets are proposed by (Zadeh, 1965) as a class of objects with a continuum of grades of membership. In this study, fuzzy MCDM techniques are used for reflecting the uncertainty and imprecision of information. This study contributes to literature by integrating the fuzzy SAW-ARAS techniques for the digital transformation technology evaluation. The critical success factors for digital transformation are determined based on literature survey, industry reports and expert opinions. Group Decision Making (GDM) approach is preferred by advising many Decision-Makers (DMs). In this regard, subjectivity in the decision-making process is eliminated. Then, a case study about digital technology evaluation is realized with the participation of DMs to illustrate the potential employment of the presented methodology.

The organization of the paper is as follows. The next section explains the research methodology which

consists of three main stages. In the third section, a case study of the research methodology is presented. In the last section, the concluding remarks and perspectives for future studies are summarized.

2 Research Methodology

The methodology is provided in Figure 1. which contains three stages. In the first stage, the critical success factors in the evaluation model and the digital technology alternatives are determined with the help of the literature review and the opinions of the experts. The factors' weights are calculated in the second stage by implementing fuzzy SAW method. In the last stage, the alternatives are ranked by using fuzzy ARAS method.



Figure 1. The stages of the research methodology

2.1. The Critical Success Factors for Digital Transformation

The continuing growth in digital technologies is important as it makes permanent competitive advantage for those organizations who implement and adopt it. To be successful in the digital transformation process, many aspects should be considered. Therefore, in this study, the critical success factors on digital transformation are determined based on experts' views, literature survey and industry reports' investigation. These factors are listed in Table 1.

Table 1. The critical success factors

Critical Success Factors	References
CSF1. Fast decision-making	(Third Stage Consulting Group, 2021) (Kruszyńska-Fischbach et al., 2021)
CSF2. Internal ownership and learning	(Third Stage Consulting Group, 2021) (Kruszyńska-Fischbach et al., 2021)
CSF3. Business process management	(Third Stage Consulting Group, 2021) (Kruszyńska-Fischbach et al., 2021) (Kutnjak, 2021)
CSF4. Flexibility	(Third Stage Consulting Group, 2021) (Kruszyńska-Fischbach et al., 2021)
CSF5. Risk mitigation	(Third Stage Consulting Group, 2021) (Kutnjak, 2021)
CSF6. Human capital management	(Third Stage Consulting Group, 2021) (Kruszyńska-Fischbach et al., 2021)
CSF7. Environmental sustainability	(Arora, 2021) (Hawash et al., 2020)
CSF8. Customer retention and growth	(Arora, 2021) (Hawash et al., 2020)
CSF9. Operational resilience	(Arora, 2021) (Kruszyńska-Fischbach et al., 2021) (Kutnjak, 2021)
CSF10. Workforce retention/reskilling	(Arora, 2021) (Kruszyńska-Fischbach et al., 2021) (Kutnjak, 2021)

2.2. Fuzzy SAW Method

Step 1. The views of the DMs' are collected. The decision matrix with fuzzy linguistic terms provided in Table 2 is constructed. Please refer to (Chou et al., 2008) for details of the method.

Table 2. The fuzzy linguistic scale (Beg & Rashid, 2013)

Linguistic term	Abb.	Fuzzy Numbers
Perfect	P	(0.83,1,1)
Very High	VH	(0.67,0.83,1)
High	H	(0.5,0.67,0.83)
Medium	M	(0.33,0.5,0.67)
Low	L	(0.17,0.33,0.5)
Very Low	VL	(0,0.17,0.33)
None	N	(0,0,0.17)

Step 2. The linguistic terms are transferred into fuzzy numbers as in Table 1.

Step 3. The DMs' weights are not equal, and it reflects the relevance levels of each DM. The fuzzy weights of the DMs are represented by the $\widetilde{\omega_t}$. The importance degrees of DMs (I_t) is computed as:

$$I_t = \frac{d(\widetilde{w_t})}{\sum_{t=1}^{k} d(\widetilde{w_t})}, t = 1,2,\dots,k \qquad (1)$$

In this case, $d(\widetilde{w_t})$ represents the fuzzy weight's defuzzified value.

Step 4. Aggregated fuzzy weights of C_j, $\widetilde{W_j} = (a_j, b_j, c_j, d_j)$, are computed as:

$$\widetilde{W_j} = (I_1 \otimes \widetilde{W_{j1}}) \oplus (I_2 \otimes \widetilde{W_{j2}}) \oplus \dots \oplus (I_k \otimes \widetilde{W_{k1}}) \quad (2)$$

Here, $a_j = \sum_{t=1}^{k} I_t a_{jt}$, $b_j = \sum_{t=1}^{k} I_t b_{jt}$, $c_j = \sum_{t=1}^{k} I_t c_{jt}$, $d_j = \sum_{t=1}^{k} I_t d_{jt}$.

Step 5. The fuzzy weights are defuzzified. The defuzzified $\widetilde{W_j}$, shown as $d(\widetilde{W_j})$, is computed as:

$$d(\widetilde{W_j}) = \frac{1}{4}\left(a_j + b_j + c_j + d_j\right), where\ j = 1, 2, \dots, n \quad (3)$$

Step 6. The normalized weights (W_j) are computed as:

$$W_j = \frac{d(\widetilde{w_j})}{\sum_{j=1}^{n} d(\widetilde{w_j})}, j = 1,2, \dots, n \quad (4)$$

$\sum_{j=1}^{n} W_j = 1$. Finally, the weight vector W=($W_1, W_2, \dots, W_n$) is established.

2.3. Fuzzy ARAS Method

Step 1: The views of the DMs' are collected. The decision matrix with fuzzy linguistic terms provided in Table 2 is constructed. Please refer to (Medineckiene et al., 2015) for details of the method.

Step 2: The matrix is normalized as:

For maxima preferable values of criteria:

$$\tilde{\tilde{x}} = \frac{\widetilde{x_{ij}}}{\sum_{i=0}^{m} \widetilde{x_{ij}}} \quad (5)$$

For minima preferable values of criteria:

$$\widetilde{x_{ij}} = \frac{1}{\widetilde{x_{ij}}}, \qquad \tilde{\tilde{x}}_{ij} = \frac{\widetilde{x_{ij}}}{\sum_{i=0}^{m} \widetilde{x_{ij}}} \quad (6)$$

Step 3: The weighted normalized matrix is constructed as:

$$\tilde{\tilde{x}}_{ij} = \tilde{\tilde{x}}_{ij} \widetilde{w}_j,\ i=0,1,\dots,m \quad (7)$$

w_j is the j^{th} criterion's weight and:

$$\sum_{j=1}^{n} w_j = 1 \quad (8)$$

Step 4: The optimality function value of i^{th} alternative is determined as:

$$\widetilde{S}_i = \sum_{j=1}^{n} \tilde{\tilde{x}}_{ij},\ i=0,1,\dots,m \quad (9)$$

Step 5: In order to find the result, the center of area technique is applied as:

$$S_i = 1/3(S_{i\alpha} + S_{i\beta} + S_{i\gamma}) \quad (10)$$

Step 6: Alternatives' utility degree is determined as:

$$K_i = \frac{S_i}{S_0},\ i=0,1,\dots,m \quad (11)$$

where S_0 is the value of most ideal criterion.

3 Case Study

As a result of the digital transformation, utilization of new digital technologies has emerged as an enabler of the development of new products and services over the past two decades (Insights & Innovators, 2022). Companies can benefit from a variety of digital technologies according to their business needs. Therefore, it is important to evaluate the digital technologies and select the most appropriate one.

In order to evaluate the digital technologies regarding their impacts on critical success factors, different technologies are determined. Then, these technologies are evaluated by using the proposed fuzzy SAW- ARAS techniques. The critical success factors were summarized in Table 1. Digital technology alternatives are selected grounded on industry reports. The alternatives can be summarized as the following (Insights & Innovators, 2022):

- **A1. Mobile Wireless Technology:** This technology's indirect economic impact is estimated to be widespread and touch almost all sectors as the technology is implemented and used to solve productivity problems.
- **A2. Cloud Computing & Distributed Infrastructure:** This technology can provide important cost declines for organizations, data integrity and recovery, minimal maintenance, on demand access, and the flexibility to customized services to the needs of individual clients.
- **A3. Artificial Intelligence & Machine Learning:** This technology replicates human competencies efficiently, effectively, and cost-effectively, thus indicating a labor-saving technology.
- **A4. Robotics & Autonomous Systems**: The importance of this technology can be considered as its strong financial influence as an industrial and disrupting socioeconomic impression through various sectors.
- **A5. Quantum Computing:** Applying this technology opens the door to numerous functions that are beyond the abilities of existing computers.
- **A6. Greentech:** Emerging green technologies are expected to help companies in meeting more aggressive environmental and sustainability goals.

DMs who have experience and knowledge in digital transformation have evaluated the critical success factors by using the fuzzy linguistic terms. These fuzzy linguistic terms and their triangular fuzzy numbers were provided in Table 2. The evaluation of DMs are provided in Table 3.

Table 3. The evaluations of DMs for factors

Factors	DM1	DM2	DM3
CSF1	H	VH	H
CSF2	M	M	VH
CSF3	VH	P	P
CSF4	H	H	L
CSF5	M	L	M
CSF6	M	H	H
CSF7	VH	VH	M
CSF8	P	P	VH
CSF9	VH	M	VH
CSF10	M	H	M

The steps of fuzzy SAW technique Eqs. (1)-(4) are implemented and the critical success factors' weights are found. Table 4 displays the weights.

At the end of the fuzzy SAW application, it is possible to say that the most important critical success factor is found as "CSF3. Business process

management", followed by "CSF7. Environmental sustainability" and "CSF9. Operational resilience".

Then, DMs evaluated digital technologies according to their insights by using comparative linguistic terms sets specified in Table 2. The evaluation of the first DM for alternatives is provided in Table 5. The steps of fuzzy ARAS technique Eqs. (5)-(11) are implemented and the ranking of the technologies is determined. Table 6 displays the results.

Table 4. The weights of factors

Factors	Weights	Ranking
CSF1	0.107	5
CSF2	0.090	7
CSF3	0.134	1
CSF4	0.082	8
CSF5	0.066	10
CSF6	0.090	6
CSF7	0.107	3
CSF8	0.134	1
CSF9	0.107	3
CSF10	0.082	9

Table 5. The evaluation of the first DM for alternatives

Factors	A1	A2	A3	A4	A5	A6
CSF1	H	VH	VH	M	VH	L
CSF2	H	H	VH	VH	M	VH
CSF3	VH	H	H	VH	M	M
CSF4	VH	VH	M	VH	H	H
CSF5	P	VH	H	H	P	VH
CSF6	M	H	H	VH	L	H
CSF7	VH	M	H	M	H	H
CSF8	H	VH	L	H	P	M
CSF9	P	H	VH	H	VH	VH
CSF10	H	H	VH	M	P	H

Table 6. The ranking of alternatives

Ai	$S_{i\alpha}$	$S_{i\gamma}$	$S_{i\beta}$	*Si*	*Ki*	Ranking
A0	0.168	0.198	0.248	0.204	1.000	-
A1	0.091	0.141	0.209	0.147	0.718	3
A2	0.090	0.140	0.215	0.148	0.726	2
A3	0.091	0.140	0.216	0.149	0.727	1
A4	0.088	0.137	0.211	0.145	0.711	5
A5	0.091	0.140	0.209	0.147	0.718	4
A6	0.071	0.126	0.201	0.133	0.650	6

The A3. Artificial Intelligence & Machine Learning is ranked as the first among other technologies (K3:0.727) and A2. Cloud Computing & Distributed Infrastructure (K2:0.6726) is ranked as the second.

4 Conclusion

Digital transformation in a post COVID-19 world is an interesting subject for both researchers and practitioners. According to the politicians and academics, digital transformation can be considered as a critical aspect in the progress of industries and economies in general. Digital transformation is not only an essential, it is also a chance for nations (Assar et al., 2022). Therefore, with the utilization of digital technologies many benefits can be provided for both industries and society.

Digital technologies have a potential to serve as an enabler of digital transformation for organizations. However, the most appropriate digital technology for an organization is dependent on many factors such as company size, industry and strategic goals. Therefore, in this paper, it was aimed to present a research methodology for evaluating the digital technologies. In this context, an integrated fuzzy SAW-ARAS technique is implemented. This paper contributes to the literature by integrating these techniques for digital transformation in post-pandemic world. At the end of the case study, the most important critical success factor is found as "Business process management". According to a survey conducted by (Arora, 2021) organizational change is the first challenge among executives that had recently gone through a digital transformation. Therefore, it is crucial to implement digital technologies by efficiently managing the business processes and the organizational change.

In future research, other fuzzy MCDM methods (e.g. VIKOR, TOPSIS, CODAS, EDAS) can be applied and the results of the existing methodology can be compared with its results. Further, other fuzzy set extensions (e.g. type-2 fuzzy sets, interval valued fuzzy sets, intuitionistic fuzzy sets, hesitant fuzzy sets etc.) may be chosen in future study.

Acknowledgments

The authors would like to express their gratitude towards industry specialists. Galatasaray University's Scientific Research Projects Commission provided support for this work (FOA-2021-1059).

References

Agostino, D., Arnaboldi, M., & Lema, M. D. (2020). New development: COVID-19 as an accelerator of digital transformation in public service delivery. *Public Money and Management*, 1–4. https://doi.org/10.1080/09540962.2020.1764206

Almeida, F., Duarte Santos, J., & Augusto Monteiro, J. (2020). The Challenges and Opportunities in the

Digitalization of Companies in a Post-COVID-19 World. *IEEE Engineering Management Review*, *48*(3), 97–103. https://doi.org/10.1109/EMR.2020.3013206

Arora, R. (2021). 2021 Digital Transformation Assessment. In *The Manufacturer*. https://www.ibm.com/downloads/cas/MPQGMEN9

Assar, S., Rupeika-Apoga, R., Petrovska, K., & Bule, L. (2022). The Effect of Digital Orientation and Digital Capability on Digital Transformation of SMEs during the COVID-19 Pandemic. *Journal of Theoretical and Applied Electronic Commerce Research 2022, Vol. 17, Pages 669-685*, *17*(2), 669–685. https://doi.org/10.3390/JTAER17020035

Bartsch, S., Weber, E., Büttgen, M., & Huber, A. (2021). Leadership matters in crisis-induced digital transformation: how to lead service employees effectively during the COVID-19 pandemic. *Journal of Service Management*, *32*(1), 71–85. https://doi.org/10.1108/JOSM-05-2020-0160

Büyüközkan, G., & Güler, M. (2020). Analysis of companies' digital maturity by hesitant fuzzy linguistic MCDM methods. *Journal of Intelligent and Fuzzy Systems*, *38*(1), 1119–1132. https://doi.org/10.3233/JIFS-179473

Chou, S. Y., Chang, Y. H., & Shen, C. Y. (2008). A fuzzy simple additive weighting system under group decision-making for facility location selection with objective/subjective attributes. *European Journal of Operational Research*, *189*(1), 132–145. https://doi.org/10.1016/j.ejor.2007.05.006

Hai, T. N., Van, Q. N., & Tuyet, M. N. T. (2021). Digital transformation: Opportunities and challenges for leaders in the emerging countries in response to covid-19 pandemic. *Emerging Science Journal*, *5*(Special Issue), 21–36. https://doi.org/10.28991/esj-2021-SPER-03

Hawash, B., Abuzawayda, Y. I., Mokhtar, U. A., Yusof, Z. M., & Mukred, M. (2020). Digital Transformation in the Oil and Gas Sector during Covid-19 Pandemic. *International Journal of Management (IJM*, *11*(12), 717–724. https://doi.org/10.34218/IJM.11.12.2020.067

Insights, L., & Innovators, F. O. R. (2022). *2022 Emerging Technology Trends*.

Kruszyńska-Fischbach, A., Sysko-Romańczuk, S., Rafalik, M., Walczak, R., & Kludacz-Alessandri, M. (2021). Organizational E-Readiness for the Digital Transformation of Primary Healthcare Providers during the COVID-19 Pandemic in Poland. *Journal of Clinical Medicine 2022, Vol. 11, Page 133*, *11*(1), 133. https://doi.org/10.3390/JCM11010133

Kutnjak, A. (2021). Covid-19 Accelerates Digital Transformation in Industries: Challenges, Issues, Barriers and Problems in Transformation. *IEEE Access*, *9*, 79373–79388. https://doi.org/10.1109/ACCESS.2021.3084801

Medineckiene, M., Zavadskas, E. K., Björk, F., & Turskis, Z. (2015). Multi-criteria decision-making system for sustainable building assessment/certification. *Archives of Civil and Mechanical Engineering*, *15*(1), 11–18. https://doi.org/10.1016/j.acme.2014.09.001

Priyono, A., Moin, A., & Putri, V. N. A. O. (2020). Identifying digital transformation paths in the business model of smes during the covid-19 pandemic. *Journal of Open Innovation: Technology, Market, and Complexity*, *6*(4), 1–22. https://doi.org/10.3390/joitmc6040104

Third Stage Consulting Group. (2021). *Digital Transformation 2021 Report* (Vol. 35, Issue 4).

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, *8*(3), 338–353. https://doi.org/10.1016/S0019-9958(65)90241-X

Evaluation of Supply Chain Tools with Fuzzy AHP-COPRAS Methodology for Intelligent Supply Chain Management

Esin Mukul, Merve Güler, Gülçin Büyüközkan
Industrial Engineering Department
Galatasaray University
Ortaköy-Istanbul, 34349, Turkey
{emukul, mguler, gbuyukozkan}@gsu.edu.tr

Abstract. *Intelligent supply chain management, which emerged with the integration of digitalized processes into supply chain management, encourages companies to redesign their supply chains to meet customer needs and expectations for improved purchasing and service quality. Most companies try to strike a balance between working capital, operating costs, and optimal service levels. For this reason, it is very important to use the right supply chain tool that provides the most efficiency in supply chain management. In this study, the evaluation of supply chain tools for intelligent supply chain management is considered as a Multi-Criteria Decision Making (MCDM) problem. However, it becomes difficult for experts to evaluate and make decisions about the system when the information is insufficient and uncertain. Therefore, in this study, the fuzzy approach is used to evaluate the MCDM problem more realistically and flexibly. The aim of the study is to propose an evaluation model and integrated analytic methods for the evaluation of supply chain tools for the intelligent supply chain management. The fuzzy Analytic Hierarchy Process (AHP) method is used to calculate the weights of the criteria and fuzzy COmplex PRoportional ASsessment (COPRAS) is used to rank the supply chain tools. Finally, an application is offered to demonstrate the proposed methodology's possible use. This paper guides researchers and practitioners with an evaluation methodology based on fuzziness for intelligent supply chain management.*

Keywords. AHP, COPRAS, fuzzy approach, supply chain tools, intelligent supply chain management

1 Introduction

Many sectors are changing and developing with Industry 4.0, which has emerged with the developments in technology. One of the issues most affected by this change and development is supply chain management. Intelligent supply chain management, which emerged with the integration of digitalized processes into supply chain management, encourages companies to redesign their supply chains to meet customer needs and expectations for improved purchasing and service quality (Khan et al., 2010; Fore et al., 2016).

Companies want to create a more efficient, fast, and high-quality communication network with their rapidly changing business processes and business partners in the digital age. In addition, they try not to break away from the competition by being more efficient, more effective, and faster integrating into the market in subjects such as inventory optimization, supply chain optimization, forecasting and demand models planning, and supply chain strategies (Büyüközkan & Göçer, 2018; Liao et al., 2019). In this context, companies need several digital tools (i.e. software and programs) in order not to fall behind in this digitalization, manage their stock correctly, determine supply chain strategies and minimize operational costs by communicating with customers actively and dynamically in this context.

In addition, most companies try to strike a balance between working capital, operating costs, and optimal service levels. However, they need to focus their time and effort on the areas that yield the greatest return. For this reason, it is very important to use the right tool that provides the most efficiency in supply chain management (Hosseini et al., 2019).

There are many tools on the market that integrate with the ERP system, thereby facilitating and optimizing the management of the supply chain. In this study, the evaluation of supply chain tools for intelligent supply chain management is considered as a Multi-Criteria Decision Making (MCDM) problem (Hwang and Yoon, 1981). However, it becomes difficult for experts to evaluate and make decisions about the system when the information is insufficient and uncertain. Therefore, in this study, the fuzzy approach (Zadeh, 1965) is used to evaluate the MCDM problem more realistically and flexibly. This approach facilitates the decision-making processes of decision-makers (DMs) in complex and uncertain situations.

The aim of the study is to propose an evaluation model and integrated analytic methods for the evaluation of supply chain tools for the intelligent supply chain management. The fuzzy Analytic Hierarchy Process (AHP) method is used to calculate the weights of the criteria and fuzzy COmplex PRoportional ASsessment (COPRAS) is used to rank the supply chain tools. Finally, an application is offered to demonstrate the proposed methodology's possible use.

In addition, this study will help companies streamline the supply chain management process, determine which system provides the best inventory optimization while minimizing all supply chain management costs and help businesses achieve high returns.

The organization of this paper is as the following: In the next section, intelligent supply chain management is presented. Then, the third section shows the proposed methodology, and the application is provided in the fourth section. Finally, the last section concludes the paper.

2 Intelligent Supply Chain Management

The managerial tasks in which the material and information flow is effectively provided between suppliers, manufacturers, distributors, retailers, and customers in this logistics system where raw materials are transformed into products or services and delivered to the end-user are defined as supply chain management. All supply chain functions should be integrated in supply chain management (Stadtler, 2008).

The benefits of supply chain management are to increase customer satisfaction, continuity in production, performance improvement in deliveries, helping to reduce stocks, efficiency, capacity increase, saving by helping to reduce costs, shortening the supply cycle time, and facilitating adaptation to sudden changes in the market (Lambert & Cooper, 2000).

It is possible to see that many sectors have changed and developed with Industry 4.0, which has emerged with the developments in technology. One of the subjects most affected by this change and development is supply chain management. In this study, the concept of "Intelligent Supply Chain Management", which emerged with the integration of digitalized processes into supply chain management, will be presented (Agrawal et al., 2019; Özek & Yıldız, 2020).

Intelligent supply chain management aims to optimize the supply chain, actively apply modern technologies and strengthen the existing system. In this context, companies can provide the opportunity to save costs, generate more income, increase the speed of reaching customers and move to the forefront of the competition thanks to the intelligent supply chain (Scuotto et al., 2017; Büyüközkan & Göçer, 2018). It differs from traditional supply chains because the latter mainly aims to minimize production, transportation, and logistics costs. It is a customer-centric platform model that leverages and maximizes real-time data from many sources and assists businesses in collaborating by integrating the complete supply chain.

Along with Industry 4.0, evolutions in production, retail, distribution, logistics, inventory planning, customer parameters, along with the developing technology, have taken place rapidly on the path of digitalization. Companies are trying to achieve the success they could not achieve in the increasing digitalization with external resources.

In this study, supply chain tools that will support intelligent supply chain management are evaluated and this evaluation is handled by analytical methods.

3 Proposed Methodology

The proposed methodology consists of three basic steps:

- *Step 1 (Problem definition):* Identification of the criteria and alternatives.
- *Step 2 (Fuzzy AHP method):* Computation of criteria's weights with fuzzy AHP method.
- *Step 3 (Fuzzy COPRAS method):* Ranking supply chain tools with the fuzzy COPRAS.

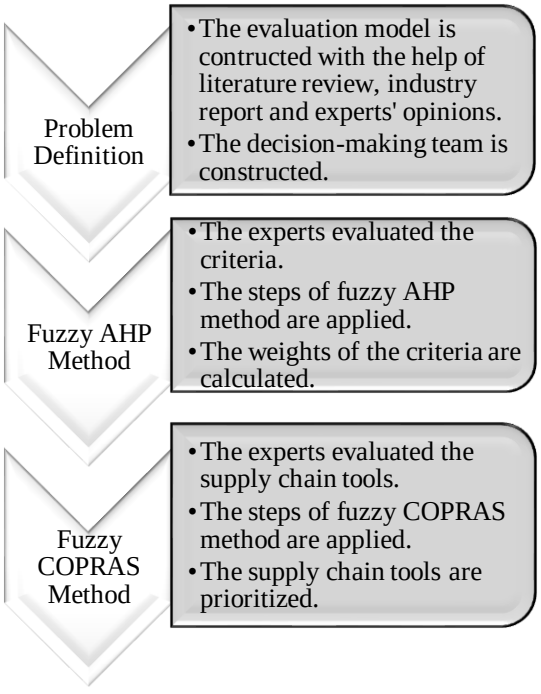


Figure 1. The steps of the proposed methodology

3.1 Proposed Evaluation Model and Supply Chain Tools

This study proposes the evaluation model by conducting a literature review, investigating industrial reports, and taking the experts' opinions. The proposed evaluation model for intelligent supply chain management is as in Table 1.

The supply chain tools to be evaluated in this study are as follows:

- *Slimstock Slim4:* Slim4 provides an end-to-end solution for demand planning, forecasting, and inventory control. It is a software add-on that can be easily, transparently, and quickly integrated with all ERP systems on the market. Portfolio companies in all industries can benefit from Slim4. The software relies on clear and reliable analysis. With Slim4, inventory is delivered at the right time, to the right place, and in the right quantity (URL1; SoftwareConnect, 2016).
- *SAP Integrated Business Planning:* Powered by SAP HANA memory technology, SAP IBP is a cloud solution that offers real-time supply chain management by providing a balance of supply and demand in the medium and short term, while providing strategic planning understanding in line with long-term goals. With SAP IBP, many modules such as demand planning, order fulfillment, inventory optimization, stock management, sales, and operational planning, control tower, production and supply planning, response and supply management are managed on a single platform (Minnock, 2020).
- *Oracle SCM Cloud:* Oracle offers a platform, infrastructure, and data services as well as software as a service. With Oracle SCM Cloud, it enables organizations to quickly respond to changing demand, supply, and market conditions. It aims to create a resilient network and process designed to lead change. It helps companies modernize their supply chain management processes by providing services to customers such as supply chain planning, product lifecycle management, and order management (URL2, Minnock, 2020).
- *JDA SCM:* JDA Software Inc. has implemented its end-to-end supply chain portfolio as a development platform to further its goal of providing an autonomous supply chain. The AI-powered supply chain management platform enables JDA to build cutting-edge applications in addition to its solutions for the first time. The supply chain management platform allows customers to connect their core business and SaaS applications and combine end-to-end planning, execution, and delivery capabilities (URL3).

Table 1. Evaluation model

Main Criteria	Sub-criteria	References
Customer based (C1)	Service quality (C11)	Büyüközkan & Göçer (2018); Liao et al. (2019); Özek & Yıldız (2020)
	Flexibility and visibility (C12)	Büyüközkan & Göçer (2018); Liao et al. (2019); Hosseini et al. (2019)
	Supplier compatibility (C13)	Scuotto et al. (2017); Büyüközkan & Göçer (2018)
	Ease of use for users (C14)	Özek & Yıldız (2020)
Technology based (C2)	Digital competence (C21)	Büyüközkan & Göçer (2018); Liao et al. (2019)
	Technology capability and integration (C22)	Büyüközkan & Göçer (2017); Scuotto et al. (2017); Büyüközkan & Göçer (2018); Özek & Yıldız (2020)
	Digital performance management (C23)	Özek & Yıldız (2020)
	Real-time monitoring technology (C24)	Agrawal et al. (2019); Özek & Yıldız (2020)
Organization based (C3)	Continuous collaboration (C31)	Scuotto et al. (2017); Büyüközkan & Göçer (2017); Büyüközkan & Göçer (2018); Hosseini et al. (2019); Liao et al. (2019)
	Respect for the privacy (C32)	Büyüközkan & Göçer (2018); Özek & Yıldız (2020)
	Adopting advanced analytics (C33)	Büyüközkan & Göçer (2017); Scuotto et al. (2017); Büyüközkan & Göçer (2018)
	Technology specialized human resources (C34)	Scuotto et al. (2017); Agrawal et al. (2019)
Financial based (C4)	Implementation cost (C41)	Scuotto et al. (2017); Agrawal et al. (2019)
	Service costs (C42)	Agrawal et al. (2019)
	Coordination costs (C43)	Scuotto et al. (2017); Sembiring et al. (2020)
	Financing efficiency (C44)	Büyüközkan & Göçer (2018); Liao et al. (2019)

3.2 Fuzzy AHP-COPRAS Methods

Saaty (1980) invented AHP, which is perhaps the most well-known and commonly utilized decision-making model. It is a strong decision-making approach for determining the priority among several criteria.

Fuzzy AHP is a decision-making strategy that allows decision-makers to make decisions in an MCDM process and facilitates decision-making in uncertain conditions (Aya, 2005). When the literature is investigated, it is discovered that various writers provide a wide range of fuzzy AHP methods.

The following stages are included in the Fuzzy AHP utilized in this paper (Büyüközkan and Çifçi, 2012):

Step 1: Create the fuzzy comparison matrices in Table 2 by utilizing triangular fuzzy numbers.

Table 2. Fuzzy scale

Linguistic expression	Fuzzy Scale
Extremely more importance (EMI)	(8, 9, 10)
Very strong importance (VSI)	(6, 7, 8)
Strong importance (SI)	(4, 5, 6)
Moderate importance (MI)	(2, 3, 4)
Equal importance (EI)	(1, 1, 2)

Step 2: α-cut matrices are built. The α-cut is thought to reflect DMs' confidence in his or her choices. The optimism index μ indicates the level of satisfaction for the judgment matrix. A higher index μ score suggests a higher level of optimism. The optimism index is a linear convex combination defined as

$$\tilde{a}_{ij}^{\alpha} = \mu\tilde{a}_{ijl}^{\alpha} + (1-\mu)\tilde{a}_{ijl}^{\alpha} \forall \alpha \in [0,1] \quad (1)$$

Step 3: Using eq. 2, matrices are normalized, and the consistency ratio (CR) for each matrix is determined.

$$\tilde{r}_{ij} = \frac{\tilde{a}_{ij}^{\alpha}}{\sum_{i}^{k} \tilde{a}_{ij}^{\alpha}} \quad (2)$$

$$CI = \frac{\lambda_{max} - n}{n-1} \quad (3)$$

$$CR = \frac{CI}{RI} \quad (4)$$

where CI refers to consistency index, λ_{max} is the largest eigenvector of the matrix, n is the number of criteria, and RI is the random index.

Step 4: The weights of the main criteria ($\widetilde{w}_i^{CR}$) are obtained using the arithmetic mean. And, these steps are applied for each sub-criteria, and global weights ($\widetilde{w}_{ij}^{G}$) are calculated by multiplying the weight of the main criteria.

$$\widetilde{w}_{ij}^{G} = \widetilde{w}_i^{CR} \times \widetilde{w}_j^{CR} \quad (5)$$

Zavadskas et al. (1994) proposed the COPRAS method and proved its validity. It evaluates alternatives by making step-by-step sequencing of alternatives in terms of importance and utility ratings. It is also an approach based on distance. Zavadskas and Antucheviciene (2007) offer the COPRAS method with fuzzy logic for the evaluation of rural building regeneration alternatives in the literature.

The fuzzy COPRAS method steps are as follows:

Step 1. DMs evaluate the alternatives by using the fuzzy scale given in Table 2. These fuzzy numbers are defuzzified with the center of area method. The crisp value x_{ij} can be found using eq. 6.

$$x_{ij} = \frac{[(U_{x_{ij}} - L_{x_{ij}}) + (M_{x_{ij}} - L_{x_{ij}})]}{3} + L_{x_{ij}} \quad (6)$$

Step 2. The normalized decision matrix is calculated with eq. 7.

$$x_{ij}^{*} = \frac{x_{ij}}{\sum_{i=1}^{m} x_{ij}} \quad (7)$$

Step 3. The weighted normalized decision matrix is calculated by using eq. 8.

$$d_{ij} = x_{ij}^{*}.w_j \quad (8)$$

Step 4. The S_i^- and S_i^+ values are calculated using eq. 9 and eq. 10 for both beneficial and non-beneficial criteria.

$$S_{i^+} = \sum_{j=1}^{k} d_{ij} \quad (9)$$

$$S_{i^-} = \sum_{j=k+1}^{n} d_{ij} \quad (10)$$

Step 5. The relative importance values (Q_i) for each alternative are calculated by using the eq. 11.

$$Q_i = S_{i^+} + \frac{\sum_{i=1}^{m} s_{i^-}}{s_{i^-} * \sum_{i=1}^{m} \frac{1}{s_{i^-}}} \quad (11)$$

Step 6. The biggest relative priority (Q_{max}) value is found.

Step 7. The performance index (P_i) of each alternative is computed with eq. 12.

$$P_i = \left[\frac{Q_i}{Q_{max}}\right] \times 100\% \quad (12)$$

4 Application of the Proposed Methodology

A company operating in the retail sector has some problems in stock management in its warehouses, especially after the COVID-19 pandemic. Especially in warehouses and stores connected to the warehouse, they try to keep stock in hand according to the outlets,

and they do not want to be out of stock. For this reason, it wants to invest in a supply chain tool supported by digital technologies.

In this context, an evaluation model consisting of 4 main criteria and 16 sub-criteria was created in this study. This model will be evaluated for 4 different supply chain tool alternatives determined after the literature review and expert opinions. The names of the tools were kept confidential during the evaluation process and they were named A1, A2, A3, and A4.

The weights of the criteria of the evaluation model will be computed with the fuzzy AHP method. The fuzzy COPRAS method will be used to rank the supply chain tools. Evaluations in this study are made by three DMs. All three DMs are sufficiently knowledgeable and experienced in the field of supply chain management. The weights of the DMs are considered equal. DM1 has 15 years of private sector experience in supply chain management. DM2 has been conducting academic studies, lectures, industrial research, and consultancy on supply chain management for 20 years. DM3 has consulting and project management experience in digital supply chain management.

4.1 Calculation of Criteria Weights with Fuzzy AHP Method

As a starting point, the evaluation model is constructed as in Table 1 with a literature review and expert opinions. Then, the fuzzy comparison matrix between these criteria is structured by using triangular fuzzy numbers in Table 2. The comparison matrix for the main criteria is shown in Table 3.

Table 3. The comparison matrix for the main criteria

	C1			C2		
C1	1,000	1,000	1,000	4,000	5,000	6,000
C2	0,167	0,200	0,250	1,000	1,000	1,000
C3	0,125	0,143	0,167	0,125	0,143	0,167
C4	0,167	0,200	0,250	0,250	0,333	0,500
	C3			**C4**		
C1	6,000	7,000	8,000	4,000	5,000	6,000
C2	6,000	7,000	8,000	2,000	3,000	4,000
C3	1,000	1,000	1,000	0,125	0,143	0,167
C4	6,000	7,000	8,000	1,000	1,000	1,000

α-cut matrices (α=0.5; μ=0.5) are constructed by using eq. 1 and these matrices are normalized as in Table 4. CR is checked by using eq. 3 and eq. 4.

Table 4. The normalized matrix for the main criteria

	C1	C2	C3	C4
C1	0,160	0,217	0,190	0,064
C2	0,480	0,578	0,635	0,596
C3	0,120	0,120	0,127	0,255
C4	0,240	0,084	0,048	0,085

These steps are applied for all sub-criteria, and final global weights for all criteria are calculated. The final criteria weights are shown in Table 5.

Table 5. The final criteria weights

Main Criteria	Weights	Sub-Criteria	Weights	Global Weights
C1	0,158	**C11**	0,542	0,0855
		C12	0,076	0,0119
		C13	0,081	0,0128
		C14	0,302	0,0476
C2	0,572	**C21**	0,332	0,1902
		C22	0,468	0,2680
		C23	0,050	0,0285
		C24	0,150	0,0856
C3	0,156	**C31**	0,090	0,0141
		C32	0,569	0,0885
		C33	0,296	0,0461
		C34	0,045	0,0070
C4	0,114	**C41**	0,355	0,0406
		C42	0,491	0,0561
		C43	0,112	0,0128
		C44	0,042	0,0048

According to the final weights, the most appropriate criteria are "Technology capability and integration (C22)", "Digital competence (C21)" and "Respect for the privacy (C32)".

4.2 Evaluation of Supply Chain Tools with Fuzzy COPRAS Method

Firstly, the matrix between evaluation criteria and supply chain tool alternatives is constructed with the help of the fuzzy scale in Table 2. Table 6 shows the DMs' evaluations for sub-criteria of customer-based (C1).

The evaluation matrices of the other criteria are also formed in this way.

Table 6. The evaluation matrix for sub-criteria of C1

	C11	C12	C13	C14
A1	SI	EI	EI	EI
A2	SI	MI	SI	MI
A3	VSI	EMI	EMI	EMI
A4	EMI	EMI	EMI	EMI

These fuzzy numbers are defuzzified with eq. 6. The weighted normalized decision matrix of the alternatives is calculated. The values of *Qi*, *Si+*, *Si-*, *Pi* are computed using eq. 9-12. Finally, Table 7 shows the results.

Table 7. The final values

	Si-	Si+	Qi	Pi	Ranking
A1	0,719	3,291	4,050	89,967	2
A2	0,981	3,945	4,502	100,000	1
A3	0,675	2,787	3,596	79,882	4
A4	0,625	2,977	3,851	85,550	3

At the conclusion of the fuzzy COPRAS method, the most appropriate supply chain tool is “Slimstock Slim4”. The second, third, and fourth alternatives are ranked as A1, A4, and A3 respectively. The names of other companies are not disclosed due to company policies and privacy.

5 Conclusion

Digitalization, which is an absolute business necessity of our age, is also reflected in supply chain management applications, in this case, it increases smart supply chain management applications.

This paper aimed to propose an evaluation model for intelligent supply chain management and rank supply chain tools using an integrated fuzzy MCDM approach. The weights of the criteria were obtained with the fuzzy AHP method, and the fuzzy COPRAS method was used to rank supply chain tools.

An application was shown to demonstrate the methodology's efficacy, and the outcomes of this examination were presented. The most appropriate criteria were found as “Technology capability and integration (C22)”, “Digital competence (C21)” and “Respect for the privacy (C32)”, and the first ranked supply chain tool was determined as “Slimstock Slim4”.

Slim4 offers a complete solution to its customers in stock optimization and inventory management. It works to maximize profits by optimizing the stocks and inventories managed by its customers. In addition to the current business rules it contains, it is supported by updates and new business rules in order to follow a successful path in market movements that differ from sector to sector, in line with the wishes and needs of the customers (URL1).

The number of supply chain tools can be considered as the limitation of the article. A more comprehensive evaluation can be made by increasing this number.

For future investigation, aggregation operators can be used to aggregate DMs' evaluations in group decision-making. On the other hand, other extensions (e.g. elicit information, pythagorean fuzzy sets, spherical fuzzy sets) of fuzzy sets may be implemented into the framework.

Acknowledgments

The authors would like to express their gratitude towards industry specialists. Galatasaray University's Scientific Research Projects Commission provided support for this work (FOA-2021-1059).

References

Agrawal, P., Narain, R., & Ullah, I. (2019). Analysis of barriers in implementation of digital transformation of supply chain using interpretive structural modelling approach. *Journal of Modelling in Management.*

Ayağ, Z. (2005). A fuzzy AHP-based simulation approach to concept evaluation in a NPD environment. *IIE transactions*, 37(9), 827-842.

Büyüközkan, G., & Çifçi, G. (2012). A combined fuzzy AHP and fuzzy TOPSIS based strategic analysis of electronic service quality in healthcare industry. *Expert systems with applications*, 39(3), 2341-2354.

Büyüközkan, G., & Göçer, F. (2017, June). An extension of MOORA approach for group decision making based on interval valued intuitionistic fuzzy numbers in digital supply chain. *Proceedings of the 17th world congress of international fuzzy systems association and 9th international conference on soft computing and intelligent systems (IFSA-SCIS)* (pp. 1-6). IEEE.

Büyüközkan, G., & Göçer, F. (2018). An extension of ARAS methodology under interval valued intuitionistic fuzzy environment for digital supply chain. *Applied Soft Computing*, 69, 634-654.

Fore, V., Khanna, A., Tomar, R., & Mishra, A. (2016, November). Intelligent supply chain management system. *International Conference on Advances in Computing and Communication Engineering (ICACCE)* (pp. 296-302). IEEE.

Hosseini, S., Ivanov, D., & Dolgui, A. (2019). Review of quantitative methods for supply chain resilience analysis. *Transportation Research Part E: Logistics and Transportation Review*, 125, 285-307.

Hwang, C. L., & Yoon, K. (1981). *Methods for multiple attribute decision making.* In Multiple attribute decision making (pp. 58-191). Springer, Berlin, Heidelberg.

Khan, M. Z., Al-Mushayt, O., Alam, J., & Ahmad, J. (2010). Intelligent Supply Chain Management. J. *Softw. Eng. Appl.*, 3(4), 404-408.

Lambert, D. M., & Cooper, M. C. (2000). Issues in supply chain management. *Industrial marketing management*, 29(1), 65-83.

Liao, H., Wen, Z., & Liu, L. (2019). Integrating BWM and ARAS under hesitant linguistic environment for digital supply chain finance supplier section. *Technological and Economic Development of Economy*, 25(6), 1188-1212.

Minnock, O. (2020). Top 10: Supply chain management software companies. Retrieved from https://supplychaindigital.com/technology/top-10-supply-chain-management-software-companies

Özek, A., & Yildiz, A. (2020). Digital supplier selection for a garment business using interval type-2 fuzzy topsis. *Textile and Apparel*, 30(1), 61-72.

Saaty, T. L. (1980). *The analytic hierarchy process*. McGraw-Hill, New York, 324.

Scuotto, V., Caputo, F., Villasalero, M., & Del Giudice, M. (2017). A multiple buyer–supplier relationship in the context of SMEs' digital supply chain management. *Production Planning & Control*, 28(16), 1378-1388.

Sembiring, N., Tambunan, M. M., & Ginting, E. (2020, April). Collaboration of sustainability and digital supply chain management of achieving a successful company. *IOP Conference Series: Materials Science and Engineering* (Vol. 830, No. 3, p. 032093). IOP Publishing.

SoftwareConnect. (2016). Optimal Supply Chain Management With Slim4 Software From Slimstock. Retrieved from https://softwareconnect.com/demand-planning/slimstock-slim4/

Stadtler, H. (2008). Supply chain management—an overview. *Supply chain management and advanced planning*, 9-36.

URL1: https://www.slimstock.com/tr/

URL2: https://www.oracle.com/tr/scm/

URL3: https://csnews.com/jda-supply-chain-management-platform

Zadeh, L. A. (1965). Fuzzy sets. Information and control, 8(3), 338-353.

Zavadskas, E. K., & Antucheviciene, J. (2007). Multiple criteria evaluation of rural building's regeneration alternatives. *Building and Environment*, 42(1), 436-451.

Zavadskas, E. K., Kaklauskas, A., & Sarka, V. (1994). The new method of multicriteria complex proportional assessment of projects. *Technological and economic development of economy*, 1(3), 131-139.

Business model development and business ecosystem evaluation in the energy domain

Jan Fauser, Dieter Hertweck
Faculty of Informatic
Reutlingen University
Danziger Str. 6, 71034 Böblingen
{jan.fauser, dieter.hertweck}@reutlingen-university.de

Abstract. *The energy turnaround, digitalization and decreasing revenues forces enterprises in the energy domain to develop new business models. Following a Design Science Research approach, we showed in two action research projects that businesses models in the energy domain result in complex ecosystems with multiple actors. Additionally, we identified that municipal utilities have problems with the systematic development of business models. In order to solve the problem, we captured together with the partners of the enterprises the requirements in a second phase. Further we developed a method which consist of the following components: Method for the creative development of a new business model in form of a Business Model Canvas (BMC). A mapping between the e3Value ontology and the BMC for modelling a business ecosystem. The Business Model Configurator (BMConfig) prototype for modelling and simulating the e3Value-Ontology. The Business model can be quantified and analyzed for its viability. We demonstrate the feasibility of our approach in business model of a power community.*

Keywords. Business ecosystem, e3Value, BMC, energy turnaround, BMConfig.

1 Introduction

As a result of the energy turnaround, digitization, the growth of renewable energies and dynamic markets, companies in the energy domain must adapt their business models to the new market conditions. Electricity or gas is not an emotional product, so, it is difficult to establish customer relationships and attract customers with another purchase criterion than price. Internet comparison portals enable customers of energy suppliers to switch their electricity or gas provider in a few minutes. In this context, business models must be developed that go beyond the sale of electricity and/or gas to tackle the high fluctuation of customers and generate constant revenues. The development of business models often takes place in innovation workshops. Methods such as design thinking or value proposition design (Osterwalder et al., 2014) are used. For conceptualizing, visualizing and structuring business models the business model canvas (BMC) (Osterwalder et al., 2010) is a well-established method in science and practice today. Many municipal utilities lack expertise in systematic customer-centric business model development (Meyer et al., 2021). Additionally, the business models, especially in the context of renewable energies are often based on complex energy systems, which is built on a different logic than centralized large-scale power plants. A number of authors in the scientific discourse and our experience in different case studies showed, that business models in the energy domain are realized as an ecosystem, and it is essential to examine one's own value chain and investigate the integration of new actors to enhance or develop a new value proposition (Bocken et al., 2014; Breuer & Lüdeke-Freund, 2014; Hellström et al., 2015; Küller et al., 2015). An isolated view of the business model, for instance in form of a BMC in the renewable energy domain is not enough. Therefore, the business ecosystem (BE) approach could be a suitable method to analyze the business models and gain a more profound understanding about their function. Moore (1996) defines a business ecosystem as: "economic community supported by a foundation of interacting organizations and individuals-the organisms of the business world." The energy system must not be reduced to its technical components alone. The term "energy system" must be understood in a broader way that includes raw materials, resources, technologies, economics, society, and law.

Thus, viewing after the systematic development of a new business model energy system business models need to be evaluated in their ecosystem. Evaluations about the possible outcomes of the proposed business model help to generate the needed information to reduce uncertainty and manage business model risks (Thompson & MacMillan, 2010). At the end of a business model innovation workshop, the question often arises as to how the creatively developed business model solutions visualized using the BMC

can be transferred into more formalized methods that can be evaluated from a business management perspective.

This paper aims to develop an artefact for the development of business models in the domain of renewable energies and validate them for their economic viability in a business ecosystem setting. Following the proposed Design Science Research (DSR) method described by Johannesson and Perjons (2014) the following research question has been formulated: *What are the components of an integrated method for transforming business models into business ecosystems and evaluate their economic* viability*?*

The remainder of the paper is organized as followed. First, we give a short overview of the DSR methodology, followed by describing our research method. In Section three we explain the problem, in section four the capture of the requirements for our developed artefact. In section five we give an overview of the components and functionality of the developed artefact and section six we demonstrate its feasibility. In the last section we summarize the main findings and discuss future research.

2 Methodology

Design Science Research (DSR) (Hevner et al., 2004) is a design-oriented research paradigm. The aim is to develop a solution in the form of artefacts based on an identified practical problem, which can then be systematically evaluated. The energy domain is facing an on-going change with a wide variety of external drivers. Therefore, a problem focused iterative methodology like the DSR is a valid research approach. It is a widely accepted research paradigm in the field of business informatics. A valuable contribution to the application of DSR is given by Hevner et al, through their conceptual framework and guidelines. However, their developed framework and explanations are very abstract. Johannesson and Perjons (2014) provide a detailed approach for the concretization of the explanations of Hevner et al. in the form of a method. This method consists of the following five activities: Explicate Problem, Define Requirements, Design and Develop Artefact, Demonstrate Artefact and Evaluate Artefact. Within DSR, it has become well established to identify four types of artefacts: constructs (vocabulary and symbols), models (abstractions and representations), methods (algorithms and practices) and instantiations (implemented prototypical systems) (Johannesson and Perjons, 2014, Hevner et al., 2004). Depending on the research project, several types of artefacts may occur in combination.

The paper at hand is part of a larger design science research project and serves as a validation of our approach to construct an artefact for a concrete problem. Following the DSR approach described by (Johannesson and Perjons, 2014), the first two activities are Explicate Problem and Define Requirements. Explicate Problem is about answering the following question: *What is the problem that some stakeholders in practice experience and why is it important?* While Define Requirements focuses on: *What artefact can be a solution for the explicated problem and which requirements on this artefact are important for the stakeholders?*

Both questions were answered in two action research projects (Baskerville, 1999). In the first project three case studies were conducted in a broad research project, which focused on decentralized and flexible solutions for future energy production and distribution with seven universities and one non-university partner. In our second action research project we analyzed, developed and tested business models for municipal utilities in rural areas for economic viability. A university partner and two municipal utilities were involved. For both research projects, we used three data collection methods for case study research to increase construct validity and reliability (Yin, 2009): semi-structured interviews, participating observation, and document study.

The third activity of the DSR framework is Design and Develop Artefact, which creates an artefact that addresses the explicated problem and fulfils the defined requirements. In order to develop an artefact, we needed an integrated approach that supports the systematic development of new business models on the one hand and the modeling and simulation of business ecosystems with different business cases on the other hand. We conducted a literature and web search to identify different approaches to develop business models, analyze, simulate, and verify the economic viability of a business ecosystem and how to combine the two approaches. As additional research method we used prototyping to analyze and test intermediate implemented states of our prototype.

In the fourth activity, “Demonstrate Artefact”, we applied our developed artefact to the identified problems in our action research projects. Thereby, the design is strongly related the demonstration takes place in several stages, which allowed a stepwise validation of the approach and the identification of new problems. The results extend the own experience and serve for discussions with partners. The last step of the DSR method Evaluate Artefact has not yet been carried out.

3 Explicate Problem

Following the DSR approach described by (Johannesson and Perjons, 2014), the first activity is Explicate Problem.

The energy turnaround in Germany and the digitalization means a radical technical change from centralized to decentralized energy production and from a central distribution model to the Smart Grid. The long-established structures of energy production in large plants and conventional distribution networks which allowed stable economies of scale and low unit

costs are over (Doleski, 2014). Drivers such as technological progress in energy generation, storage or control, but also organizational and cultural change through re-municipalization, strong citizen participation or the development of consumers into prosumers, have massively changed the face of the energy industry in the course of the energy turnaround. In this context, energy suppliers are facing a variety of challenges. (i) Increasing numbers of actors are entering the market to offer products and services along the energy value chain. (ii) As a result of the increase of renewable energies in the energy mix, revenues from many energy supplies are decreasing (Frantzis et al., 2008). (iii) Internet comparison portals enable customers of energy suppliers to switch their electricity or gas provider in a few minutes (Meyer et al., 2021). Additionally, electricity or gas is not an emotional product, as a result, it is becoming increasingly difficult to establish customer relationships and attract customers with another purchase criterion than price. (iv) The constantly changing legal framework makes the implementation of new business models difficult due to the lack of long-term planning security (Engelken et al., 2016). (v) The growth of renewable energies and new technologies creates new business model opportunities. However, the mindset and the skills of many employees of utilities are a barrier to dealing with new technologies (Meyer et al., 2021). (vi) Increasing customer demands e.g., for innovative and renewable value (Valocchi et al., 2014). These and other factors have a crucial impact on the profit of utilities and threaten their existence. From an ecological perspective, it is essential to develop new business models in the energy domain. Only through the right economic stimulation, the energy turnaround can be successfully implemented and thus make a valuable contribution to achieving the 2°C target set in the Paris Climate Agreement (*The Paris Agreement*, 2015).

We (Fauser et al., 2019) examined three different energy business models in a larger action research projects in the domain of renewable energy. We discovered that energy systems cannot be reduced only to their technical components. The concept of the "energy system" must be understood in a broader form, including raw materials, resources, technologies, economy, society and law. Another issue discussed with the case study partners is the economic viability of the energy system business model. Due to the complexity of multiple actors involved in the business model and the governmental regulation and subsidies, the calculation of different scenarios to determine the best economic value of the business model is difficult. The scientific literature also states the fact that distributed energy systems rely on a very different logic compared to large-scale centralized power plants (Magnusson et al., 2005; Mirata et al., 2005; Richter, 2012), imply a need for considering the underpinning business models. The development of business models in the renewable energy domain often requires a broader and systemic perspective (Hellström et al., 2015). Serval scholars discuss that a BM in the energy domain will be established as an ecosystem, thereby it is also necessary to investigate how actors integrate their own value chain with those of others (Bocken et al., 2014; Breuer and Lüdeke-Freund, 2014; Hellström et al., 2015; Küller, Hertweck and Krcmar, 2015). An investigation of isolated BMs is not enough in the domain of renewable energy. Therefore, the business ecosystem approach could be suitable to understand their business. We identified the following problems:

In our second action research project, we wanted to examine our developed artefact (BMConfig) (Fauser et al., 2019), which we developed in the three previous mentioned case studies in a larger action research project, for modeling, analysis and simulation of different business cases in a business ecosystem setting. The aim of this project was to develop business models for municipal utilities in rural areas and to test them for their economic viability. A municipal utility is a company that is usually majority-owned by the municipality and mainly provides services of public interest. One of the municipal utilities has 350 employees and provides the following services to their customers: Electricity, gas, district heating, drinking water and wastewater management. In addition, the company operates several parking garages and parking lots in the city area as well as the thermal baths, the city indoor swimming pool and the three outdoor baths. The second municipal utility in our project also focuses on energy, heat and water supply and the operation of a parking garage and it has 50 employees. The first problem we identified was that the participating municipal utilities have limited or no methodological knowledge for the systematic development of new business models. This relates to an interview study we conducted, that energy suppliers have a lack of competencies in the IT area and around the systematic business model development of customer-centric business models (Meyer et al., 2021).

For the systematic development of the business models, which should be developed from the perspective of a focal actor. We conducted a business model innovation workshop for the initial development of a business idea and the conceptualization of a business model. In the next step, we developed a method how we can transfer the creatively developed business model solutions visualized using the BMC into a more formalized methods that can be evaluated from a business management perspective. According to Osterwalder et al. (2010) or Wirtz (2016) managers evaluated feasibility and profitability of the proposed BMI, before implementing it. Based on the state-of-the-art research on distributed energy systems (Funcke & Bauknecht, 2016), but especially on the previous mentioned case studies, we showed that business models in the renewable energy domain are mostly realized in the form of a complex business ecosystem (Koppenhoefer et al., 2017, 2018). Therefore, a

description in the representation of a BMC is not sufficient to develop a holistic understanding about the business model.

In the first action research project and the state of the art we were able to show that business models in the energy domain are mostly based on a business ecosystem. The business ecosystem must be visualized for analysis and to ensure the understanding of all stakeholders involved. We derived the problems 1-3 from the three different case studies.

- Problem 1: The multiple actor's structure of the business model ecosystems and their graphical visualization.
- Problem 2: The relationships between the actors and their value exchanges need to be transparent to fully understand the business model.
- Problem 3: Determination and comparison of the economic viability of various specifications of the energy system business model.
- We identified the following problems in our second research project business models for municipal utilities in rural areas.
- Problem 4: Systematic business model development. The municipal utilities have problems with the systematic development of customer-centric business models. At the same time, the municipal utilities were not familiar with methods such as VPC or, in some cases, BMC.
- Problem 5: Form a business model to a business case. From the representation of a business model in the form of a BMC, a variety of different variants of the business model can be implemented. The question arises how different business cases can be calculated for a business model and which components must be considered.

4 Define Requirements

The second activity defines the requirements. First, the artefact is outlined. This means deciding upon the type of artefact and its basic characteristics. Due to the high number of different identified problems in both research projects, which makes it difficult to transfer into an integrated approach, for instance a model or an instantiation. In our opinion an artifact as a method addresses the identified problems the most. According to (Johannesson & Perjons, 2014) a method defines guidelines and processes for how to solve problems and achieve certain goals. A Method can describe how an artifact is created. Its contains a collection of procedures, which are based on concepts and whose results are represented by a notation (Goldkuhl, 1997) The combination of procedures, concepts and notation forms a method component. Thus, the individual method components are linked in a meaningful way and it is possible to enrich the different activities of the method with models, instances or methods that lead to the achievement of a certain goal. In our case the individual method components should reflect the analysis, ideation and design process for the creation of an initial business model, a mapping between the BMC to transfer into our artefact BMConfig for modelling and simulating business ecosystem different business cases.

In conclusion, this argumentation leads to the following research goal, which is to be addressed through the following conception of the artefact:

Support for the development of new business models and modelling and evaluation of business ecosystems in the energy domain and their economic viability.

Based on the research objective and the underlying identified problems, the next step is to derive requirements for the artefact. The requirements of the artefact were derived from the case studies. We carried out interviews with the involved stakeholders in each case study. That result in a list of requirements, which should be addressed by the artefact.

- Systematic approach for the development of customer centric business models.
- Transformation guidelines from BMC to a business ecosystem approach.
- Graphical visualization of the network.
- Presentation of value-related exchange relationships between the actors.
- Good/easy understandable for non-domain experts.
- Possibility to model elements of value networks: tangible and intangible value flows, processes, resources, organizations (multiple actors).
- Calculation of different business cases.
- Presentation of multiple business models.
- Presentation of the results of the simulation.

5 Design and Develop Artefact

The third activity was to design and develop the outlined artefact.

Figure 1 summarizes our method and its individual components. First, the DIGITRANS method (*Digitrans*, 2018) for a systematic business model business model innovation is described. In the next step, the BMC is transformed into a business ecosystem using a mapping process. Then a quantification of the model with costs, prices quantities and a time specification for how many years/months the ecosystem is to be simulated is carried out. By simulating and analyzing the individual business cases, a validation of the business proposal can be carried out. The individual components of the method are described below.

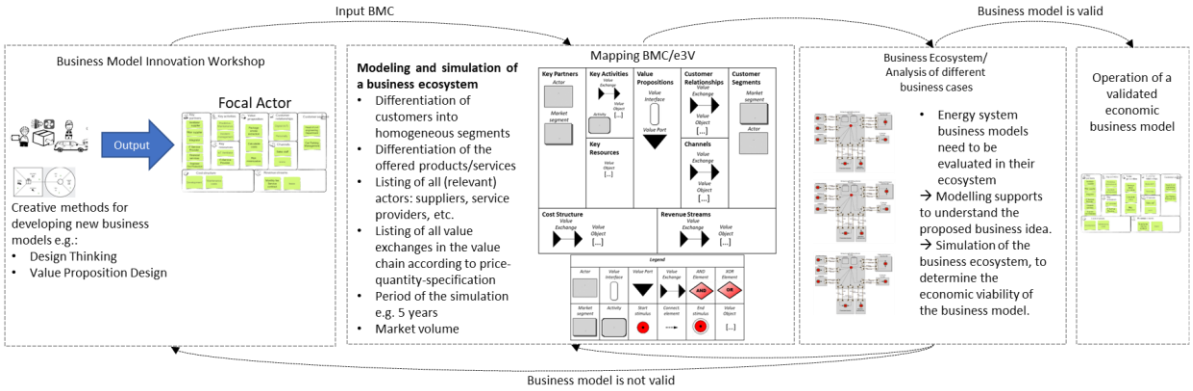


Figure 1. Method integration for the transfer and evaluation of business models in business ecosystems

The business model development process supports the systematic design of the creative process to develop a new business model(Jonda, 2004; Köster, 2014). Wirtz & Daiser (2018) conducted a systematic literature review of business model innovation processes and derived seven generic BMI process steps: Analysis, Ideation, Feasibility, Prototyping, Decision making, Implementation, and Sustainability. In most cases, the development of business models does not follow a linear process, but it is instead continuous adaptation.

In our method we used the DIGITRANS method (*Digitrans*, 2018), which supports the development of business models for SMEs. The method is divided into the innovation and transformation phase. The innovation phase focuses on developing of a new digital business idea, while the transformation phase is addressing the sustainable implementation and realization of the new digital business model in the company's overall organization. The conceptual design and development of the business models takes place in the innovation phase. The transformation processes were only conceptually elaborated in the Digitrans project, but not empirically validated. So, we only used the innovation phase for our method. The Digitrans innovation phase addresses many of the aspects identified by (Wirtz & Daiser, 2018). The innovation phase is divided into two sub-phases analysis and design. The team for the development of the new business model should consist of cross-functional employees of the company. In the analysis phase the environment and strategy of the enterprise involved is analyzed to identify potentials or market trends. A different aspect of the analysis phase is to build empathy to understand the customer, which is crucial to a customer-centric business model development method. This allows to set aside different assumptions form the team members and to gain insights into the customers' situation. The new developed value proposition should add value for the customer. Potential customers can also be integrated into the development process at an early stage. In the design phase the whole ideation process starts based on the findings form the analysis phase. The new developed ideas should reflect the customer needs. Thereby the whole process is supported by different methods testing, feedback loops, mockups and other methods to encourage the team members to think out of the box. As well the iterative working processes will form an integral part of this phase on the way to improve or modify the business model. After having completed the whole innovation phase a new business model conceptualized in form of a BMC is outlined.

The next step of the method is to transfer the developed business model, which is documented in the form of a BMC, into a business ecosystem. To model and simulate the business ecosystem, we use our own prototype BMConfig, which is based on the e3Value method (Gordijn, 2002) and enables the modelling of value networks or BEs and addresses many of the identified requirements in Chapter 4. In the BMConfig one or more business ideas are realized by a network of actors that jointly create, distribute, and consume value in the context of cross-organizational business models. Thereby it stands for an approach to discuss innovative business ideas and explore their environment with the aim of making a statement about the profitability of a value network or BE. In an initial literature and internet research, we identified that the e3Value-ontology were used in different scientific contribution to model value networks in the energy domain (Kartseva et al., 2004; Koppenhoefer et al., 2018; Küller et al., 2015). A value network is a network of organizations that form a value creation system in which suppliers, partners and customers collaborate to create value (Peppard & Rylander, 2006). The BE approach goes beyond value networks and includes actors who are not directly involved in value creation process (Iansiti & Levien, 2004). Therefore, the BE approach appears to meet the requirements in the energy domain much better.

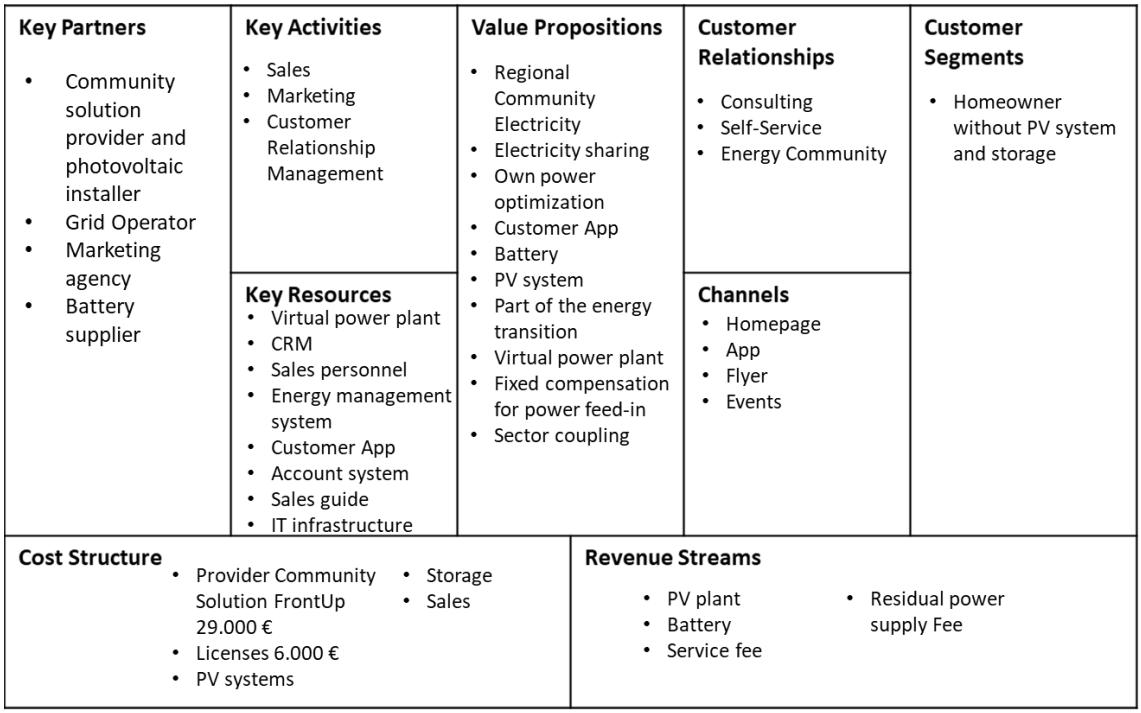


Figure 2. BMC Power Community

Figure 1 shows the individual concepts of the e3Value ontology. The basic elements of e3Value are defined as follows: The Actor is defined as an independent economic entity. A Market Segment is a generalization of an Actor and groups several Actors with similar characteristics. A Value Activity is an operational activity whose execution serves to generate value. It is intended to illustrate transparency about operations within an Actor. Value Ports enable the offering or consumption of value objects. A Value Interface summarizes the exchange of values. For each offer of value objects to the environment, it also contains an adequate reception, so that trade between the actors is created. The value ports are a part of the value interface. Value Exchanges connect two Value Ports for the transfer of values between two actors. Value Objects illustrate which values are exchanged via the Value Exchanges. These are, for example, goods, services or money, whereby a Value Object must represent a value for one or more actors. Scenarios are used to calculate different business cases of the business model. For this purpose, value interfaces are linked via connection elements. A start stimulus marks the beginning of a scenario. The end is represented by an associated stop stimulus. AND/OR operators enable the mapping of logical operations for a scenario and thus the possibility to represent several business models in one business ecosystem.

A mapping between the e3Value-ontology and the BMC is shown in Figure 1. This builds on the prior work of Gordijn et al. (2005) and Caetano et al. (2017). Jaap Gordijn the developer of the e3Value ontologies and Alexander Osterwalder and Pigneuer the authors of the Business Model Ontology (BMO) and the BMC, compare in (Gordijn, Osterwalder and Pigneur, 2005) the BMO and the e3Value-ontology. Similarities, differences and integration possibilities between the two ontologies were analyzed. This should improve the representation, design and analysis of business models. Caetano et al. (2017) uses the approach of (Gordijn, Osterwalder and Pigneur, 2005) and transfers it to the BMC.

The two concepts of the BMC Customer Segment and Key Partners can be translated 1:1 into an e3Value model. For modelling either Actors or Market Segments can be used. The concept Key Activities are the central activities a company has to perform to deliver the Value Proposition. This can be represented in e3Value by a Value Activity, which an actor performs to increase its value, or by Value Exchanges containing Value Objects, which represent the procurement of required resources from a Key Partner. A Key Resource is represented by a Value Object in e3Value, which is acquired from a Key Partner. The Value Proposition is the value that a company offers to its customers. In e3Value the Value Proposition is represented by a Value Interface and the corresponding Value Ports. The Customer Relationships and Channels are represented by Value Exchanges and Value Objects. Revenue Streams and Cost Streams are related to the Value Exchanges and Value Objects and have to be provided with price and quantity accordingly.

The BMC allows a business model to be represented at a higher level of granularity, taking a more strategic perspective (Caetano et al. 2017). Using the BMConfig, a modelling which is closer to the actual operationalization of the business is enabled. Therefore, when transforming a BMC to the e3Value model, additional information about the business

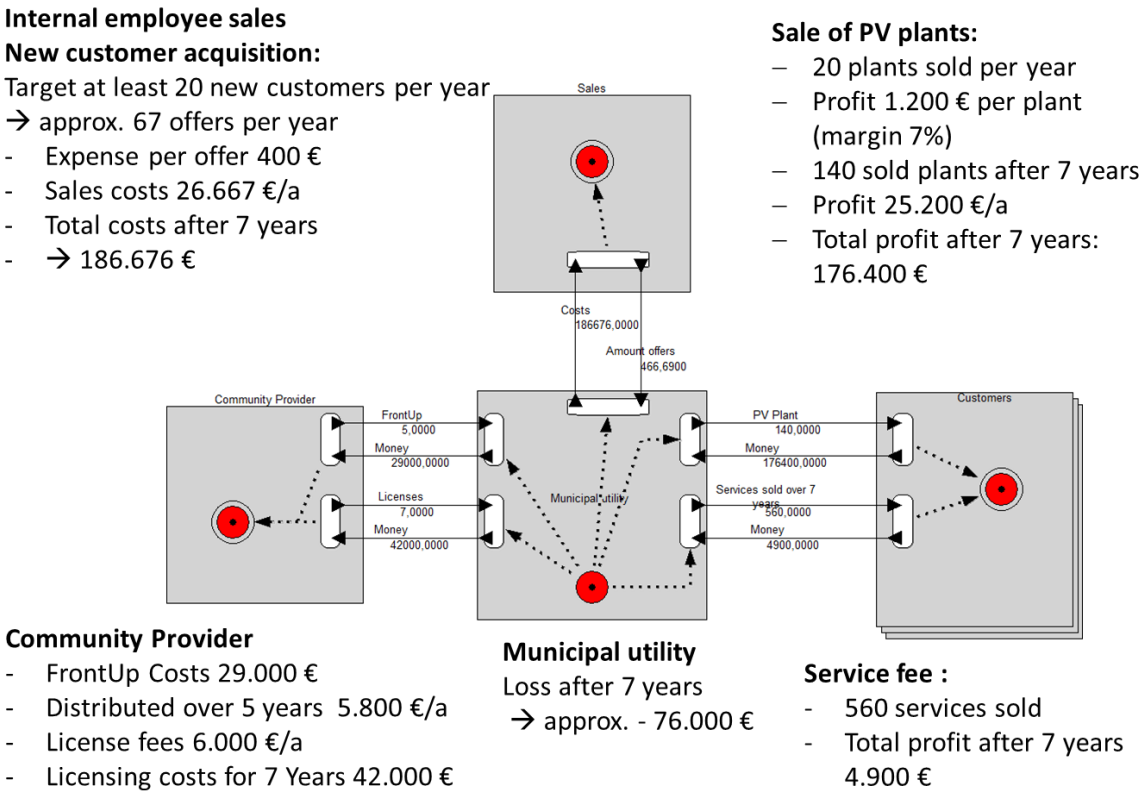


Figure 3. e3Value-Modell: Power Community

model is needed. The transformation of the BMC into an e3Value ontology enables an analysis based on the BE approach and, through simulation, an investigation of the economic viability of the business model.

6 Demonstrate Artefact

In the activity Demonstrate Artefact, the developed Mapping is demonstrated on a case, thereby proving its feasibility (Johannesson and Perjons, 2014).

Using the presented mapping, the developed BMC was transferred into a e3Value modeling in the BMConfig. The power community offer of a company was considered, which was designed as a white label business model especially for the clientele "rural municipal utilities". The BMC for this is shown in figure 2. The provider of the electricity community solution supplies the municipal utilities with a white-label product with its own app, a module for optimizing in-house consumption of electricity from photovoltaic plants and the components for setting up their own electricity community. For the transfer, the BE is modeled from a company's point of view with a specific value proposition. In the first step, the customers and key partners from the BMC were transferred to the e3Value model. Afterwards, we discussed which actors are related and which values they exchange. For this purpose, the Value Exchanges and Value Objects were defined. Through the fine-granular modeling, additional actors and exchange relations were identified, which enable an even more realistic business scenario. For the simulation, costs and prices of the individual services were defined. In addition, the market volume, sales volumes and the desired duration with which the business model should be simulated were determined. Figure 3 shows a simplified model of the power community. The largest cost driver in this business model is photovoltaic plant sales. As a result, a simplified business model was modeled and simulated for an initial analysis and for presentation purposes. The sales department is modeled as an independent actor for a better overview. The business model is simulated for a period of seven years. For the acquisition of 20 new customers per year, the municipal utilities calculate that approximately 67 offers have to be written. The expenses per offer are approximately 400 €. The front-up costs for the solutions are amortized over five years. License costs totaling 42,000 € will be generated over the seven years. 140 PV systems and 560 services are sold over the seven years. The business case examined shows a loss of approximately 76,000 € after seven years.

The transfer of the business model into the BMConfig makes it possible to compare the representation of different specifications of the business model and thereby to reflect on design decisions and to identify relevant components for the further implementation of the business model. New insights are gained through the concretization of various parameters and regarding the simulation. In the

context of the electricity community business model, the following three key barriers were identified, through the modeling and simulation and the resulting discussion: the high distribution costs, the higher price for PV plants compared to the competition, and the low scaling due to the regional roots of the municipal utilities. The analysis shows that selling PV systems is uneconomical for the municipal utilities at the current point in time and that further investigations are necessary.

7 Conclusion

New business models are needed to realize the energy turnaround. As a result of decreasing revenues, digitization and new market entries, enterprises in the energy domain must develop new business models. In four investigated case studies, we identified 5 main challenges, which occurred by developing and analyzing an energy system business model and determine the economic viability. Thereby the stakeholders struggle to fully understand the business ecosystem there are part of. In addition, difficulties in the systematic development of new business models were identified. In order to address the identified problems and requirements, we outlined the artefact a method. We captured the following requirements in our action research projects. A method for systematic development of business model from scratch, the graphical modelling of the business ecosystem and the relations between the actors. Furthermore, the simulation of different business cases to identify the business model with the most economic value was essential to the stakeholders. We developed an artefact that addresses the identified problems and requirements and demonstrated its feasibility on a case study in the field of power community.

The validation and the identification of further requirements which should be addressed in our method is part of our future research with the aim to improve our artefact systematically.

References

Baskerville, R. L. (1999). Investigating Information Systems with Action Research. *Communications of the Association for Information Systems*, *2*(3), 1–32.

Bocken, N. M. P., Short, S. W., Rana, P., & Evans, S. (2014). A literature and practice review to develop sustainable business model archetypes. *Journal of Cleaner Production*, *65*, 42–56.

Breuer, H., & Lüdeke-Freund, F. (2014). Normative innovation for sustainable business models in value networks. *XXV ISPIM Conference-Innovation for Sustainable Economy and Society, 8-11 June 2014, June*, 1–17.

Caetano, A., Antunes, G., Pombinho, J., Bakhshandeh, M., Granjo, J., Borbinha, J., & da Silva, M. M. (2017). Representation and analysis of enterprise models with semantic techniques: an application to ArchiMate, e3value and business model canvas. *Knowledge and Information Systems*, *50*(1), 315–346.

Digitrans. (2018). https://digitrans.me/psm/home

Doleski, O. D. (2014). Entwicklung neuer Geschäftsmodelle für die Energiewirtschaft - das Integrierte Geschäftsmodell. In C. Aichele & O. D. Doleski (Eds.), *Smart Market: Vom Smart Grid zum intelligenten Energiemarkt* (pp. 643–703). Springer.

Engelken, M., Römer, B., Drescher, M., Welpe, I. M., & Picot, A. (2016). Comparing drivers, barriers, and opportunities of business models for renewable energies: A review. *Renewable and Sustainable Energy Reviews*, *60*, 795–809.

Fauser, J., Koppenhöfer, C., & Hertweck, D. (2019). Modelling and evaluating of business model ecosystems in the energy domain. In V. Strahonja & D. Hertweck (Eds.), *Central European Conference on Information and Intelligent Systems Proceedings 30th International Scientific Conference , 2019 : October 2nd - 4th, 2019, Varaždin, Croatia* (pp. 341 – 348).

Frantzis, L., Graham, S., Katofsky, R., Sawyer, H., Frantzis, L., Graham, S., Katofsky, R., & Sawyer, H. (2008). Photovoltaics Business Models. *National Renewable Energy Laboratory*, *February*.

Funcke, S., & Bauknecht, D. (2016). Typology of centralised and decentralised visions for electricity infrastructure. *Utilities Policy*, *40*.

Goldkuhl, G. (1997). *METHOD INTEGRATION AS A LEARNING PROCESS*. 15–26.

Gordijn, J. (2002). E3-value in a Nutshell. *International Workshop on Ebusiness Modeling*, 1–12.

Gordijn, J., Osterwalder, A., & Pigneur, Y. (2005). Comparing two Business Model Ontologies for Designing e-Business Models and Value Constellations. *18th Bled EConference - EIntegration in Action, Osterwalder 2004*, 1–17.

Hellström, M., Tsvetkova, A., Gustafsson, M., & Wikström, K. (2015). Collaboration mechanisms for business models in distributed energy ecosystems. *Journal of Cleaner Production*, *102*, 226–236.

Hevner, March, Park, & Ram. (2004). Design Science in Information Systems Research. *MIS Quarterly*, *28*(1), 75.

Iansiti, M., & Levien, R. (2004). Strategy as ecology. *Harvard Business Review*, *82*(3), 68–78, 126.

Johannesson, P., & Perjons, E. (2014). An Introduction to Design Science. In *Springer International Publishing Switzerland*.

Jonda, M. (2004). *Szenario-Management digitaler Geschäftsmodelle*. Oldenburg.

Kartseva, V., Gordijn, J., & Tan, Y.-H. (2004). Value-Based Business Modelling for Network Organizations : Lessons Learned From the Electricity Sector. *ECIS 2004*.

Koppenhoefer, C., Fauser, J., & Hertweck, D. (2017). Managing and Controlling Decentralized Corporate Energy Systems - Transferring Best-practice Methods to the Energy Domain. *Proceedings of the 19th International Conference on Enterprise Information Systems*, 532–540.

Koppenhoefer, C., Fauser, J., & Hertweck, D. (2018). Multi-Model-Approach Towards Decentralized Corporate Energy Systems. In B. Otjacques, P. Hitzelberger, S. Naumann, & V. Wohlgemuth (Eds.), *From Science to Society: New Trends in Environmental Informatics* (pp. 117–128). Springer International Publishing.

Köster, O. (2014). *Systematik zur Entwicklung von Geschäftsmodellen in der Produktentstehung*. Universität Paderborn.

Küller, P., Hertweck, D., & Krcmar, H. (2015). Energiegenossenschaften - Geschäftsmodelle und Wertschöpfungsnetzwerke. In A. Zimmermann & A. Rossmann (Eds.), *Digital Enterprise Computing 2015 (LNI)* (pp. 15–26). Gesellschaft für Informatik.

Magnusson, T., Tell, F., & Watson, J. (2005). *From CoPS to mass production ? Capabilities and innovation in power generation equipment manufacturing*. *14*(1), 1–26.

Meyer, D., Fauser, J., & Hertweck, D. (2021). BUSINESS MODEL TRANSFORMATION IN THE GERMAN ENERGY SECTOR: KEY BARRIERS AND DRIVERS OF A SMART AND SUSTAINABLE TRANSFORMATION PROCESS IN PRACTICE. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *VIII-4/W1-2021*, 73–80.

Mirata, M., Nilsson, H., & Kuisma, J. (2005). *Production systems aligned with distributed economies : Examples from energy and biomass sectors*. *13*, 981–991.

Moore, J. F. (1996). *The Death of Competition: Leadership and Strategy in the Age of Business Ecosystems*. HarperCollins.

Osterwalder, A., Pigneur, Y., Bernarda, G., & Smith, A. (2014). *Value Proposition Design: How to Create Products and Services Customers Want*. John Wiley & Sons.

Osterwalder, A., Pigneur, Y., Smith, A., & Movement, T. (2010). Business Model Generation. In *Booksgooglecom* (1st ed., Vol. 30, Issue 5377). John Wiley & Sons.

Peppard, J., & Rylander, A. (2006). From Value Chain to Value Network:. Insights for Mobile Operators. *European Management Journal*, *24*(2–3), 128–141.

Richter, M. (2012). Utilities' business models for renewable energy: A review. *Renewable and Sustainable Energy Reviews*, *16*(5), 2483–2493.

The Paris Agreement. (2015). https://unfccc.int/files/essential_background/convention/application/pdf/english_paris_agreement.pdf

Thompson, J. D., & MacMillan, I. C. (2010). Business models: Creating new markets and societal wealth. *Long Range Planning*, *43*(2–3), 291–307.

Valocchi, M., Juliano, J., & Schurr, A. (2014). Switching perspectives: Creating new business models for a changing world of energy. *Green Energy and Technology*

Wirtz, B. W., & Daiser, P. (2018). Business Model Innovation Processes: A Systematic Literature Review. *Journal of Business Models*, *6*(1), 40–58.

Wirtz, B. W., Pistoia, A., Ullrich, S., & Göttel, V. (2016). Business Models: Origin, Development and Future Research Perspectives. *Long Range Planning*, *49*(1), 36–54.

Yin, R. (2009). *Case study research: Design and methods* (4th ed.). Sage.

Systems for Smart Mobility

Boris Tomaš, Lovro Posarić
WiFi access point localization in urban multi-sensor environment

Filip Ribić, Tomislav Erdelić, Tonči Carić, Martina Erdelić
Solving Capacitated Time-Dependent Vehicle Routing Problem

Filip Vrbanić, Edouard Ivanjko, Slaven Gašpar
Virtual Reality and Transport: A Short Review

WiFi access point localization in urban multi-sensor environment

Boris Tomaš, Lovro Posarić
Faculty of Organization and Informatics, IOT Laboratory
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{btomas, lposaric}@foi.hr

Abstract. *Urban sensor networks produce large amount of data including available WiFi networks information and current GPS location. This work focuses on finding usable information in this WiFi "noise" in urban environments, envisioning the new WiFi access point localization mapping technique. Main challenges of such service are identified and their solution is proposed as an application with specific algorithm for WiFi access point localization. The application architecture is a simple client server model with external actors responsible for data acquisition. Source of the data is an urban sensor network, which is sort of wireless network specifically set up in urban mobile environments.*

Keywords. trilateration, localization, mapping

1 Introduction

In an urban multi-sensor environment there is a large amount of gathered data. After the initial gathering stage, these data need to be synchronized with a central database at the end of data gathering, but ideal would be real time data synchronization. However, for the real time synchronization stable and constant Internet access is required. In urban environments there are many open access WiFi access points (hotspots) that provide free Internet access, location of those devices is fairly constant, exceptions are mobile hotspots on mobile devices and in special occasions dedicated hotspots for conferences, fairs, exhibitions,... Location of fixed hotspots is mostly unknown and it is difficult to determine location of hotspot with high precision because hotspots are mostly in private homes and apartments. Localizing hotspot would mean that someone should visit every home in an area and interview the residents whether they have or don't have WiFi hotspot. Hence, different method needs to be used in order to create a localization mapping of available access points.

2 Problem

Mobile object during movement in the urban sensor environment may require stable and constant internet access using existing WiFi infrastructure. Problem is that this infrastructure is not know. Location of the WiFi access point (AP) needs to be known, but APs themselves don't advertise their exact location. On the other hand, urban multi-sensor networks (UMN) like one described in (Rodrigues, Aguiar, Vieira, Barros, & Cunha, 2011) collect various data types like GPS location, velocity, acceleration, time, WiFi information (SSID, MAC address, signal and noise strength,...) and many more. UMN are usually implemented in the urban environments that do have many WiFi hotspots deployed around the area. UMN provide valid data input for the localizing AP in urban environments. The end goal is to create the mapping of access point locations present in the UMN.

Application that may use localization of AP is the assignment of WiFi AP to the roaming object in UMN (or any other network). This assignment would be necessary because of the offloading existing mobile data networks networks with WiFi is more efficient (Deshpande, Hou, & Das, 2010). Every time a moving object is changing its assigned AP it costs time because of authentication process and assignment procedure called handshaking which lasts for about 330ms (Park, Han, & Kim, 2009). During this time vehicle travelling 80km/h can travel up to 7.3m which can be crucial for some applications using urban sensor networks.

3 Related work

For the indoor localization, WiFi positioning techniques based on Wireless Local Area Network (WLAN) or Wireless Personal Area Network (WPAN) is commonly employed presently (Mok & Retscher, 2007). Given the ubiquity of WiFi networks and WiFi access points, as well as the fact that nearly all smartphones nowadays have a built-in WiFi module, WiFi positioning has become a prominent tool for indoor po-

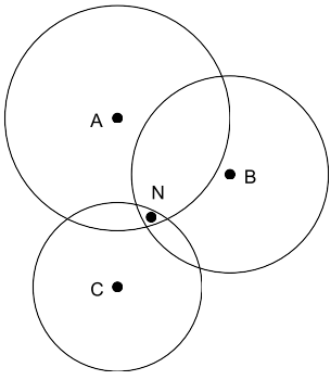


Figure 1: Circle intersections with clustering (Kaminsky, 2007)

sitioning. In open environments the most famous positioning technology is the GPS/GLONASS positioning. It is only even more reinforced by the fact that it does not require additional special-purpose hardware and to top that off, location estimation can be easily attained by measuring the received signal strength (RSS) from a Wi-Fi access point (Li et al., 2019). This signal strength can be converted to the distance which can represent circle radius, more on this later on. Alternative to the signal strength is the use of radio signal propagation over time (Mirisola, 2003). As an approach based on signal strength, this technique is also not bullet proof from reflections and multipathing. Time measurement of signal propagation requires very precise measurement equipment that can not be found in the mainstream WiFi systems. Distance, using time, can be calculated by multiplying the speed of light with measured timespan, this distance can also represent the radius of eventual coverage estimation circle In both cases WiFi AP position can be determined using relatively easy calculations like the special case of multilateration - trilateration: "The trilateration problem is to find the coordinates of node $N = (nx, ny)$ from the given information. A complicating factor is that the known nodes' coordinates and distances typically include measurement errors. Two methods of solving the trilateration problem are nonlinear least squares and circle intersections with clustering." (Kaminsky, 2007) Main prerequisite for the good trilateration is to have a minimum of 3 non linear sources (Konrad & Wölfel, 2012). In urban multi-sensor environments, like urban travel trajectory WiFi scans, this requirement is not always achievable.

In Figure 1, three points can be seen: A, B and C, each of those points have a circle around them. Trilateration procedure is based on solving the equation set of 3 circle equations. The ideal result would be a single point, however in real-time scenarios result usually is a polygon and location point (N) is the centroid of resulting polygon.

$$(x - x_1) + (y - y_1) = r_1^2 \quad (1)$$

$$(x - x_2) + (y - y_2) = r_3^2 \quad (2)$$

$$(x - x_3) + (y - y_3) = r_3^2 \quad (3)$$

It is required that at least three circles are used in the calculation. If only two circles are used, it will be accurate in one dimension (the distance between the APs), but won't be able to accurately detect the location in the second dimension. On the flip side, more than five or six APs could limit the effectiveness by adding unnecessary noise and interference in the environment. In these related cases, trilateration is used to determine a location of a object in space (open or closed) using the fixed location sources' radiation. In this paper, trilateration is used in a case where the object location is relatively known (using GPS), but the location of radiating WiFi hotspots is unknown. Basically it is the same principle, but in a reverse direction.

3.1 WiFi AP localization

Each moving object scans its surroundings for the existing WiFi networks. For each WiFi AP scan, relevant data is gathered:

- GPS *location*;
- MAC address;
- SNR Signal Noise Ratio in dBm.

Procedure follows the case shown in Figure 1 where for each scan circle the coverage area is calculated. Assume that A, B and C are the GPS location of the moving object. Radius represents the distance to the probable WiFi location. Radius (r) is calculated using *free-space path loss* formula:

$$r = 10^{\frac{\ln(10)(T-R-K-20\log(f))}{10n}}$$

where:

- R - receive power level, this value is being gathered using WiFi scanning equipment,
- T - transmit signal power strength is 17dBm. According to IEEE 802.11b standard[1] the maximum power output level is 20dBm and the minimum gives 13dBm, this gives average of 17dBm, while the actual value depends on the firmware used by the router that powers the access point,
- K - path loss constant, value is -147.55,
- f - the WiFi frequency and is set to 2450 MHz,
- n - the path loss exponent, it is set to 3 because we are measuring distances in the highly urban environment. Typically used values are: 2 for free space, 2.7 to 3.5 for urban areas, 3.0 to 5.0 in suburban areas and 1.6 to 1.8 for indoors when there is line of sight to the router,

[1] `http://www.cisco.com/en/US/docs/optical/15000r7_0/15327/reference/guide/2770spcx.html`

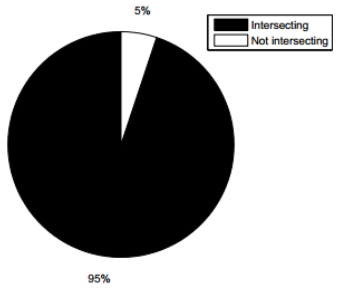


Figure 2: Circle intersection states for assigned AP

- r - is the distance (radius), the result in meters.

After the radius has been calculated, a circle around each moving object can be constructed. This circle represents "how well does an object hear a WiFi AP" not the location estimation of a WiFi AP. To determine the AP location it would be necessary to undertake the trilateration procedure. Figure 1 shows a case with 3 scan points from moving object each with its own radius. First it would be good to investigate the correlation between circles defined by point A, B and C and the appropriate radii. Possible correlation states are:

- One point intersection: This would be the best case scenario but it is rare;
- Polygon intersection: This is the most common case where the intersection is a polygon like shown in Figure 1;
- No intersection. If there is no intersection, it is not possible to carry out trilateration between the three circles as the mutual AP can only exist in the intersection

Figure 2 shows the current state of intersections circles in the system. To simplify, states 1 and 2 make no difference for the trilateration procedure. Important is the existence or no existence of intersection.
If there is an intersection then probable location of AP is the centroid of intersecting polygon (or point). In Figure 1 this is the point N.

Trilateration starts from the starting circle and then calculates intersection with all the other circles. For each circle the intersection is calculated between the circle and the intersection itself. In turn, this makes the intersection (polygon) shrink with every circle considered.
Problem is how to select the starting circle if there are some circles that do not intersect with any other circle, or there may be a clusters of intersecting circles that may result in two or more intersecting polygons. This paper identifies two strategies:

- Most Intersections First (MIFS). This strategy selects the starting circle by selecting one with the most intersections. This means that for each circle it should be calculated with how many other circles it intersects.
- SIF: Smallest Circle First (SIF): This strategy selects the starting circle as one with the smallest radius (strongest signal) with at least one possible intersection.

Evaluation of each strategy will be done later. After the intersection (polygon) is found, then centroid of a polygon is the WiFi AP location estimation. Next, it would be necessary to define the radius of WiFi AP coverage estimation. Radius is calculated using the same path-loss equation (3.1) Input signal strength in this case is the lowest signal strength of the scan circle that contributed in the final intersection forming.

4 Implementation

This procedure follows following steps:

- Calculate the distance (*radius*) using the *path loss* equation defined before: 3.1 using scanned SNR as input.
- Create 12 side polygon with the centre in *location* point and *radius* (distance from centre to the middle of a polygon side). This polygon represents a circle[2] inside which there is probably a WiFi AP that is being scanned.
- Get the list of all WiFi AP that are scanned during trip.
- For each WiFi AP get all the scan circles; and for each circle calculate count of intersections it has with all other circles for the current WiFi AP.
- Select the starting circle using different algorithm variations defined before: (see: 3.1) *MIFS* or *SRF* Starting circle is intersecting polygon (I_p)
- For each circle for the current WiFi AP get the intersection with I_p. Intersection result is stored in the same variable I_p. When all circles are iterated then result is a polygon with the most probable WiFi AP location. Figure 3 shows the result of this step. Blue circles from A to F are scan circles, intersection of those circles is shown as the red polygon "I".
- Centroid of I_p represents the most probable location of current WiFi AP location.
- Find the weakest signal circle, one that is the largest and that forms the intersecting shape that is assigned to this WiFi AP. Radius of this circle is used as a coverage radius for the WiFi AP. Retrieving distance using path-loss model using SNR of weakest signal circle would be redundant because this has been already done. Coverage area is represented as a circle that is also a 12 side

[2] it is referred to as a circle although it is a dodecagon.

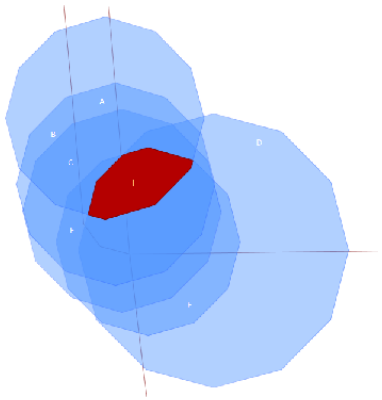


Figure 3: Intersecting polygon/shape (I_P)

polygon that has a radius equal to the radius of the weakest circle and center is equal to the centroid location of I_P.

Database used for this work is PostgreSQL with PostGIS extensions for geo-spatial functions.This implementation uses the geometrical intersection function of underlying database. It is used instead of custom method being implemented that only solves the set of two circle equations mathematically, because, in future work, scan area does not need to be a circle at all, it can be any shape. This way, described algorithm and implementation will work even if the scan shape is not the circle but some complex polygon.

5 Results

There are total of 4 algorithm variations (A):

- MIFS epe+ (Most intersections first with the GPS error over $10m$);
- SRF epe+ (Smallest radius firs with the GPS error over $10m$);
- MIFS epe- (Most intersections first with the GPS error below $10m$);
- SRF epe- (Smallest radius first with the GPS error below $10m$);

Analysis of algorithm variations is done on a test case scenario[3] where the real location of 35 WiFi AP is known.

5.1 Metrics and validation

In the test case scenario for each AP there are 4 locations based on 4 algorithm variations. Distance between each variation and real location is calculated: Distance measurements (d) are:

[3]Grounds of Campus S.João in Porto, Portugal

$d_1 = d(A_1, A_2)$	Where Ai is:
$d_2 = d(A_1, A_3)$	$A_1 = MIFS\ epe+$
$d_3 = d(A_1, A_4)$	$A_2 = SRF\ epe+$
$d_4 = d(A_2, A_3)$	$A_3 = MIFS\ epe-$
$d_5 = d(A_2, A_4)$	$A_4 = SRF\ epe-$
$d_6 = d(A_3, A_4)$	
$d_7 = d(R, A_1)$	
$d_8 = d(R, A_2)$	
$d_9 = d(R, A_3)$	
$d_10 = d(R, A_4)$	

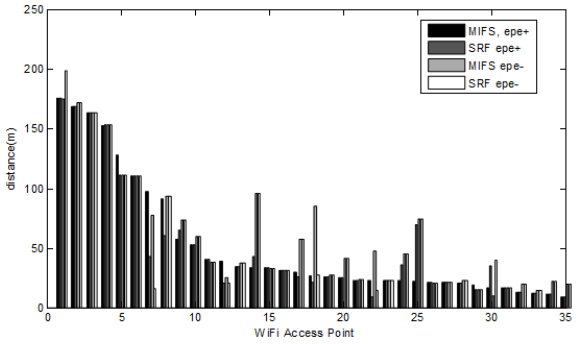


Figure 4: Distance to the real location by algorithm type

Figure 4 shows the distances from the estimated location (for each algorithm combination) to the WiFi AP real location. Test scenario is made of 35 WiFi AP and it can be seen that the distance difference is mostly insignificant (few meters), but there are "spikes" which indicate that the algorithms 3 and 4 (MIFS epe- and SRF epe-) provide worse results than algorithms with GPS error. Figure 5 shows the number of measurements taken for each WiFi AP. Total scans line shows the total number of scans taken, *Total-snr5%* shows the number of scans without 5% of WiFi scans with lowest signal strength and *Total-snr5%-epe* is the number of scans without 5% of the worst signal and without scans with the GPS error above 10m. Each "spike" from Figure 4 is backed by the spike in Figure 5 which means that in the case of exception there is a huge number of scans being measured. On the other hand it can be seen

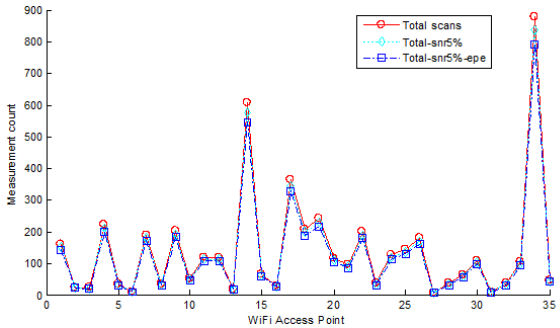


Figure 5: Measurements count for each WiFi AP

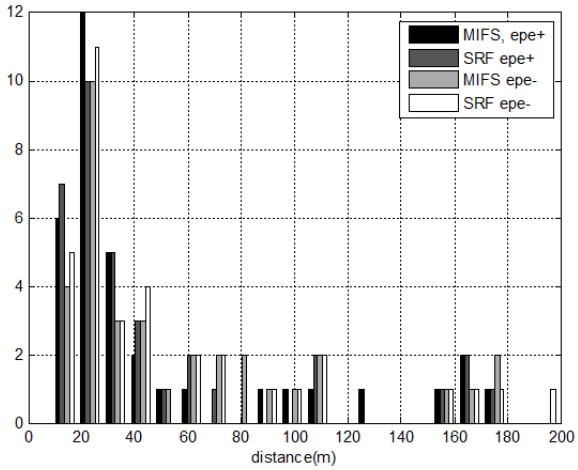


Figure 6: Histogram of distances from estimated to real location by algorithm variations

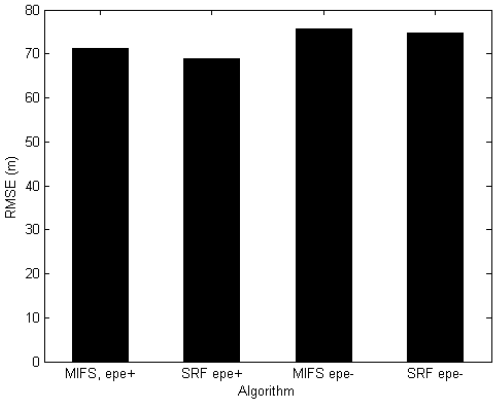


Figure 7: Root-mean square error for location estimation using algorithm variations

that the number of measurements(scans) do influence deviation of estimated location to the real location. Although, there is no rule except that in case of high distance deviation number of scans is low (less than 210 scans), on the other hand small deviation (high precision) does not necessary need to be backed with large amount of measurements(scans). There is, however, an interesting case of AP 34 where distance deviation is small 20m and measurements count is very high, around 850 measurements. Also it can be seen that the algorithms with GPS error produced more precise results in this case. Which may be interpreted as that the GPS error is averaged out by the large number of measurements. Algorithms with GPS error produce better results, likely due to the number of scans undertaken.

Finally, Figure 6 shows the histogram of mentioned test case. It can be seen that the algorithm variations with GPS error included do produce better results regarding distance of the estimated location to the real location.

Figure 7 shows root-mean squared error that is calculated as distance between the estimated location and the real location. It shows that SRF with GPS error does produce better location estimation.

6 Conclusion

This paper describes solutions to the problem of crowsourcing data in urban sensor environment and localization of access points in space, however there are different approaches that may provide different results. As a part of the future work one such approach is the probabilistic approach; instead of geometric analysis, calculation of probability is done using input data. It would be interesting to see the difference between geometric and probabilistic location estimation. Part of the future work is a final implementation of this solution; it should be deployed as a web service that can be used in many applications.

Acknowledgments

This work has been fully supported by Croatian Science Foundation under the project IP-2019-04-4864.

References

Deshpande, P., Hou, X., & Das, S. R. (2010). Performance comparison of 3G and metro-scale WiFi for vehicular network access. *Proceedings of the 10th annual conference on Internet measurement - IMC '10*, 301. Retrieved from `http://portal.acm.org/citation.cfm?doid=1879141.1879180` doi: 10.1145/1879141.1879180

Kaminsky, A. (2007). Trilateration. , *1*(2), 8–11.

Konrad, T., & Wölfel, P. (2012). WiFi Compass. , *2012*.

Li, S., Hedley, M., Bengston, K., Humphrey, D., Johnson, M., & Ni, W. (2019, 03). Passive localization of standard wifi devices. *IEEE Systems Journal*, *PP*, 1-4. doi: 10.1109/JSYST.2019.2903278

Mirisola, L. G. B. (2003). The Localization Problem on Sensor Networks.

Mok, E., & Retscher, G. (2007). Location determination using WiFi fingerprinting versus WiFi trilateration. *Journal of Location Based Services*, 1–15. Retrieved from `http://www.tandfonline.com/doi/abs/10.1080/17489720701781905`

Park, H.-S., Han, S.-H., & Kim, J.-D. (2009). Vehicular client roaming and location-based

handoff through multiple WLAN APs in a container terminal. *Proceedings of the 2009 International Conference on Hybrid Information Technology - ICHIT '09*, 465–472. Retrieved from http://portal.acm.org/citation.cfm?doid=1644993.1645079 doi: 10.1145/1644993.1645079

Rodrigues, J. G. P., Aguiar, A., Vieira, F., Barros, J., & Cunha, J. P. S. (2011, October). A mobile sensing architecture for massive urban scanning. *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 1132–1137. Retrieved from http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6082958 doi: 10.1109/ITSC.2011.6082958

Solving Capacitated Time-Dependent Vehicle Routing Problem

Filip Ribić, Tomislav Erdelić, Tonči Carić, Martina Erdelić
Faculty of Transport and Traffic Sciences, Department of Intelligent Transport Systems
University of Zagreb
Vukelićeva 4, 10000, Zagreb, Croatia
{fribic, terdelic, tcaric, merdelic}@fpz.unizg.hr

Abstract. *A vehicle routing problem aims to determine a set of vehicle routes to deliver goods to the customers. Most of the variants of the problem consider static traffic conditions, while in the real world, travel times depend on the departure time. In this paper, a capacitated time-dependent vehicle routing problem is observed, which considers time-dependent travel times between the customers, and delivery with vehicles that have limited load capacity. To solve the problem, an iterative local search metaheuristic is applied that couples the ruin-recreate principle with the common local search procedures. Several scenarios were developed to analyze the impact of time-dependent travel times on the solution quality in terms of the total travel time and customer configuration.*

Keywords. time-dependent vehicle routing problem, total distance traveled, total time, heuristics

1 Introduction

Delivering goods to customers is a complex logistic process. In it is important to deliver the goods on time and meanwhile to reduce the total delivery costs in terms of total travel time and total traveled distance (Toth and Vigo, 2001). The key question in such process is to determine which routes to choose and how different route choices affect the change in total travel time and total distance traveled. Vehicles used in such processes are often limited by the load capacity in terms of vehicle cargo volume or maximum vehicle load-bearing. Another important aspect of the delivery are traffic conditions, or more precisely time spent on the route. Due to the traffic conditions, route travel time depends on the departure time, especially in large urban centres where traffic congestions occur typically in rush periods (Erdelić et al., 2021). As traffic congestions significantly increase the travel time, the total route travel time can change significantly, and thus increase the total delivery costs (up to 3%), (Kellner, 2016).

In the literature, the well-known problem of delivering goods is modelled as NP-hard Vehicle Routing Problem (VRP), (Toth and Vigo, 2001). The goal of the VRP is to determine the least-cost delivery routes from a warehouse to customers, subject to side constraints (Laporte, 1992). These side constraints often resemble real-life limitations of the problem: limited vehicle load capacity (Capacitated VRP, CVRP), pickup and delivery of goods (VRP with Pickup and Delivery, VRPPD), customer delivery time windows (VRP with Time Windows, VRPTW), heterogeneous fleet, etc., (Erdelić and Carić, 2019). Driven by the impact of traffic congestions on routing decisions, a Time-Dependent VRPW (TD-VRPTW) was formulated (Figliozzi, 2012), which considers time-dependent travel times on the road network. The delivery period is discretized into smaller time periods, usually representing rush and non-rush periods. In non-rush periods typically, the free flow speed is considered, while in the rush periods, the speed significantly decreases. These time periods and speeds are often determined by knowledge. On the other hand, there are research papers that present the exact time periods and speed values derived from large historical vehicle tracking data (Erdelić et al., 2021).

In this paper, the Time-Dependent VRP (TD-VRP) is observed, which can be considered as a subproblem of both TD-VRPTW and CVRP. In this problem, vehicles are limited by their load capacity, travel at time-dependent travel times and do not have to visit a customer within its time window (only the warehouse working times are respected). As VRP is an NP-hard problem, the TD-VRP is also NP-hard. This means that exact procedures are only capable of solving problems with a relatively low number of customers (up to 50 customers), and most often, the metaheuristics and heuristics are applied to solve larger problems (Vidal et al., 2013). Over the years, a vast number of heuristic and metaheuristic procedures have been applied to efficiently solve the VRP problems (Erdelić and Carić, 2019). Usually, the best results are achieved with a metaheuristic procedure that guides the search, such as tabu search, simulated annealing, large neighborhood search, variable neighborhood search, genetic al-

gorithm, ant colony, iterative local search (Vidal et al., 2013). The problem-specific heuristic procedures that intensify the search are used within the metaheuristic procedure, such as local search operators (relocate, exchange, 2-Opt, etc.), nearest neighbor heuristic, Clark-and-Wright savings method, etc., (Erdelić and Carić, 2019).

Compared to the related research papers of Figliozzi (2012); Carić and Fosin (2020) and Rožić et al. (2015), this paper is one of the few that considers the TD-VRP problem without time windows and its influence on the total route duration and customer configuration; thus, investigating the real-life problems in which customers do not have time windows, and can be visited at any time during the day. The applied Iterative Local Search (ILS) heuristic to solve the problem has not yet been applied on such problem. Additionally, we investigated the impact of different speed values in different time periods on solution quality in the form of several specific scenarios.

The rest of the paper is organized as follows. The description of the problem, together with the test instances, are presented in Section 2. The solution method used to solve the problem is described in Section 3 together with the heuristic for the initial solution, Local Search (LS) operators and ILS. The results of the paper are presented in Section 4. Finally, the conclusion of the paper is given in Section 5.

2 Problem description

In this section, a TD-VRP problem modelled as a Mixed Integer Program (MIP) is described. Additionally, an example of test instances used to test the solution method is presented.

2.1 MIP model

TD-VRP is formulated as a MIP model by Equations 1-7 with the goal of minimizing the overall total travel time. Let $V = \{1, \ldots, N\}$ be a set of geographically scattered customers that need to be served. Vertices 0 and $N + 1$ denote the warehouse instances, and every route begins at vertex 0, and ends at vertex $N + 1$ ($V_{0,N+1} = V \cup \{0\} \cup \{N + 1\}$). Graph G is defined as $G = (V_{0,N+1}, A)$, where A is the set of arcs $A = \{(i, j)|i, j \in V_{0,N+1}, i \neq j\}$. The binary variable $x_{ij} \in \{0, 1\}$ (Equation 1) is equal to 1 if arc (i, j) is traversed in the solution, and 0 otherwise. Each vehicle has a load capacity C. Each customer i has a service time s_i and load demand q_i. The warehouse has a working time $[0, l_0]$. Additionally, two more decision variables for customers ($i \in V_{0,N+1}$) are used: τ_i - begin time of service and u_i - remaining load capacity. The arc value $t_{ij}(x)$ represents the time function that determines time needed to traverse the arc with the departure time x. The objective function is given by Equation 2 as the sum of travel times per arcs.

Equations 3 and 4 ensure the arc connectivity of customers, meaning that each customer can have only one entry and one exit arc which results in the constraint that each customer has to be visited only once. Equation 5 ensures travel time feasibility of arcs between customers i and customer j. If arc (i, j) is traversed, then the begin time at customer j, τ_j, has to be equal to the sum of begin time at customer i, τ_i, time-dependent travel time between i and j (depends on the departure time) $t_{ij}(\tau_i + s_i)$, and service time of customer i, s_i. Equation 6 ensure arcs load flow, in similar way as equation 5. Equation 7 ensures that the leaving warehouse instance has a remaining load capacity equal to the vehicle load capacity C.

$$x_{ij} \in \{0, 1\},\ \forall i \in V_0, j \in V_{N+1}, i \neq j \tag{1}$$

$$\min \sum_{i \in V_0} \sum_{j \in V_{N+1}, i \neq j} t_{ij}(\tau_i + s_i) x_{ij} \tag{2}$$

$$\sum_{j \in V_{N+1}, i \neq j} x_{ij} = 1,\ i \in V \tag{3}$$

$$\sum_{i \in V_{N+1}, i \neq j} x_{ji} - \sum_{i \in V_0, i \neq j} x_{ij} = 0,\ j \in V \tag{4}$$

$$\tau_i + (t_{ij}(\tau_i + s_i) + s_i) x_{ij} - l_0 \cdot (1 - x_{ij}) = \tau_j, \quad \forall i \in V_0, \forall j \in V_{N+1}, i \neq j \tag{5}$$

$$0 \leq u_j \leq u_i - x_{ij}(q_i + C) + C, \quad \forall i \in V_0, \forall j \in V_{N+1}, i \neq j \tag{6}$$

$$u_0 = C \tag{7}$$

The time of one working day is discretized in different time periods in order to take into account traffic congestions that occur in the so-called rush hours. This means that the function of travel time on arc $t_{ij}(\tau_i + s_i)$ is discretized into a finite number of time intervals $k \in \{T_0, \ldots, T_K\}$, and to each time period a static speed v_k is assigned. The problem considers that the distance on arc d_{ij} does not change, and the travel time on arc in time period k can be computed as $t_{ij}(k) = d_{ij}/v_k$. The departure time customers i expressed as $\tau_i + s_i$ is used to determine the period k in which the vehicle starts to travel from customer i to customer j.

2.2 Data

To test their methods and compare them to other methods, researchers typically solve the benchmark instances. In this paper, one instance from the well-known Solomon VRPTW test instances is used (Solomon, 1987). Each customer in instance has a location $C(x, y)$ in the Cartesian coordinate system, service time and demand. The warehouse working time consist of opening (0) and closing time (l_0). The distance between the customers is computed by Euclidean

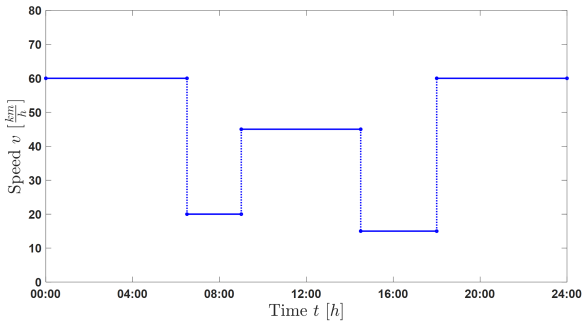


Figure 1: Speeds

distance given by Equation 8. The warehouse working time $[0, l_0]$ is discretized into five-time buckets of equal duration: $[0, 0.2l_0\rangle$, $[0.2l_0, 0.4l_0\rangle$, $[0.4l_0, 0.6l_0\rangle$, $[0.6l_0, 0.8l_0\rangle$, and $[0.8l_0, l_0\rangle$. In periods of congestion, speeds are reduced in order to better represent the actual state of the traffic system, while in other intervals, speeds are closer to the free flow speed. The example of speeds per discretized intervals for one path in one working day is presented in Figure 1. In total, five time periods are observed on the x-axis with corresponding speed values on the y-axis. As it can be seen, during the night and evening, a free-flow speed with the value of 60 km/h is present, while during the morning and afternoon rush hours, the speed drops to 20 km/h, and 15 km/h respectively. Between the rush hour periods, the speed is somewhere in between with the value of 45 km/h.

Used instance contains a subset of 25 customers presented in Figure 2. The customers are presented with black circles, where the size of the circle represents the customer demand, the higher the demand is, the larger the circle is. The warehouse is presented as a red circle. The name of the instance $C101$ indicates that customers are groupped in clusters.

$$d(C_1, C_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (8)$$

The values of distances and travel times on arcs (paths) between customers are commonly stored in 2D matrices, representing the shortest path between each

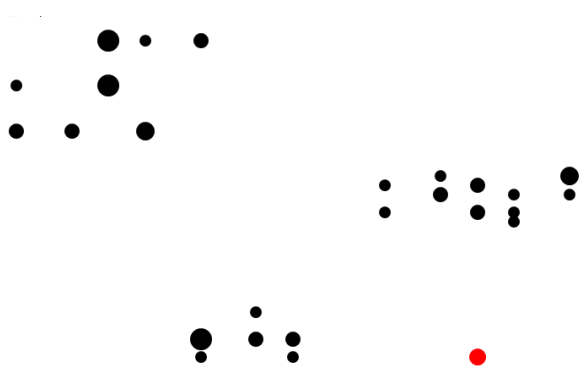


Figure 2: Instance C101_25

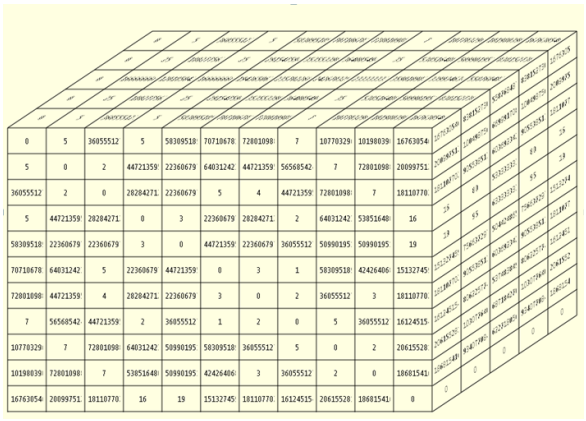


Figure 3: 3D travel time matrix

customer pair (Erdelić et al., 2021). Such matrices are computed in a preprocessing step to reduce the computation time, as computing the shortest path between customer pairs is time-consuming. Here, as the Euclidean distance was used, there is no need for the computation of the matrices, but for the sake of the complicity, the matrices will be used. Including time-dependent travel times results in the 3D travel time matrix, where the third dimension represents the number of discretized time periods. The example of a 3D travel time matrix is presented in Figure 3. For example, a value of five in the first row and second column represents a travel time of five between the warehouse and the first customer in first time period.

3 Methodology

The methodology section consists of three parts: (i) creation of the initial solution, (ii) LS procedure, and (iii) ILS procedure. All applied heuristic methods search in the feasible solution space, which is constrained by the vehicle load capacity and by the total route duration due to the warehouse working hours.

3.1 Initial solution heuristic

The heuristic used for the creation of the initial solution is based on the time-oriented nearest neighbour heuristic (Solomon, 1987). The heuristic in each iteration selects the customer that is the closest to the current customer in terms of the travel time.

The procedure for creating an initial solution is presented by the Algorithm 1. The input to the algorithm is a list of all customers V. First, all customers, except the warehouse, are added to the list of unvisited customers. Next, a new vehicle route is opened, and a warehouse is added to it. Further on, the algorithm searches through all unvisited customers for the unvisited customer who has the lowest arrival time and can be added to the vehicle route, respecting the load capacity of the vehicle and total route duration (enough

Algorithm 1 Heuristic for the creation of the initial solution

Input: List of customers V

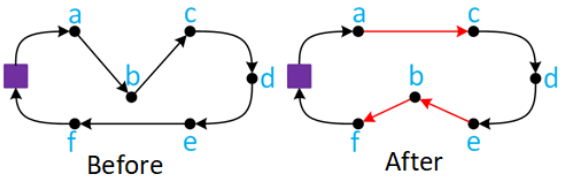
1: Add all customers V, except warehouse, to the list of unvisited customers
2: Open a new route, add warehouse to it and set the warehouse as the current customer
3: **while** there are no unvisited customers **do**
4: Find the timely closest unvisited customer c_{clos} to the current customer
5: **if** load capacity and the total duration of the vehicle route is not violated by adding a customer c_{clos} **then**
6: Add customer c_{clos} to the current vehicle route
7: Remove customer c_{clos} from the list of unvisited customers
8: Reduce the available load capacity of current vehicle
9: **else**
10: Close the current vehicle route by adding a warehouse at the end
11: Open a new route, add warehouse to it and set the warehouse as the current customer
12: **end if**
13: **end while**

time to return to the depot) is not exceeded. In case in which the observed customer exceeds the capacity of the vehicle or the total route duration, the vehicle is returned to the warehouse, and a new vehicle is opened. This means that the warehouse is added after the last visited customer and a new route with a warehouse is opened. By adding a customer to the vehicle, the capacity of the vehicle increases by it's demand. If a customer can be added to the vehicle route, it is removed from the list of unvisited customers and saved to the list of visited customers of the current vehicle. The process is repeated until all customers are visited. The final result of the initial solution is the number of vehicles needed to visit all customers, the configuration of customers per vehicle routes with the total traveled time and the total traveled distance.

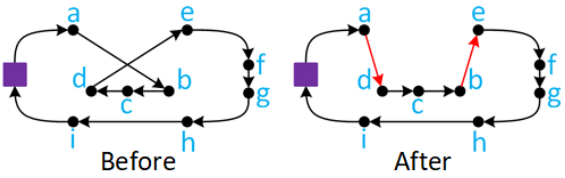
3.2 LS

The created initial VRP solution is often 20-30% far from the optimal solution to the problem (Erdelić and Carić, 2019). By searching the neighborhood of the current (initial) solution, a better solution can be found, commonly called local optima. The local optima is obtained by operators that calculate the difference in total travel time made by vehicles after switching customers among themselves or by inserting and removing customers from certain vehicle routes, but only if:

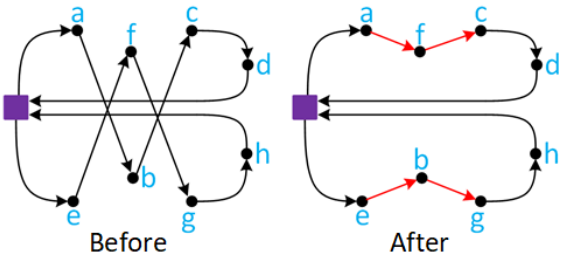
(i) a feasible solution regarding the load capacity and warehouse working hours is produced,



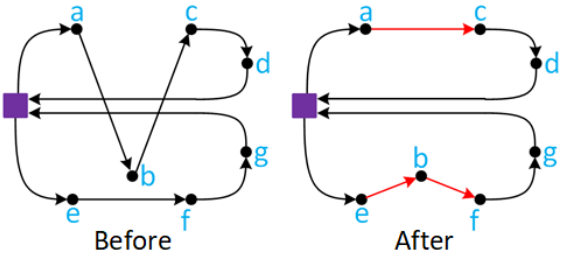
(a) Intra relocate



(b) Intra 2-Opt



(c) Inter exchange



(d) Inter relocate

Figure 4: LS operators

(ii) saving in the total travel travel time objective is achieved.

Depending whether the changes are performed on a single route or between multiple routes intra and inter LS operators are used. In total, four LS operators are applied, presented in Figure 4. After applying each of the operator on the current solution, the local optima in search composite neighborhood is found.

3.2.1 Intra operators

Intra operators are used to change the order of customers within the same route. The vehicle capacity limit with the intra operators does not need to be checked due to the rotation of the customer within the same route, so the capacity requirement remains unchanged.

Intra relocate is an operator that moves a customer from one position in a route to another position within the same route. As presented in Figure 4(a) three arcs in the same route are removed, and three new red ones are inserted. Time savings is defined as the difference between the total travel time before the change and after the change.

Intra 2-Opt operator is used to remove the arcs intersections within the route. As shown in Figure 4(b) two arcs that intersect are removed, and two new red ones are inserted, resulting in a decrease of the total travel time. Additionally, the part of the route between the removed arcs needs to be reverted (instead of $b \rightarrow c \rightarrow d$, the reversed order is $d \rightarrow c \rightarrow b$).

3.2.2 Inter operators

Inter operators are used to move customers between different routes in the solution. It is mandatory to check the vehicle capacity limit and total route duration with the inter operators.

Inter exchange operator replaces positions of two customers in different vehicle routes. Replacement is only performed if the vehicle capacity limit is not exceeded and time savings are achieved. As shown in Figure 4(c) two customers are exchanged with in total four arcs removed and four red ones inserted (two per route).

Inter relocate operator switches customers from one position in the route to another position in another route. As shown in Figure 4(d) one customer is removed from one route to another route, meanwhile removing three arcs and inserting three red ones.

3.2.3 Search strategy

LS operators are called in the following order: *inter exchange*, *inter relocate*, *intra 2-Opt* and *intra relocate*. Such order is selected to first improve the solution by moving customers between different routes and then to improve the customer configuration within each route. This makes sense, as there is no point in the route improvement with intra operator if a particular customer will be removed from the route with inter operator as it leads to an overall better solution. The operators are applied in such order, as long as there was at least one improvement by either of the operators. This means that the loop for the LS procedure repeats until there is no improvement. Within a search of a single operator, a *best-of* technique is used, meaning that the best move from all the possible moves found by a particular operator is conducted.

3.3 ILS

LS operators explore the narrow solution space around the current solution, which often tends to lead to the same local optima. To escape the local optima and search in the other parts of the solution space, the ruin-recreate principle to diversify the search is applied. Here, a random removal and insertion of customers in the solution is used. The description of ILS metaheuristic based on the ruin-recreate principle is given by Algorithm 2. The input to the algorithm is the currently best solution s, achieved either by the initial construction heuristics or LS procedure. In each iteration, n_{EJ} percent of customers are randomly selected from the solution and removed. Random removed customers are tried to be reinserted at some positions in vehicle routes. When inserting customers in vehicles, it is mandatory to check the capacity limitations of the vehicle and total route duration. After the evaluation of insertions, if either constraint is violated, the insertion of customers in the solution will not be performed, and the procedure for insertion will repeat itself to at maximum n_{INS} iterations. If a feasible solution is found, a LS is performed in an attempt to find a better solution. During the algorithm execution, two solution instances are tracked: current best solution s and temporal solution s'. The ruin-recreate principle is always performed on the current best solution. To prevent the data loss regarding the configuration of current best solution s in case when the ILS was not able to find a better solution than the best so far, at the start of each iteration, the best solution s is copied to s'. In the case when the total travel time of solution s' is better than the total travel time of s, the configuration of s' is copied to s. The whole procedure of removal, insertion and improvement is repeated for n_{ILS} iterations.

Algorithm 2 ILS

Input: Best solution s

```
while repeat n_ILS iterations do
    Copy solution s to s'
    Randomly remove n_EJ percent of customers from the solution s'
    while repeat n_INS iterations do
        Determine random positions of customer insertions in vehicles' routes in solution s'
        if solution s' would be feasible then
            Perform determined insertions on solution s'
            Perform LS on s'
            if solutions s' is better than s then
                Copy solution s' to s
            end if
            Break loop
        end if
    end while
end while
```

Due to the random selection of customers' removals and insertions after each restart of the program, there will be a different solution that may be better than the first solution obtained. It is important to note that this procedure searches the solution space beyond the local

optima in order to find a better solution that may then be a globally optimal solution, but this is not guaranteed.

4 Results

In this section, the proposed methodology is evaluated on the test instance described in Section 2.2, and the analysis of time-dependent speeds is conducted. The parameters used for solving the problems are the following: $n_{ILS} = 30$, $n_{EJ} = 10$, $n_{INS} = 10\%$. The ILS is implemented as a single-thread code in the C# programming language. All tests were performed on a machine with Intel(R) Core(TM) i5-9300H CPU (2.24 Ghz) and 8 GB of RAM.

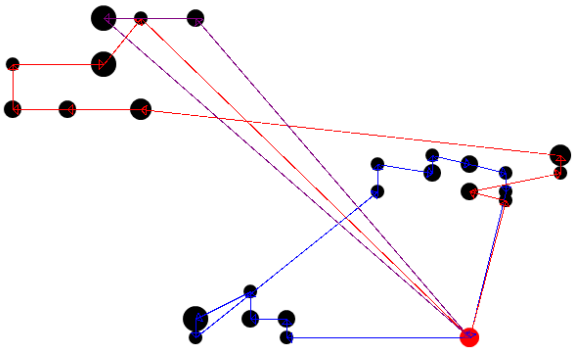
4.1 Evaluation of a search process

To solve the observed instance of the VRP, the discretization of time intervals and speeds is given by Table 1 (Scenario A). The warehouse working hours are set between 0 and $l'_0 = l_0/4$. The closing time is shortened compared to the l_0 value from the Solomon instance (Solomon, 1987) because the closing time l_0, originally used for the problem with time windows, is too large for the observed TD-VRP problem where time windows are not included. This means that all deliveries in TD-VRP routes are always done by the time l'_0.

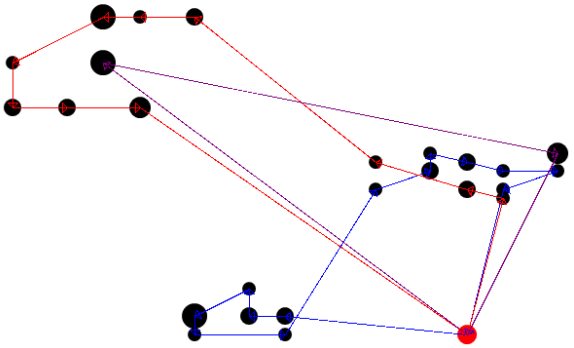
To show how the ILS works, the examples of solutions for scenario A, are presented in Figure 5. The metrics used to evaluate the solution are the Number of Vehicles (NV) in the solution, cumulative Total Travel Time (TTT) of all vehicles in the solution and cumulative Total Traveled Distance (TTD) of all vehicles in the solution. First, the initial solution is created with the time-oriented nearest neighbor heuristic (Figure 5(a)), resulting in NV=3, TTT=617.14, and TTD=266.29 (no real-world units are used). The values of TTT and TTD were decreased by a single application of the LS procedure (Figure 5(b)) to TTT=481.37 and TTD=256.53, with the same number of vehicles. The ILS strategy further decreased the values to TTT=433.88 and TTD=215.17 (Figure 5(c)). The results of the scenario A indicate that the proposed procedure is able to efficiently solve the problem and decrease TTT and TTD values compared to their initial values.

Period k	1	2	3	4	5
Time T_k	$0.05l'_0$	$0.15l'_0$	$0.2l'_0$	$0.4l'_0$	l'_0
Speed v_k	1.0	0.2	0.3	0.2	1.0

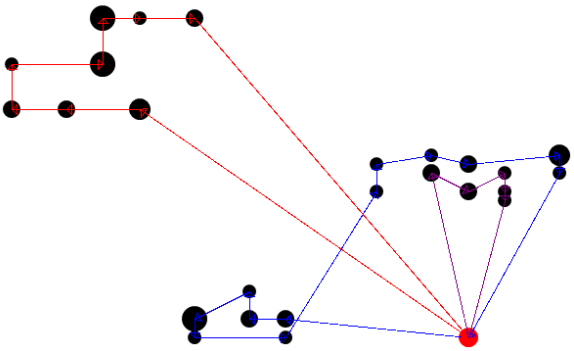
Table 1: Speed values in scenario A



(a) Initial solution - NV:3, TTT: 617.14, TTD: 266.29



(b) LS - NV:3, TTT: 481.37, TTD: 256.53



(c) ILS - NV:3, TTT: 433.88, TTD: 215.17

Figure 5: Example of search stages

4.2 Effect of time-dependent speeds

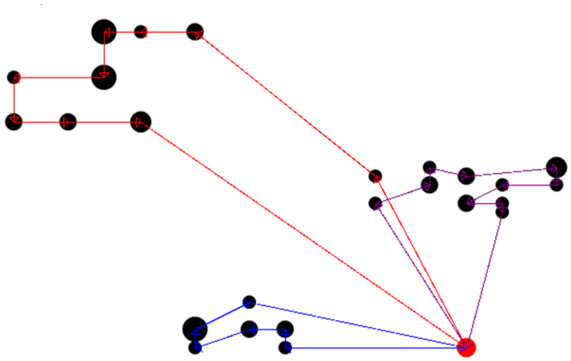
To investigate the effect of time-dependent speeds on the solution quality and customer configuration, another scenario B is observed. The time periods and speeds in scenario B are set in such way that they represent static traffic conditions, so the values of speeds are all equal to 1 in all time periods. In this case, the ILS resulted in NV=3, TTT=205.34, and TTD=205.34. As speed values are equal to 1, the TTT and TTD have the same values. Compared to scenario A, which uses higher speeds, the TTT in scenario B reduced by 52.67%, while the TTD reduced by 4.57%. As it can be seen, lower speeds during the mid-day time period primarily affect the TTT and not the TTD. Therefore, the knowledge of time-dependent travel times has a signif-

Period k	1	2	3	4	5
Time T_k	$0.05l'_0$	$0.15l'_0$	$0.2l'_0$	$0.4l'_0$	l'_0
Speed v_k (C)	0.1	0.2	0.2	0.2	0.1
Speed v_k (D)	0.9	0.1	0.7	0.1	0.9
Speed v_k (E)	0.8	0.6	0.7	0.6	0.8

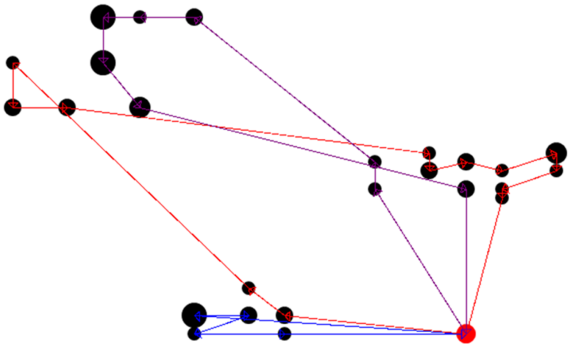
Table 2: Speed values in scenarios C, D and E

icant impact on the delivery duration and can lead to an increase in the transportation cost in terms of driver salary and penalties for late arrivals, especially in time-precise deliveries.

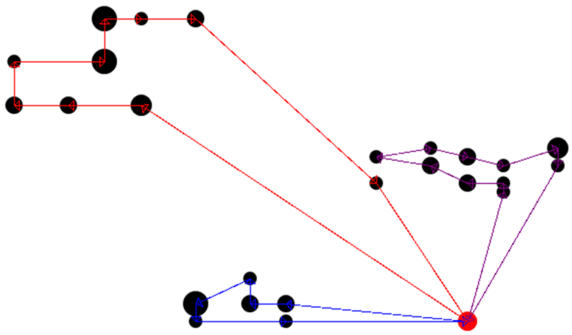
Additional three scenarios are tested: constant congestion (scenario C), rush hours (scenario D) and light congestion (scenario E), with speed values presented in Table 2. The same time intervals are kept as in scenario A. The results are presented in Figure 6, and are the following: (i) scenario C - NV=3, TTT=1363.01, and TTD=188.84 (Figure 6(a)); (ii) scenario D - NV=3, TTT=413.55, and TTD=229.29 (Figure 6(b)); and (iii) scenario E - NV=3, TTT=280.57, and TTD=188.53 (Figure 6(c)). Again the TTD is slightly affected by the different speeds while the TTT is highly affected, especially in scenario C where the travel times increased six times compared to the scenario B, which considers static traffic conditions. As we can see on the figure 6 the configuration of customers per vehicle routes also changes, with the significant change in the scenario D.



(a) Scenario C - NV:3, TTT:1363.01, TTD:188.84



(b) Scenario D - NV:3, TTT:413.55, TTD:229.29



(c) Scenario E - NV:3, TTT:280.57, TTD:188.53

Figure 6: Results for scenarios C, D and E

5 Conclusion

In this paper, a variant of the VRP called capacitated TD-VRP is considered. TD-VRP takes into account time-dependent traffic conditions in the form of time-dependent travel times and vehicles with a limited load capacity. To solve the problem, an ILS metaheuristic is used that combines the ruin-recreate principle to diversify the search and LS procedure to intensify the search. The initial solution, used as input to the metaheuristic, is created using a time-oriented nearest neighbor heuristic. The results on one instance containing 25 customers show that the proposed metaheuristic is able to efficiently solve such problems. The conducted analysis regarding the different speed values in different periods of the delivery time (horizon) showed that time-dependent speeds have little effect on the total distance traveled, while they have a significant effect on the total travel time, especially in heavily-congested scenarios.

For future research, the idea is to investigate the metaheuristic behaviour on instances with several thousand customers, as well as the improvement of the total travel time computation regarding the transitions between two neighboring time periods.

Acknowledgments

The research has been partially supported by Croatian Science Foundation under project IP-2018-01-8323 and project KK.01.1.1.01.0009 (DATACROSS) within the activities of the Centre of Research Excellence for Data Science and Cooperative Systems supported by the Ministry of Science and Education of the Republic of Croatia. Authors are also very grateful to the Faculty of Transport and Traffic Sciences, University of Zagreb for conference publication grant.

References

Tonči Carić and Juraj Fosin. Using congestion zones for solving the time dependent vehicle rout-

ing problem. *PROMET - Traffic&Transportation*, 32(1):25–38, 2020. doi: 10.7307/ptt.v32i1.3296. URL https://traffic.fpz.hr/index.php/PROMTT/article/view/3296.

Tomislav Erdelić and Tonči Carić. A survey on the electric vehicle routing problem: Variants and solution approaches. *Journal of Advanced Transportation*, 2019:1–48, May 2019. doi: 10.1155/2019/5075671. URL https://www.hindawi.com/journals/jat/2019/5075671/.

Tomislav Erdelić, Tonči Carić, Martina Erdelić, Leo Tišljarić, Ana Turković, and Niko Jelušić. Estimating congestion zones and travel time indexes based on the floating car data. *Computers, Environment and Urban Systems*, 87:101604, 2021. ISSN 0198-9715. doi: https://doi.org/10.1016/j.compenvurbsys.2021.101604.

Andres Miguel Figliozzi. The time dependent vehicle routing problem with time windows: Benchmark problems, an efficient solution algorithm, and solution characteristics. *Transportation Research Part E: Logistics and Transportation Review*, 48(3):616–636, 2012. ISSN 13665545. doi: 10.1016/j.tre.2011.11.006.

Florian Kellner. Insights into the effect of traffic congestion on distribution network characteristics a numerical analysis based on navigation service data. *International Journal of Logistics Research and Applications*, 19(5):395–423, 2016. ISSN 1469848X. doi: 10.1080/13675567.2015.1094043.

Gilbert Laporte. The vehicle routing problem: An overview of exact and approximate algorithms. *European Journal of Operational Research*, 59 (3):345 – 358, 1992. ISSN 0377-2217. doi: https://doi.org/10.1016/0377-2217(92)90192-C. URL http://www.sciencedirect.com/science/article/pii/037722179290192C.

Lovro Rožić, Juraj Fosin, and Tonči Carić. Solving the time dependent vehicle routing problem using real-world speed profiles. *Central European Conference on Information and Intelligent Systems, 2015*, 2015.

Marius M. Solomon. Algorithms for the vehicle routing and scheduling problems with time window constraints. *Operations Research*, 35(2):254–265, 1987. doi: 10.1287/opre.35.2.254. URL https://doi.org/10.1287/opre.35.2.254.

Paolo Toth and Daniele Vigo, editors. *The Vehicle Routing Problem*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2001. ISBN 0-89871-498-2.

Thibaut Vidal, Teodor Gabriel Crainic, Michel Gendreau, and Christian Prins. Heuristics for multi-attribute vehicle routing problems: A survey and synthesis. *European Journal of Operational Research*, 231(1):1 – 21, 2013. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2013.02.053. URL http://www.sciencedirect.com/science/article/pii/S0377221713002026.

Virtual Reality and Transport: A Short Review

Filip Vrbanić, Edouard Ivanjko
University of Zagreb, Faculty of transport and traffic sciences
Department of Intelligent Transport Systems
Vukelićeva street 4, 10000 Zagreb, Croatia
{filip.vrbanic, edouard.ivanjko}@fpz.unizg.com

Slaven Gašpar
RIS SYSTEMS
Zavižanska Street 18, 10000 Zagreb,
Croatia
slaven.gaspar@ris-systems.com

Abstract. *The development of virtual reality has been very intensive in recent years. Available systems combined together with very fast wireless communication networks have reached industry-level robustness and are being used more and more in industry, and not only for entertainment. One of the important areas for application is also the transport sector. In the transport sector, virtual reality-based systems improve the schooling of vehicle operators, serve as driver assistant systems, guide workers in warehouses and enable visualization of complex traffic solutions. This paper gives a short overview of the current state and application of virtual reality in the transport sector.*

Keywords. Virtual Reality, Road Traffic, Smart City, Logistics

1 Introduction

Virtual Reality (VR) systems started to evolve more intensively with the broader application of computers for creating various graphic representations. Prior to that era, there existed only simple systems of some sort of virtual reality related to visualization of the surrounding world. Ancient people made drawing in caves to visualize the living conditions, two closely spaced pictures created the effect of stereo vision to present distance events to people in 3D, SENSORAMA attempted to simulate operating different vehicles (Heilig, 1992) without any interactivity, etc.

The development of computing power and fast information sharing enabled the inclusion of VR technology into the industry. The concept of augmented reality, where additional information is presented to workers, has gained special attention (Fite-Georgel, 2011). Tele-training of workers is also one such application (Boulanger, 2004). Even improvement of the very complex system of old industry sectors, bringing them into the current concept of industry 4.0 is in focus. Shipyards give a relevant example (Fraga-Lamas et al., 2018).

The transportation sector is also one that benefits from VR technology. Well-known applications are various simulators or realistic setups for vehicle operator training (car, bus or truck drivers, pilots, and train drivers) (Gąsiorek et al., 2020). Such a concept enables training in different (dangerous) scenarios with the option that a respective scenario can be repeated as many times as needed in the same conditions until the vehicle operator trainee has mastered the correct response. Also, new areas like traffic planning can use the VR technology where traffic simulators can be connected with a virtual environment enabling testing of new solutions prior to building them in the real world (Hasan et al., 2021). Advanced driver assistant systems are also one of the areas of VR application, making our roads safer and driving more easier. Logistics also benefit from this technology where warehouse workers can be guided more efficiently to the respective item location or receive help how to apply new mobile robot based systems (Kalinov et al., 2021).

With the inclusion of VR into the transport sector in different areas combining new simulation approaches, there is a need for an analysis of the possibilities of these technologies. This is the aim of this paper to create a short review of the most important VR applications in the transport sector, ranging from better simulation approaches, concepts for the intelligent city, and new better logistics.

This paper is structured as follows. After the Introduction section, a brief description of the development of virtual reality is given in the second section. The following third section elaborates on the use of virtual reality for simulations. A description of intelligent city applications is given in the fourth section. Application in the logistics sector is continued in the fifth section. The paper ends with the sixth section containing the conclusion and perspective on transport sector applications.

2 Development of Virtual Reality

As mentioned, the very beginning of VR dates to ancient times when people made visualizations of their environment in caves creating the first visualization effects. This resulted in realistic murals and stereoscopic images presenting important events at the end of the 19 century. The development of VR systems as we know them today was driven by the improvement of computers and the increase in their computing power. Ad-

ditional miniaturization led to the first head-mounted displays increasing the interactivity of the user with the simulated virtual world. Thus, mass production of VR systems began at the end of the 20 century with many applications in entertainment, training, and simulations for military and space exploring agencies.

At the beginning of the 21 current century, several companies brought to the market VR systems easily connectable to a standard PC. Some systems were produced to be compatible with smartphones. Expect of tracking only the eyes and view direction of the user, today's systems can also track the movement of the hands; with 6 degree of freedom wireless tracking sensors, the whole body can be tracked, and the user gestures can be deciphered. New lightweight glasses for augmented reality give the opportunity to present a lot of different additional information to the user. Thus, the application of VR in every sector dealing with complex systems is possible by using state-of-the-art VR systems.

3 Virtual Reality Based Simulators

Virtual reality test systems have been present in the automotive industry for a long time and have been widely used. Their typical use is testing during Research and Development (R&D) processes, as well as testing in certification processes, such as emission certification. Application areas of virtual reality in the automotive industry can be divided into two main groups; (i) purely simulated test systems and (ii) Hardware-In-the-Loop (HIL) test systems.

Purely simulated test systems are used in the early phase of the R&D process. With this type of test system, everything is simulated and modeled in virtual reality - vehicle chassis, engine, road, traffic, pedestrians, etc. Due to advances in simulation technology and the development of cars with automated and autonomous driving, this type of simulation is becoming more important, and its usage is increasing significantly.

HIL testing is a type of testing where Unit Under Test (UUT), i.e., engine, small vehicle part,or the entire vehicle, is not simulated but is real, while the rest of the components are simulated in virtual reality - vehicle chassis, gearbox, road, and wind resistance, etc. This type of testing is used when more precise testing and measurement are needed, and UUT simulation cannot be applied because current technologies do not support the required level of reality approximation. HIL testing is typically used in the final stage of the R&D process for calibration, optimization, and certification of the vehicle itself. The two most common cases of HIL testing are engine testing on an engine test bed and vehicle testing on a chassis dyno test bed. Engine testing on an engine test bed is real, and the rest, wheels, vehicle chassis, road, and wind resistances,are placed into virtual reality. In the case of vehicle testing on the chassis dyno test bed, the entire vehicle is real, and the rest, road and wind resistances, traffic, pedestrians, and infrastructure are placed into virtual reality. Emission certification is an example in which testing on the HIL system is used in the certification process. This type of testing and certification is being used globally, and it is conducted on a chassis dyno test bed.

On the other hand, systems for automated and autonomous driving are very complex, consisting of multiple hardware and software systems and subsystems with a lot of sensors. Purely simulated test systems play a huge and important role in R&D process because the number of tests, test cases, and testing hours are thousands or tens of thousands. With the current technology development, validation and certification of these systems are only possible by driving on the roads or in "Fake cities." This approach to testing has numerous limitations. Some of them are:

- Very dependent on external conditions, like day, night, rain,
- Limited by geographic position and season (e.g., testing in rain or snow is limited to certain periods of the year, if possible at all),
- Limited number of possible testing scenarios and test cases,
- Difficult measuring and data acquisition,
- Difficult and slow software debugging process,
- Potentially very dangerous for all persons involved in the testing process (especially the test driver or surrounding test pedestrians),
- Test automation is limited or not possible at all.

The consequence of all these limitations is that validation and certification processes are very time-consuming and costly. An additional and very important limitation is that the test scenarios and test cases are neither repeatable nor reproducible. By testing in general and especially by certification testing, repeatability and reproducibility are of key importance. All of the previously mentioned obstacles and limitations could be eliminated by testing and validating in virtual or augmented reality. Current simulation technology is still not developed enough to be applied in such tests, and currently, automotive companies are putting a big effort into the development of virtual reality test systems, which provide more realistic validation and/or certification testing possibilities.

On the other hand, VR simulators can also be utilized in a Microscopic Traffic Simulation (MTS) framework, including human control using VR technologies (Hasan et al., 2021). One such approach developed in (Hasan et al., 2021) allows up to 10 geographically distributed users to run and guide vehicles in MTS online using the widely available retail VR devices. The main goal of the framework is set

to the improvement of precision of MTS by developing as close to real VR immersive scenes and environments. The use of low-cost, affordable VR devices in the proposed distributed MTS expands the usability of MTS for many users, thus, significantly improving MTS results by incorporating realistic human behavior. A different approach to VR simulation based on a driving simulator game-based learning system that assists divers in practicing driving behavior and real-world traffic situations was developed in (Muhammad Amirul Fawwaz Nik Abdullah et al., 2022). One key aspect of the research was to analyze human behavior when interacting with driving simulator game-based learning. Therefore, the authors developed five assessment criteria, including simulation satisfaction, interaction with the simulator, motivation to complete challenges, challenges to enhance driving skills, and effectiveness of the simulation framework.

VR systems play an important role in the automotive industry, and with the development of vehicles with automated and autonomous driving, their importance is growing exponentially. At the moment, VR technology is behind the market needs, and therefore, companies are investing a lot in research and development. Testing and validation in or with VR is the future of testing, and it is a field with great potential for scientific, engineering, and business activities.

4 Intelligent City Applications

Intelligent cities represent the new strategy used to improve current cities' problems that include traffic, safety, pollution, resource management, and waste processing/recycling. The design of a smart city includes more efficient transport and intelligent mobility, a feasible environment, and controlled urbanization. Therefore, new technologies, in particular, VR, can be utilized in the development of various aspects that contribute to building a smart, sustainable city. The active user immersion and participation in the constructed three-dimensional virtual environment help in designing an intelligent city. The construction of the virtual scene according to the perception of the real environment in an immersive and interactive way gives insights into current city problems. Most importantly, it offers virtual solutions before implementation in the real world. A vast majority of recent VR frameworks lie upon gathered relevant and real-time data from sensors.

The implementation aspect of VR in developing a road traffic virtual environment that replicates the real-world in various conditions (Bourhim & Cherkaoui, 2019). This environment can be simulated by altering various traffic parameters that include: signal length, traffic flow, the type of vehicle, time and speed, and many more. According to (Bourhim & Cherkaoui, 2019), this framework can help in the development of new solutions using VR technology for validation, evaluation, and simulation of multiple materials, environmental, financial, and traffic-related scenarios. The development of such a VR environment is beneficial for road planning and design in a 3D virtual city for testing and comparing different design schemes that involve modifying the physical properties of traffic infrastructure (Yan et al., 2021).

The application of VR technology can also be beneficial to train drivers, as discussed in (Zhang, 2021). The idea of VR is to convey the input driving signals, such as throttle opening and steering position, into the VR framework system. Therefore, drivers can also be trained to adapt to new traffic regulations and imposed intelligent traffic control systems. Furthermore, the prototype of a VR educational platform was developed to address the design and implementation of VR systems for learning through serious games and the creation of a mobility web observatory for disseminating information through digital activism and data visualization and analysis (León-Paredes et al., 2020). The main idea of the VR prototype is to offer a new interactive road learning platform. The main goal is to train citizens and collect data to detect and analyze the participants' routines and mishaps when driving, transit, walking, or transporting.

Another aspect of VR application lies in the control and monitoring of traffic flows for users. The development of such a system allows users to analyze traffic flows in multiple scenarios and roads to support better decision-making to control and monitor traffic flow in a smart city (Yao et al., 2018).

In (Roldán-Gómez et al., 2022), VR based system was proposed, developed, and validated as a smart monitoring system for the smart city based on an aerial robotic swarm in a simulation environment. Drones were used to develop a behavioral algorithm to monitor traffic, pollution, and ecological aspects in a smart city. The immersive VR interface can perceive the state of the city, determining and detecting the most relevant events in a city.

One of the newest concepts to foster the intelligent city concept is the digital twin approach. In this case, a whole city or only a part of it can be modeled in a virtual environment, and the digital twin is constantly improved during operation. Thus, city operators have the whole time accurate insight into what is happening with improved visualization. The advantage to common VR-based city models is that real-time traffic measurements can be included in the execution of the VR model, ensuring a constant correction of the simulated traffic flows and mobility entities' behavior. The latest example of such a constantly improving digital twin is the Geneva motorway example described in (Kušić et al., 2022).

5 Logistics Applications

The development and introduction of VR into the transport sector also branch to the logistics sector. The need for improved logistics distribution is especially evident due to the fast-growing e-commerce in recent years. Therefore, the standard logistics distribution path display system that is built on a geographic information system only displays the information about the transfer center and goods on a two-dimensional map. The VR-based e-commerce logistics distribution path display system has a better user experience and reliability with the interactive and intuitive features was developed in (Min, 2020). The VR distribution center scene was created in Unity3D software by combining the two-dimensional road satellite map and the positioning information on the satellite map to establish a plane gridded scene.

VR can also be used to set up a virtual logistics laboratory, as shown in (Li et al., 2009). The proposed VR structure of three-dimensional storage laboratory mainly consists of four modules, including automatic storage and retrieval system, automatic conveyor system, automatic picking system, and warehouse management system (Li et al., 2009). The developed VR model can be used to record, search, sort, modify, and print. Furthermore, checking-in management module incorporates management of goods, weighing, checking and acceptance of goods, stock pre-allocation, and more. Two more modules include warehouse management and customer relationship management.

VR was successfully implemented in the port container terminal that serves as an intermodal global transportation network (Shu et al., 2007). The port logistics information platform incorporates port management sub-systems and an information-sharing platform to help visualize data. Data mining and visualization are combined and integrated with communication among port entities. VR-based container operation management system was developed based on an existing real-time system that drives objects in motion in a 3D scene. Furthermore, the developed system can simulate the movement of objects in a 3D scene operated by a human through an interface. In (Chen et al., 2010), authors analyzed the existing problems in a port container terminal simulation and developed a methodology to construct the container terminal VR system faster. The advantages of using component technology to deploy a container terminal VR system, according to the authors, include an efficient, reusable, and flexible method to develop component VR container terminal simulation. In a recent paper (Liu et al., 2021), authors developed a platform for port safety training. The platform includes an interactive training learning function in a VR scene to increase the interest of safety training and solve the problem of insufficient learning motivation. Furthermore, the interactive assessment is developed in the VR scene to evaluate the quality of safety training results. Lastly, the port VR provides an immersive sense of substitution to help develop required safety skills quickly and effectively for a port management system.

6 Conclusion

The development of new VR technologies opens new possibilities in various industry areas and not just entertainment. One of the industry sectors, which is getting more and more benefits from VR, is the transport sector. In addition to the well-known vehicle operators learning application of VR, new approaches related to better visualized and more accurate simulations and new guiding systems for workers in highly automatized warehouses are being put into use. The latter is especially beneficial in optimal item placement warehouses where only the warehouse management system knows where each respective item is.

Also, a synergy of other new technologies is happening with VR in the transport sector. The best example is the digital twin technology with which an accurate modeling/calibration backbone to the virtual world can be made. Thus, more realistic simulation environments for training vehicle operators, traffic management centers, and urban planning departments can be made. It can be expected that VR will become a new important tool to alleviate solving current problems in the transport sector, making the current transport network more green, intelligent, integrated, and resilient to various disturbances.

Acknowledgments

This research has been supported by the University of Zagreb Faculty of Transport and Traffic Sciences, and the University of Zagreb Support Program for scientific and artistic research short-term support: "Advanced Methods for Traffic Control."

References

Boulanger, P. (2004). Application of augmented reality to industrial tele-training. In *First canadian conference on computer and robot vision, 2004. proceedings.* (p. 320-328). doi: 10.1109/CCCRV.2004.1301462

Bourhim, E. m., & Cherkaoui, A. (2019). How can the virtual reality help in implementation of the smart city? In *2019 10th international conference on computing, communication and networking technologies (icccnt)* (p. 1-6). doi: 10.1109/ICCCNT45670.2019.8944508

Chen, Y., Chang, D., Mi, W., & Lu, H. (2010). Study on container terminal virtual reality system based

on component technology. In (p. 299â302). IEEE Press.

Fite-Georgel, P. (2011). Is there a reality in industrial augmented reality? In *2011 10th ieee international symposium on mixed and augmented reality* (p. 201-210). doi: 10.1109/ISMAR.2011.6092387

Fraga-Lamas, P., Fernández-Caramés, T. M., Blanco-Novoa, O., & Vilar-Montesinos, M. A. (2018). A review on industrial augmented reality systems for the industry 4.0 shipyard. *IEEE Access*, *6*, 13358-13375. doi: 10.1109/ACCESS.2018.2808326

Gąsiorek, K., Odachowska, E., Matysiak, A., & Pędzierska, M. (2020). Virtual reality technologies in the training of professional drivers. comparison of the 2d and 3d simulation application. In M. Siergiejczyk & K. Krzykowska (Eds.), *Research methods and solutions to current transport problems* (pp. 133–142). Cham: Springer International Publishing. doi: 10.1007/978-3-030-27687-4_14

Hasan, M., Perez, D., Shen, Y., & Yang, H. (2021). Distributed microscopic traffic simulation with human-in-the-loop enabled by virtual reality technologies. *Advances in Engineering Software*, *154*, 102985. doi: 10.1016/j.advengsoft.2021.102985

Heilig, M. L. (1992). EL Cine del Futuro: The Cinema of the Future. *Presence: Teleoperators and Virtual Environments*, *1*(3), 279-294. doi: 10.1162/pres.1992.1.3.279

Kalinov, I., Trinitatova, D., & Tsetserukou, D. (2021). Warevr: Virtual reality interface for supervision of autonomous robotic system aimed at warehouse stocktaking. In *2021 ieee international conference on systems, man, and cybernetics (smc)* (p. 2139-2145). doi: 10.1109/SMC52423.2021.9659133

Kušić, K., Schumann, R., & Ivanjko, E. (2022). Building a motorway digital twin in sumo: Real-time simulation of continuous data. In *Accepted for presentation on 64th international symposium elmar2022.*

León-Paredes, G. A., Bravo-Quezada, O. G., Sacoto-Cabrera, E. J., Pizarro-Gordillo, O. F., Vintimilla-Tapia, P. E., Bravo-Torres, J. F., & Cabrera-Chica, W. P. (2020). Virtual reality and data analysis based platform for urban mobility awareness as a tool for road education. In *2020 ieee andescon* (p. 1-6). doi: 10.1109/ANDESCON50619.2020.9272084

Li, R., Zhu, Z., Liu, H., Gao, L., & Wu, Y. (2009). Research on three-dimensional interactive storage laboratory system under virtual reality environment. In *2009 international conference on computational intelligence and software engineering* (p. 1-3). doi: 10.1109/CISE.2009.5363664

Liu, W., Cheng, L., Liu, Z., Yang, Y., & Li, L. (2021). The development of port safety training platform based on virtual reality technology. In *2021 ieee 7th international conference on virtual reality (icvr)* (p. 207-214). doi: 10.1109/ICVR51878.2021.9483808

Min, L. (2020). Design of e-commerce logistics distribution path display system based on virtual reality technology. In *2020 international conference on virtual reality and intelligent systems (icvris)* (p. 38-41). doi: 10.1109/ICVRIS51417.2020.00017

Muhammad Amirul Fawwaz Nik Abdullah, N., Huzaifah Ahmad Sharipuddin, A., Mustapha, S., & Mohammed, M. N. (2022). The development of driving simulator game-based learning in virtual reality. In *2022 ieee 18th international colloquium on signal processing applications (cspa)* (p. 325-328). doi: 10.1109/CSPA55076.2022.9781976

Roldán-Gómez, J. J., Garcia-Aunon, P., Mazariegos, P., & Barrientos, A. (2022, 01). Swarmcity project: monitoring traffic, pedestrians, climate, and pollution with an aerial robotic swarm. *Personal and Ubiquitous Computing*, *26*(4), 1151-1167. doi: 10.1007/s00779-020-01379-2

Shu, F., Mi, W., & Xu, Z. (2007). The information sharing platform for port container terminal logistics using virtual reality. In *2007 ieee international conference on automation and logistics* (p. 2570-2575). doi: 10.1109/ICAL.2007.4339013

Yan, G., Chen, Y., & Gupta, D. (2021). The application of virtual reality technology on intelligent traffic construction and decision support in smart cities. *Wirel. Commun. Mob. Comput.*, *2021*. doi: 10.1155/2021/3833562

Yao, J., Kong, X., Sandrasegaran, K., & Liu, L. (2018). Develop a virtual system using traffic flow model for intelligent transportation in smart city. In *2018 18th international symposium on communications and information technologies (iscit)* (p. 110-115). doi: 10.1109/ISCIT.2018.8587953

Zhang, Z. (2021). Research on the application of virtual reality technology in smart city construction. In *2021 2nd international conference on artificial intelligence and information systems.* New York, NY, USA: Association for Computing Machinery. doi: 10.1145/3469213.3470311

Trends in Software Engineering

Matija Novak, Iva Levak
Using graph databases in source code plagiarism detection

Luis de-Marcos, José-María Gutiérrez-Martínez, Adrián Domínguez-Díaz, Sergio Caro-Álvaro, Daniel Rodríguez
Evaluating the UX of a Microservice Web App: Cyberranges in the µDevOps Project

Nadica Hrgarek Lechner, Vjeran Strahonja, Zlatko Stapić
Threat Modeling Methods in the Medical Device Industry: An Integrative Literature Review

Jelica Stanojević, Uroš Šošević, Miroslav Minović, Miloš Milovanović
An Overview of Modern Cross-platform Mobile Development Frameworks

Žana Zekić, Zlatko Stapić
Software Quality Assessment Standards and Metrics: A Systematic Literature Review

Robert Logozar, Matija Mikac, Danijel Radošević
Exploring the Access to the Static Array Elements via Indices and via Pointers — the Introductory C++ Case

Using graph databases in source code plagiarism detection

Matija Novak, Iva Levak
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin
matija.novak, iva.levak@foi.unizg.hr

Abstract. *Modern plagiarism detection tools calculate the percentage of the similarity between two given source code files. In academia, the process of checking for potential plagiarized students solutions can be challenging in terms of resources due to the large number of combinations between many students. In such conditions, the reliability of plagiarism detection tools may be put to risk. Every plagiarism detection tool produces a similarity report as files containing the results of the analysis for each pair of analyzed source code files. While such a report is useful for a one-time checking, sometimes it is needed to store the result data for future use. In our previous work, the results were stored in a relational database and a list of relevant queries was defined for meaningful analysis. Nevertheless, the large number of pair-wise impacts the storage and query execution speeds. In this paper, we present a solution to this problem by importing the similarity analysis data into a graph database and evaluate the difference in the query execution speed between a graph and a relational database.*
Keywords. plagiarism, graph databases, similarity detection

1 Introduction

In academia, students are required to do homework assignments as part of their course evaluation. One problem that teachers have with homework assignments is plagiarism. Plagiarism is "the act of taking the writings of another person and passing them off as one's own" according to (Encyclopædia Britannica Inc, 2015). There is a similar problem called collusion which means "working together to produce assessed work in circumstances where this is forbidden" (Barrett and Cox, 2005). The main difference is that in collusion the original author knows and participates in the process of cheating. While there are many definitions what constitutes plagiarism in this paper the definition from (Novak, 2020) will be used which is defining specifically source-code plagiarism and it seems the most accurate in this case: "Plagiarism, in programming assignments, is the act of taking a significant amount of source-code parts (up to the entire source-code) from other students or from the Internet and using it without noting its origin. A 'significant amount' means that the similarity between two solutions of a programming assignment is high enough that an expert (teacher, ethical board, etc.) considers specific student work as sufficiently 'real' plagiarism to accuse the student of plagiarism.".

The main problem when dealing with plagiarism is how to find out that plagiarism is taking place. In small classrooms this is not a big issue just by grading the assignments the teacher notices similarity and can act accordingly. But in big classrooms (100 or more students) it is impossible to remember every solution and compare them. In such situations, similarity detection tools are a significant assistance. Such tools are often called plagiarism detection tools yet it's not really what they perform rather for what purpose they are used. So-called plagiarism detection tools calculate the percentage of the similarity between two given files. Some tools are for text and some are specialized for source-code. A good comparison can be found in (Ragkhitwetsagul et al., 2016; Novak et al., 2019). As already mentioned, in this paper the focus is on source-code similarity since the chosen courses are programming courses.

Plagiarism detection tool produces a similarity reports as files and while such a report is useful for a one-time checking, sometimes it is needed to store the result data for future use. In our previous work, the results were stored in a relational database and list of relevant queries was defined for meaningful analysis. Nevertheless, the large number of pair-wise impacts the storage and query execution speeds. In this paper, we present a solution to this problem by importing the similarity analysis data into a graph database and evaluate the difference in the query execution speed between a graph and a relational database.

Section 2 gives the related work. In section 3 the plagiarism detection process is presented and more details about concrete problem are given. Section 4 describes the methodology and section 5 describes the transition to graph database. Section 6 presents and discusses the query execution results in both databases and section 7 concludes.

2 Related work

In works like (Joy and Luck, 1999; Martins et al., 2015; Đurić and Gašević, 2013) various similarity detections are performed with the purpose of detecting plagiarism in academia source-code assignments. The first works date back to 80's like (Donaldson et al., 1981). But, when looking for plagiarism detection engines or similarity detection engines that use graph database, there is not much work done. There is a paper by (Arora et al., 2021) where they use graph databases to visualize the code structure so it can be analysed. Likewise, in (Tyagi et al., 2022) they are "converting a Java program into a specialized dependency graph". On the other hand there are a lot of papers like (Esser and Fahland, 2021; Timón-Reina et al., 2021; Šestak et al., 2021) and books (e.g. (Robinson et al., 2015)) about graph databases and their usage in various areas (e.g. Facebook's Open Graph, Google's Knowledge Graph, Twitter's FlockD) as stated by (Miller, 2013). In this paper we want to apply graph databases in the context of plagiarism detection and see if we can get good results. In difference to the previously mentioned work in this paper we use the typical similarity detection tools (e.g. Sherlock, jPlag) for detection and graph database in the analysis's phase. More details about the tools are given in the next section.

3 Plagiarism detection process

In (Đurić and Gašević, 2013; Kermek and Novak, 2016) the process of plagiarism detection is described, which can be divided into four main phases: preparation phase, preprocessing phase, detection phase and result analysis phase. The preparation phase is usually not complicated, especially if students submit assignments in Learning Management Systems like Moodle, it includes downloading the code and passing it to the preprocessing phase. Preprocessing phase deals with cleaning the data before detection and an in-depth analysis of the effectiveness of preprocessing was done in (Kleiman and Kowaltowski, 2009; Đurić and Gašević, 2013; Novak, 2020). Afterwards, the similarity is calculated using tools like SIM (Grune and Huntjens, 1989), Sherlock (Joy and Luck, 1999), JPlag (Prechelt et al., 2002) etc. These tools produce output in various forms, but usually the data is displayed in a table form (Prechelt et al., 2002; Grune and Huntjens, 1989) or graph form (Joy and Luck, 1999; Novak et al., 2021). This process normally takes a while depending on the number of submissions and the code lengths of each submission. At last, the data is analysed by the teacher by looking at the highest similarities mainly in side-by-side comparison and deciding if plagiarism is taking place.

While our process is the same in the first three phases the analysis phase differs. In our case, students can submit their assignments at different points of time usually before an exam. According to the faculty regulations, a student can retake the exam (if he does not pass) up to eight times in four years. Therefore students from different academic years submit their projects before each exam. In order to make sure that students did not use the projects or parts of the projects from students that already passed, the course plagiarism detection needs to be performed before each exam. The comparisons are done between student projects from the same academic year since they have the same assignment. The project assignment is to build up a full-featured web application that in the end has around 1500-5000 lines of code (LOC). In order to not repeat detections that were already done, the data is stored in a relational database. Then before each exam, a query is performed to get similarities larger than 30% for students that are taking the exam. If high similarities occur, then projects similarities are analysed manually and if plagiarism is taking place the student can not take the exam and is reported to the ethical board.

Conducting plagiarism analysis happens right before every exam for each applied student. Number of examined students can vary from one student to roughly one hundred. The issue here is that over the years the database became quite large (over 1.200.000 data) and the query executions became slow. By looking at the MySQL process status it was noticed that sometimes a query would take up to 1 minute to execute. This can be problematic if 50 students apply for an exam so it takes up too much time to execute all queries for all applied students. Naturally, the execution not only depends on the number of students that applied but also on the server load, the number of plagiarised matches, etc. It is not possible to influence the server load, also the number of plagiarised matches is not something that can be changed and many other factors are out of our reach. But maybe the query executions can be improved by cleaning up the database of old data, which didn't solve the problem. Afterwards, we tried to optimize the queries so that only one complex query is submitted instead of multiple queries. While it seemed better at first, by monitoring the query times in MySQL process status we noticed that some queries still took around 40 seconds. The query that was used looked like this:

```
SELECT nameP,similarity,fileType
FROM (SELECT
  CONCAT(‘username1‘,’-’,‘username2‘)
    as nameP,
    ‘username1‘ , ‘username2‘ ,
    ‘fileType‘ ,
    SUM( ‘similarity‘ ) as ‘similarity‘
  FROM ‘matches‘
  WHERE
  (username1=’studentA’
    OR username2=’studentA’)
  GROUP BY ‘username1‘ , ‘username2‘,
  ‘fileType‘
```

```
order by similarity DESC) as temp
WHERE ‘similarity‘ > 20
```

Mentioned query returns information about which student has significantly high similarity to observed student. Nested query produces results which includes observed student and his observed pairs and their total of similarities by file type.

It might be that the query can be further optimised but since there is a work about graph databases that shows their advantage over the relational database like (Cheng et al., 2019) the idea is to migrate the data to a graph database and see if the query execution will be significantly faster.

4 Methodology

To verify if graph database is a good choice in this situation we performed an experiment. To have comparable results the experiment was done for both databases on the same machine with 4GB of RAM and 4CPU cores. We are aware that it's recommended to have at least 8GB of RAM for Neo4j but we wanted to perform experiment in environment which can be used by an individual teacher. We are aware that needs for used tools are not met, but if we manage to improve the times on such hardware then it is the expectation that the times will be even better on a server.

In the experiment we will measure time to execute the query for 1 student, 33 students and two times the same 33 students which makes a total of 66 students. Also, further improvement of the MySQL query was performed and at last the query used for testing in MySQL looks like this:

```
SELECT CONCAT(‘username1‘,’-’,‘username2‘)
    AS imeP,
    SUM(similarity) as similarity,fileType
FROM matches
WHERE (username1=’studentA’
   OR username2=’studentA’)
GROUP BY ‘username1‘ ,
   ‘username2‘, ‘fileType‘
HAVING SUM(similarity) > 20
ORDER BY similarity DESC
```

First of all, query is filtering all data by examined student username. Query then produces totals of similarities between examined student and its pair student. Finally, query filters only the results in which similarity is greater than some number, in this case twenty.

Before conducting the experiment, all data needs to be transitioned to the graph database.

5 Transition to graph database

In order to transition to a graph database the data was exported from a relational database and stored

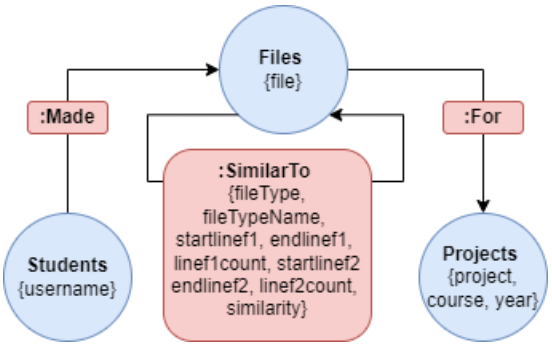


Figure 1: Graph Database schema

first as a CSV file. CSV file was then used for importing data into graph database. Neo4j [1] was used for graph database and installing it was very simple but the main problem was importing all the data into the database. The data for import was stored at $/var/lib/neo4j/import$ directory so it can be loaded with:

```
LOAD CSV WITH HEADERS FROM
"file:///matches.csv" as line
CREATE(:Plagijat{file1: line.file1,
file2:line.file2,
similarity:toInteger(line.similarity),
username1:line.username1,
username2:line.username2,
project:line.projekt,
course: line.predmet,
year:toInteger( line.godina)});
```

This was the first attempt which did not actually create graph rather it created independent nodes where one node was corresponding to one row in the relational database. Keywords WITH HEADERS allowed us to load CSV files with header row. To create graph in database we used next database schema as presented in Figure 1.

In database schema there are three types of nodes: Students, Files and Project which are connected with relationships :Made and :For. Main relationship :SimilarTo is recursive and defines file pairs similarities. In relation with created database schema, loading of data was with next query.

```
LOAD CSV WITH HEADERS
FROM "file:///load1.csv" AS line
WITH line
  MERGE (f1:Files {file:line.file1})
  MERGE (f2:Files {file:line.file2})
  MERGE (p:Projects{project:line.projekt,
    year:toInteger(line.godina),
    course:line.predmet})
WITH f1, f2, p, line
  MERGE (f1)-[:For]->(p)<-[:For]-(f2)
WITH f1, f2, line
```

[1] https://neo4j.com/

```
MERGE (f1)-[:SimilarTo
  {fileType:line.fileType,
  fileTypeName:line.fileTypeName,
  startlinef1:line.startlinef1,
  endlinef1:line.endlinef1,
  linef1count:line.linef1count,
  startlinef2:line.startlinef2,
  endlinef2:line.endlinef2,
  linef2count:line.endlinef2,
  similarity:line.similarity}]->(f2)
```

Rather then CREATE, clause MERGE was used as it intends to create nodes and relationships that don't already exist in database. Data in CSV file were divided in sequences of 100 000 lines per file because of heap memory. With given query, two of three mentioned nodes were created, Files and Projects, and two of three relationships, :For and :SimilarTo. Nodes and relationships are created with properties corresponding to schema in Figure 1. Relationship :For defines which file belongs to which project, and relationship :SimilarTo defines plagiarism detection results between two files. In consequence of big data, student usernames were loaded separately. First, usernames with corresponding file names were excluded from CSV files and loaded with next query.

```
LOAD CSV WITH HEADERS
FROM "file:///studentFiles.csv" AS line
WITH line
  MERGE (f:Files{file:line.file})
  MERGE (s:Students
    {username:line.username})
WITH f,s
  MERGE (s)-[:Made]->(f)
```

Query above was used for importing student usernames as node named Students and connecting usernames and matched files with relationships :Made. To do the testing we used Node since it has packages for both MySQL and Neo4j in the repositories:

```
sudo npm -g install mysql2
sudo npm -g install neo4j-driver
```

6 Results and discussion

For tests, we have performed the same queries on a relational database and on a graph database. Not exactly the same, since they used different query types but with the same intent, conditions and data comparisons. The idea was to get all similarities bigger than some percentage for a particular student. While executing queries, the time was measured for how long it takes to get a result response.

In Table 1 the average execution time is presented. The test was performed with only 1 student, then with 33 different students and then with 2 times 33 students which means that every student was queried twice. The total amount of data was 1.200.000 in both databases. The first attempt with was importing data into GraphDB without much considering the graph and as one can see the GraphDB query is slower than the MySQL optimized query as there is no graph to work with. The query in the graph database looked like this:

```
MATCH(p:Plagijat) WHERE
  (p.username1="studentA"
    OR p.username2="studentA")
  WITH {type:p.type,
       user1:p.username1,
       user2:p.username2,
       suma:SUM(p.similarity)}
  AS u WHERE u.suma>20
  RETURN u
```

As database didn't have relationships, this query is very similar to MySQL query, with disctincion of not having to explicitly specify grouping of data. Grouping is in fact formed from the non-aggregation columns in scope when aggregation function like SUM is used. Query searches for all data with specified student username and returns data with sum of similarities of students per file type.

As mentioned, second experiment was to create a real graph database with relationships. In this case the benefit is in having a graph representation to analyse one student relations to others. This time the query looked like this:

```
MATCH(s:Students{username:"studentA"})
  --()-[st:SimilarTo]->()--(s1:Students)
WITH {type:st.type,
  user1:s.username,
  user2:s1.username,
  suma:SUM(toInteger(st.similarity))}
AS u WHERE u.suma>20 RETURN u
```

Previous query searches for all relationships :SimilarTo from student with username "studentA" to all students. Although not explicitly written, in query there are used relationship :Made and nodes Files. Full MATCH clause would look like:

```
MATCH(s:Students{username:"studentA"})
  -[:Made]->(:Files)-[st:SimilarTo]
  ->(:Files)-[:Made]->(s1:Students)
```

Query firstly selects node :Students with username "studentA" then finds all the files that student made, and secondly finds files of another students which files where compared to files of "studentA". Lastly it returns calculated similarity between two students.

One can see from Table 1 that with added relationships, execution time is improved. For better comparison of MySQL optimized query and Neo4j query with relationships, query execution times for one student are presented as graph in Figure 2. In Figure 3 the results are presented for executing a queries for 33 students,

Table 1: Results summary

Tool	Students	AVG Total Time
MySQL	1	0.496s
Neo4j with one node	1	1.628s
Neo4j with relationships	1	1.092 s
MySQL	33	18.525s
Neo4j with one node	33	19.172s
Neo4j with relationships	33	2.468s
MySQL	2*33	38.101s
Neo4j with one node	2*33	41.718s
Neo4j with relationships	2*33	2.679s

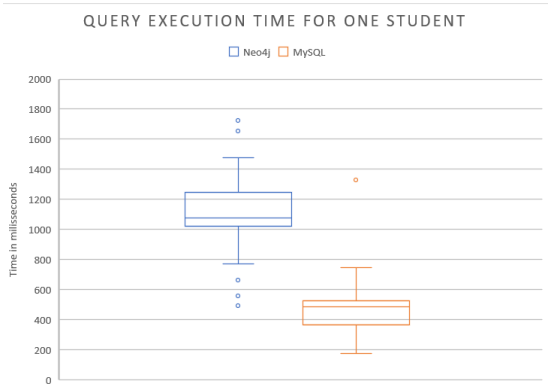


Figure 2: Query execution time for 1 student

and in figure 4 the results are presented for executing a queries for 66 students. From this analysis we conclude that MySQL performed better when executing only one query but when the number of queries is larger then Neo4j starts to outperform MySQL.

One of the big advantages of using graph database is that tools like Gephi are not needed, as used in (Novak et al., 2021) where the authors display the benefits of graph representation of student file similarities. With graph database one can directly have a graph representation without any need for other tools. In Figure 5 we can see connections of selected student with other students.

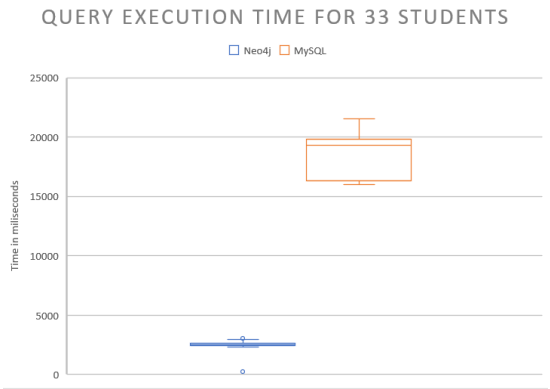


Figure 3: Query execution time for 33 students

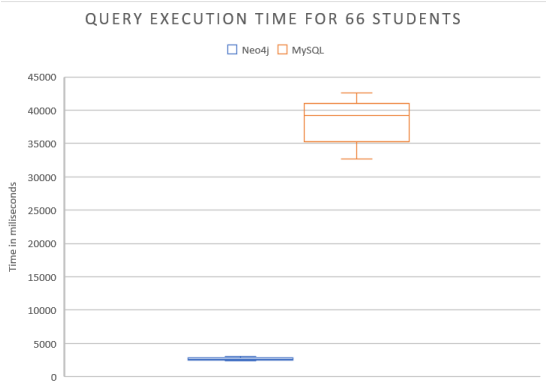


Figure 4: Query execution time for 66 students

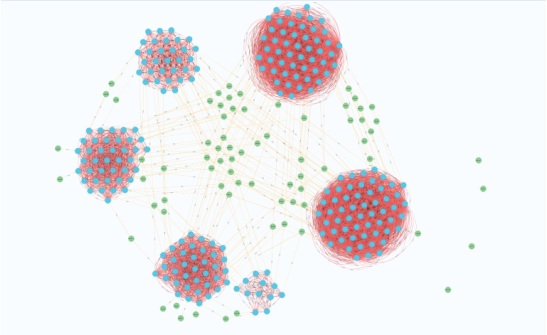


Figure 5: Node connections for one username in Neo4j

7 Conclusion

In this paper, we have shown that graph databases provide a lot of benefits for analysing data in the context of plagiarism detection. Firstly, it represents data connections which can be used for analysis so there is no need for installation and configuration of additional tools for creating graphs. Furthermore, our testing results show that with the correct graph database scheme, executing presented queries and getting the result set of data from a graph database outperforms the relational database when the number of queries increases. With only one query MySQL had an average of 0.495 seconds execution time while Neo4j had 1.098 seconds execution time. On the other hand with 33 students the average execution time of MySQL queries increased to 18.525 seconds while average execution time of Neo4j queries was only 2.468 seconds.

In future works, the idea is to expand research by additional testing on created graph database to determine the real impact of graph databases in the process of plagiarism detection. Moreover, additional tests have to be performed on better hardware with different data in order to have a better understanding and to be able to generalize the findings. Current setup was used intentionally since this is a hardware configuration that most teacher should have available. Based on current results, although experiment was performed on

lower-end hardware and because Neo4j outperformed MySQL, we expect that with better hardware the differences will even more in favor of Neo4j since it demands at least 8GB of RAM.

References

Arora, R., Motilal Maurya, A., and Sharma, Y. (2021). Application of java relationship graphs (jrg) to plagiarism detection in java projects: A neo4j graph database approach. In *2021 The 4th International Conference on Software Engineering and Information Management*, pages 46–51.

Barrett, R. and Cox, A. L. (2005). At least they're learning something: the hazy line between collaboration and collusion. *Assessment & Evaluation in Higher Education*, 30(2):107–122.

Cheng, Y., Ding, P., Wang, T., Lu, W., and Du, X. (2019). Which category is better: benchmarking relational and graph database management systems. *Data Science and Engineering*, 4(4):309–322.

Đurić, Z. and Gašević, D. (2013). A Source Code Similarity System for Plagiarism Detection. *The Computer Journal*, 56(1):70–86.

Donaldson, J. L., Lancaster, A.-M., and Sposato, P. H. (1981). A plagiarism detection system. *ACM SIGCSE Bulletin*, 13(1):21–25.

Encyclopædia Britannica Inc (2015). Plagiarism. Accessed on 2015-09-10 Available at http://www.britannica.com/topic/plagiarism.

Esser, S. and Fahland, D. (2021). Multi-dimensional event data in graph databases. *Journal on Data Semantics*, 10(1):109–141.

Grune, D. and Huntjens, M. (1989). Het detecteren van kopieen bij informatica-practica. *Informatie*, 31(11):864–867.

Joy, M. and Luck, M. (1999). Plagiarism in programming assignments. *IEEE Transactions on Education*, 42(2):129–133.

Kermek, D. and Novak, M. (2016). Process Model Improvement for Source Code Plagiarism Detection in Student Programming Assignments. *Informatics in Education*, 15(1):103–126.

Kleiman, A. and Kowaltowski, T. (2009). Qualitative Analysis and Comparison of Plagiarism-Detection Systems in Student Programs - Technical Report IC-09-08. Technical report, Institu to de Computação, Universidade Estadual de Campinas.

Martins, V. T., Henriques, P. R., and da Cruz, D. (2015). An AST-based Tool, Spector, for Plagiarism Detection: The Approach, Functionality, and Implementation. *Communications in Computer and Information Science*, 563:153–159.

Miller, J. J. (2013). Graph database applications and concepts with neo4j. In *Proceedings of the southern association for information systems conference, Atlanta, GA, USA*, number 36 in 2324.

Novak, M. (2020). *Effect of source-code preprocessing techniques on plagiarism detection accuracy in student programming assignments*. PhD thesis, Sveučilište u Zagrebu. Fakultet organizacije i informatike.

Novak, M., Joy, M., and Kermek, D. (2019). Source-code similarity detection and detection tools used in academia: A systematic review. *ACM Transactions on Computing Education*, 19(3):27:1–27:37.

Novak, M., Joy, M. S., and Mirza, O. M. (2021). Improved plagiarism detection with collaboration network visualization based on source-code similarity. In *2021 IEEE Technology & Engineering Management Conference-Europe (TEMSCON-EUR)*, pages 1–6. IEEE.

Prechelt, L., Malpohl, G., and Philippsen, M. (2002). Finding plagiarisms among a set of programs with JPlag. *Journal of Universal Computer Science*, 8(11):1016–1038.

Ragkhitwetsagul, C., Krinke, J., and Clark, D. (2016). Similarity of Source Code in the Presence of Pervasive Modifications. In *IEEE 16th International Working Conference on Source Code Analysis and Manipulation*, pages 117–126. IEEE.

Robinson, I., Webber, J., and Eifrem, E. (2015). *Graph databases: new opportunities for connected data*. " O'Reilly Media, Inc.".

Timón-Reina, S., Rincón, M., and Martínez-Tomás, R. (2021). An overview of graph databases and their applications in the biomedical domain. *Database*, 2021.

Tyagi, D., Arora, R., and Sharma, Y. (2022). Application of java relationship graphs to academics for detection of plagiarism in java projects. In *ICT Systems and Sustainability*, pages 761–772. Springer.

Šestak, M., Heričko, M., DruÅ¾ovec, T. W., and Turkanovič, M. (2021). Applying k-vertex cardinality constraints on a neo4j graph database. *Future Generation Computer Systems*, 115:459–474.

Evaluating the UX of a Microservice Web App: CyberRanges in the µDevOps Project

Luis de-Marcos, José-María Gutiérrez-Martínez, Adrián Domínguez-Díaz, Sergio Caro-Álvaro, Daniel Rodríguez

Universidad de Alcalá. Dpto. de Ciencias de la Computación.

Edificio Politécnico. Campus Universitario. Alcalá de Henares. 28805. Madrid. Spain

{luis.demarcos, josem.gutierrez, adrian.dominguez, sergio.caro, daniel.rodriguezg}@uah.es

Abstract. *Microservice architectures are becoming increasingly important since they facilitate agile and modular production cycles to deliver applications using collections of loosely coupled and fine-grained services. The µDevOps is a research project formed by an international network of organizations including industry and academia that aims to tackle current challenges of microservice development operations. This paper presents the µDevOps project and the initial research carried out to evaluate the user experience of a microservice web application that delivers cybersecurity learning. Results point to critical elements of three main functionalities: library of scenarios, scenario information and entering scenario. Since there are several currently available solutions that offer similar services, the user experience of the microservice web app may play a critical role in determining which application will get a dominant role in the market.*

Keywords. User experience, microservice, usability, web application, cyberrange.

1 Introduction

Software industry has quickly moved toward ultra-agile and modular production cycles to deliver flexible, scalable, and user-centric applications in a timely way. A DevOps (Development Operations) process is a recent paradigm that puts together important advancements at architectural design, development, and deployment level. At architectural level, Microservices are an important evolution of service-oriented architectures aiming at structuring applications as a collection of loosely coupled and fine-grained services: it fosters a high modularity to make the application easier to understand, develop and test, and to facilitate the deployment, execution, and maintenance. At development and deployment level, DevOps stresses the agile code production by engaging the operational phase earlier in the development cycle: it fosters communication, collaboration and integration between development and IT operation teams, to favour rapid and continuous delivery cycles. The whole process is supported by new technologies (e.g., containers for deployment and cloud platforms at infrastructural level), which greatly alleviate several manual tasks, saving technical stakeholders time to deliver greater value. Even though Microservices and DevOps originate independently, they share the same set of principles and cultural background, stressing concepts like agility, flexibility, scalability, automation, user-oriented development, and cloud-based provisioning. Today they are viewed as a complement to each other, and likely, this production paradigm will underlie many software applications in the next years.

While researchers and practitioners have well caught the advantages of the interplay between the development and IT operation teams, this is by far not true for the third pillar of µDevOps, namely the quality assurance team. Software quality (SQ) is a pivotal asset and a key business driver for today's IT industry (Jones & Bonsignour, 2011). Innovation in this field has significant market opportunities for non-academic participants in the project. SQ embodies a wide set of non-functional requirements (in this context identified as Quality-of-Service (QoS) attributes) deemed crucial by end users, such as: security, reliability, performance, availability, usability. "Failures" in satisfying these requirements entail a large part of the business risk associated with a software product, as they can determine completely its success or failure. The continuous awareness, during development or operation, of user-perceived quality in the context where the system operates is paramount for decision-makers.

The µDevOps project aims to address all these challenges by forming an international and inter-sectoral network of organisations working on a joint research programme in the field of Software Quality Assurance (SQA) for Microservice Development Operations (we coin it as µDevOps) engineering processes. This paper presents µDevOps project and the initial findings of the work done to evaluate the

UX of a microservices web application. Since the UX plays a significant role in the success of an application, it is deemed critical to assess and create an infrastructure to continuously evaluate user experience. The rest of the paper is structured as follows. Section 2 introduces the µDevOps project. Section 3 presents the state of the art. Section 4 briefly outlines the approach. Section 5 reports initial findings. The paper concludes with final remarks and future research.

2 The µDevOps Project

Boundaries between development and operations are blurred and a highly dynamic environment demands for a new view of the role of testing for SQA. The µDevOps project grounds on the idea that the operational-time contextual information is destined to be predominant, and any realistic quality practice will not work if disregards it. The µDevOps research and innovation action intends to tackle the described challenges by promoting a cultural shift that places testing as a key cornerstone between the development and operation phases

This would amplify the intertwining and synergy between development and operations in the µDevOps paradigm, but requires a mind-set change that this project intends to promote. Toward this long-term goal, the proposers will implement a joint research program based on knowledge exchange between academic and industrial partners. The knowledge exchange program will revolve around the study, definition, implementation, and validation of a context-driven risk-based SQA testing process in µDevOps, including four scientific objectives:

- Context Learning: the objective is to study and develop new modelling and learning strategies to dynamically characterize the operational context of the microservice-based software under test at runtime
- Continuous in vivo testing: the objective is to define and develop techniques for in vivo tests generation and execution for µDevOps environment
- Risk Assessment to risk-based SQA in µDevOps: the objective is to study and develop metrics and methods to provide quantitative measures of the business risk of using a given functionality.
- µDevOps Development & Testing as a Service (D&TaaS): the objective is to make everything easily accessible and usable by a cloud infrastructure where both µDevOps-based development and testing will be offered as a service.

We structured the research programme in Work Packages. Figure 1 shows the work packages and the workflow of the project, providing an overview of the inter-relationship between WPs.

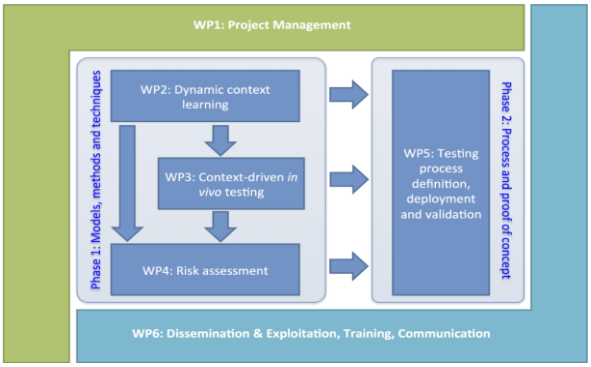


Figure 1. µDevOps project. Workflow of work packages

3 Testing and SQA in Microservices

Testing has been extensively investigated in the software architecture area, but it is still an open problem when dealing with microservice-based systems (Jones & Bonsignour, 2011). Indeed, testing microservice-based systems is extremely challenging, mainly because tools must be agnostic to the programming language and runtime environment of each microservice and because of the extremely rapid evolution of the business logic of each microservice within the system (Di Francesco et al., 2019). A limited number of testing approaches have been proposed. Schermann et al. (2016) propose “Bifrost”, a formal model for defining and automatically enacting live testing strategies for supporting multi-phase release strategies of microservice-based systems, with acceptable performance overhead under many release strategies in parallel. Release strategies are defined using a YAML-based Domain-Specific Language (DSL), thus allowing strategies to be transparently shared, reused, and versioned across projects and organisations. Heorhiadi et al. (2016) presented “Gremlin”, a purely network-oriented, systematic resiliency-testing framework inspired by software-defined networks. In Gremlin, a centralized control plane allows operators to provide high-level outage scenarios and assertions on how microservices should react during outages. Then, Gremlin automatically translates this information into a fixed set of fault injection rules applied to the network messages exchanged between microservices via network proxies. Experiments show that Gremlin can provide low-latency feedback to operators, and that the learning curve for creating scenarios and assertions is minimal. An architecture for automating acceptance testing in the context of microservice-based systems is proposed in (Schermann et al., 2016). The architecture heavily relies on Behaviour-Driven Development (BDD), where acceptance tests are defined as scenarios conforming to a simple syntax, even accessible by business stakeholders. The approach ensures that all microservices meet

requirements related to reusability, auditability, and maintainability. For what concerns SQA of microservices, performance (e.g., (Mazedur & Gao, 2015; Holger, 2016)) and maintainability (e.g., (Düllmann et al., 2017a; Hasselbring, 2017)) are the most investigated quality attributes in the state of the art. Many methods and techniques focussing on system performance have a special focus on scalability (Jones & Bonsignour, 2011), suggesting that researchers consider scalability as a sub-problem of performance when architecting with microservices. For example, O'Connor et al., 2017 presented a polyglot (i.e., supporting microservices developed in different languages, e.g., Java, JavaScript) auto scaling technique for the IBM Bluemix platform. The technique enables engineers to describe policies and set thresholds for scaling microservice-based applications based on their CPU, memory, and heap usage, and supports a shared multi-tenancy deployment model and it is delivered as managed cloud services. For what concerns maintainability, the relatively high scientific interest it is attracting might be related to the highly distributed nature of microservice-based systems and their inherent focus on services independence. As an example, Nuha et al. (2016) presented "MicroART", a technique for semi-automatically reconstructing the architecture of microservice-based systems to ease their maintenance and comprehension. MicroART is based on the Model-Driven Engineering paradigm (Seelam et al., 2015) and it reconstructs the software architecture of the system by mining the GitHub repositories with microservices source code and reusing existing monitoring platforms (e.g., New Relic). Other qualities are less explored with respect to performance and maintainability, including services compatibility (e.g., (Granchelli., 2017, Schmidt & Douglas, 2006)), security (e.g., (Sara et al., 2017, Viennot et al., 2015)), portability (e.g., (Yuqiong et al., 2015, Yahia et al., 2016)), or organisational alignment (Linthicum, 2016).

Relevant DevOps practices that relate to testing and SQA are continuous delivery (Justus & Zimmermann, 2016) and continuous monitoring (Jambunathan & Kalpana, 2016). In continuous delivery pipelines, testing activities are automated, using technologies as Cucumber or Selenium (Farley & Humble, 2010). However, it is an open challenge how to select the right load tests in such pipelines: with frequent releases, the capacity for executing load tests become critical and the need to select the right tests for maximum insight becomes crucial. One solution adopted in practice is canary testing, in which a new release is delivered to a few users for live testing. Ongoing efforts aim at providing more systematic support, such as the architecture-based performance engineering platform CASPA (van Hoorn et al., 2009). Currently, the main aim is to identify regressions over time (Aderaldo et al., 2017). No solutions have been proposed yet to systematically test systems for unexpected contexts. Overall, the activities are still usually separated between development and operation: while models and predictions may be used at development time to support design decisions, measurements-based approaches such as application performance management and monitoring are used at operation time. Initial approaches to reconcile both have been suggested, such as iObserve (Düllmann et al., 2017b) and CIPM (Mazkatli & Koziolek, 2018). Security testing in DevOps is also a relevant concern. There are tools helping to extract and automate tests to check for security issues, e.g., BDD-Security, Mittn by F-Secure, Contrast Security and Gauntlt. These are usually coupled with other means, like code analysis (e.g., tools like Veracode to scan the code to find vulnerabilities), runtime checks, cloud infrastructure best practices to check for configurations security best practices (e.g., Microsoft Azure Advisor, evident.io). DevSecOps refers to the practices to build security testing into DevOps. Reliability testing is much less investigated. An approach indirectly useful for reliability is the WESSBAS approach for the usage profile modelling (Vögele et al., 2016), which extracts probabilistic workload from measurements at operation time to inform models and decision-making at development-time.

4 Case Study: The CyberRanges Web Application

A cyber range is a virtual environment commonly used to provide a secure, legal environment for cybersecurity education, practice, and training. Isolation from threats is ensured by providing trainees the ability to recognize and respond to real-world challenges in a controlled environment. This approach guarantees that client infrastructure and data is never at risk because of the cybersecurity training (Aries Security, 2020).

CyberRanges[1] is a web application (Figure 2) developed by Silensec that delivers cyber security training and capability development exercises using technology and services for the design, delivery, and management of simulation-based, deep-dive experiences. It can be used to learn, train, test, measure, and improve the digital dexterity and cyber resilience of professionals, teams, and organizations by using a platform and technology that resembles a real-world scenario. CyberRanges uses a microservice architecture that strongly focuses on the functionality of the platform.

Since there are several competitors offering cyber range solutions, to get a dominant position is not only required to provide the better offering, in terms of functionality, library of scenarios and user base. But it

[1] https://www.cyberranges.com/

is also necessary to offer a seamless user experience that provides a meaningful and relevant interaction. Competitors to consider and compare include openly accessible solutions like TryHackMe[2] and HackTheBox[3], and commercial solutions that require a paid account like RangeForce[4] or ImmersiveLabs[5]. Similarly, online teaching and learning marketplaces like Udemy provide compelling and intuitive user experiences but lack of simulation required for cyber ranges.

The development team of the company assessed the UX strengths and weaknesses of the CyberRanges application based on the skills and competences of their staff. They identified a set of needs that could be addressed by the researchers. These needs can be summarized as follows: (1) Evaluate and redesign the UX for a more user-friendly experience, particularly simplifying the interface for the construction of scenarios and the UX inside them by suggesting changes to the actionable UI elements. For instance, changes are required in the interface to enter a scenario, in the library of scenarios (e.g., provide a more organic filter, better scrolling, and more on-screen information) as well as a possible redesign the statistics web page. (2) Get a better understanding of the process to build improve UX including the stages and steps (e.g., analysis, validation, etc.) to make it agile and incorporate feedback effectively. (3) Define the data to bring out and use for continuous improvement of UX (e.g., logs). Secondary less urgent needs include improving the analytics by suggesting new charts or changes to existing ones and analyzing the gamified elements and gameful design of the platform.

The case study presented here offers significant research possibilities in terms of the experimental testing and evaluation of the platform. Since cyberranges.com is going to publicly available soon, this opens the possibility for massive data collection. It would be possible soon to design and run experiments related to UX and user interaction.

[2] https://tryhackme.com
[3] https://www.hackthebox.com
[4] https://www.rangeforce.com
[5] https://www.immersivelabs.com

5 Approach for User Experience Evaluation

The sheer number of existing usability and UX evaluation methods includes, among others, observation, self-reporting, visual designs, idea descriptions, interaction analysis, lab tests, field studies and market feedback. The selection of one or other usually relies on the phase of product development, period of experience to evaluate and desired feedback. For instance, a visual design can be useful to assess the emotional expressions and reactions of a non-functional prototype or concept, while a field study (e.g., experience sampling or day reconstruction method) can be used to evaluate the long-term experience of a functional prototype.

Among the myriad of approaches, the μDevOps project opted for expert evaluation using sources like Yablonski's (2020) laws of UX and usability engineering. Given current state of the application and deployment plans for the subsequent months, we chose a method that can expedite the feedback required by the developers to improve the UX before launch. The usability evaluation was run by a group of three experts. Evaluation was performed during one month during Spring 2022. Experts meet initially to determine the goals, parameters, and steps for the study. They followed the template suggested by the Nielsen Norman Group[6]. The usability evaluation focused on the following functionalities: library of scenarios, scenario information, and entering scenario. These core functionalities to be subject of evaluation were pointed by the company development team based on their perceptions of what are more critical for their users. Each expert evaluated the user interface separately and they all joined to write the final report. The remaining of the paper summarizes the findings on which all evaluators agreed in their final report.

To further improve the evaluation process in the future, we can also explore the existing logs of the web application and implement the necessary changes in those logs to gather meaningful information about the user interaction that can be used to further improve the user experience.

[6] NN/g Checklist for Planning Usability Studies. https://www.nngroup.com/articles/usability-test-checklist/

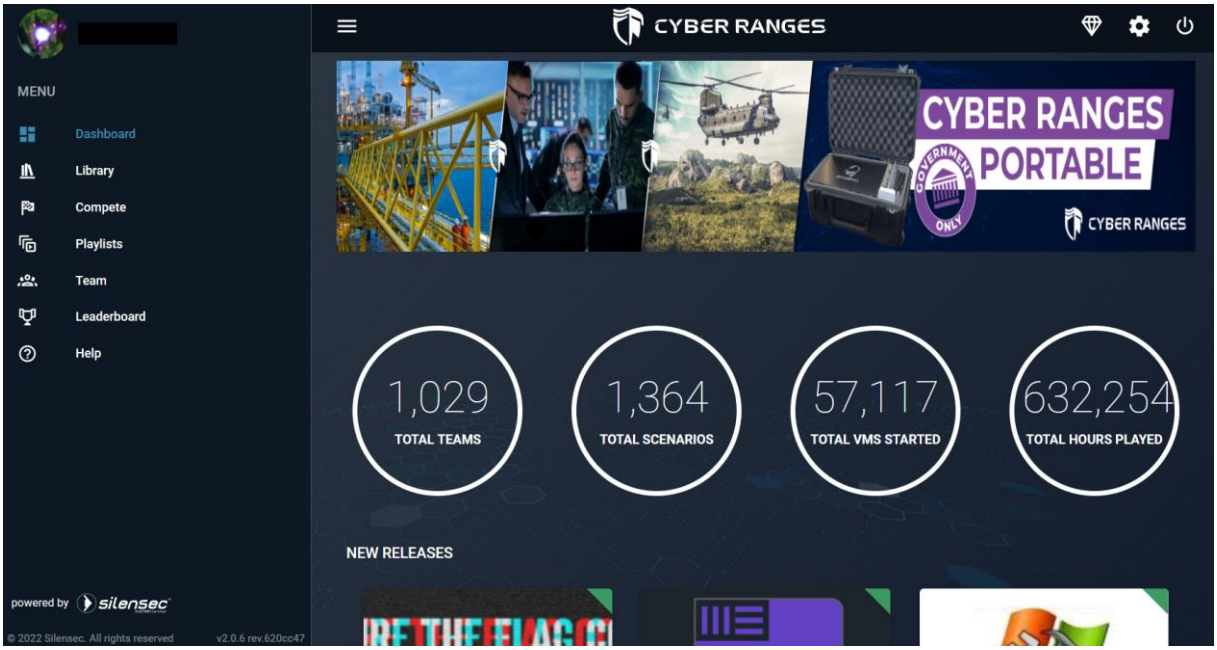


Figure 2. Cyberranges web app dashboard

6 Initial Findings

This section provides examples of the results of the evaluation of the user experience for three functionalities of the CyberRanges application: library of scenarios (Figure 3), scenario information (Figure 4) and entering scenario (Figure 5). We report here the functionalities that returned more issues according to the experts' evaluation.

The library of scenarios implements horizontal browsing for the elements which presents two usability issues. First, common usability heuristics suggests that horizontal scrolling is not recommended to present information. Second, the interface elements (buttons) that enable the horizontal browsing are only visible upon rollover of the mouse cursor. Also, in the library of scenarios assessment it was noticed that, upon application of a filter, the presentation of the list of scenarios changes substantially. The suggestion is to present the filtered output using a similar layout.

In the interface that shows information about scenarios, assessment found that the categories shown are different to the ones showed in the library. Apparently, the library contains tags for the whole collection while the scenario information screen shows the keywords. This is confusing. Suggestions include: (1) include the tags of the library also in the scenario information screen, (2) make clear that they represent different sets, and (3) make them clickable so that the user can access all the scenarios with same tags and keywords with just one action. In the scenario information interface, experts also suggest including an element (e.g., a button) to return to the library. In the current implementation the user needs to use the main menu on the right.

In the scenario loading user interface, it was suggested to improve the visibility of the loading bar. Recommendations are particularly focus on the fact that loading a scenario usually take a long time. The current bar's size and location on the top of the screen may prevent the user to see it initially, thus missing the important information it shows at a critical moment. The bar contrast with big 'Loading…' label at the bottom right that captures initial user attention. A wider suggestion is to rearrange the scenario loading UI. The screen presents parts that are mostly empty. Although the interface has a good approach, using tabs to present the mission, rules and leaderboards, designers may consider including further information and fill the empty spaces if needed, since the user may spend a substantial amount of time in this screen while the scenario loads. This may probably require motivating scenario creators to include all the additional information, which may be done by informing them of the space and time available. Similarly, the scenario loading screen could be used to provide required technical information that the user needs to complete it. For instance, the configuration of the VPN is shown on the main screen, but it would also be a good idea to include other reminders or links while loading.

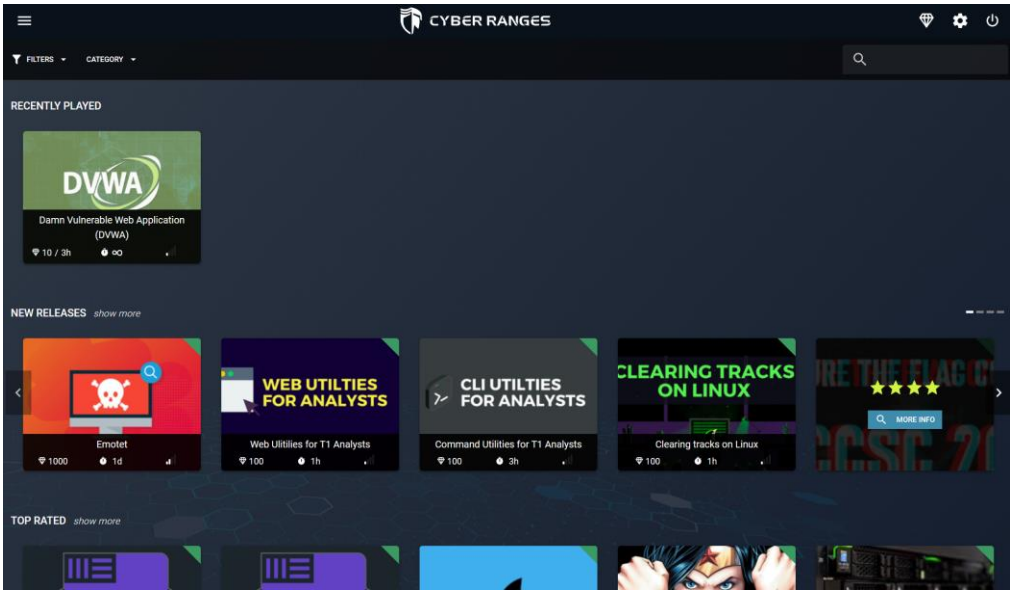


Figure 3. Scenario information

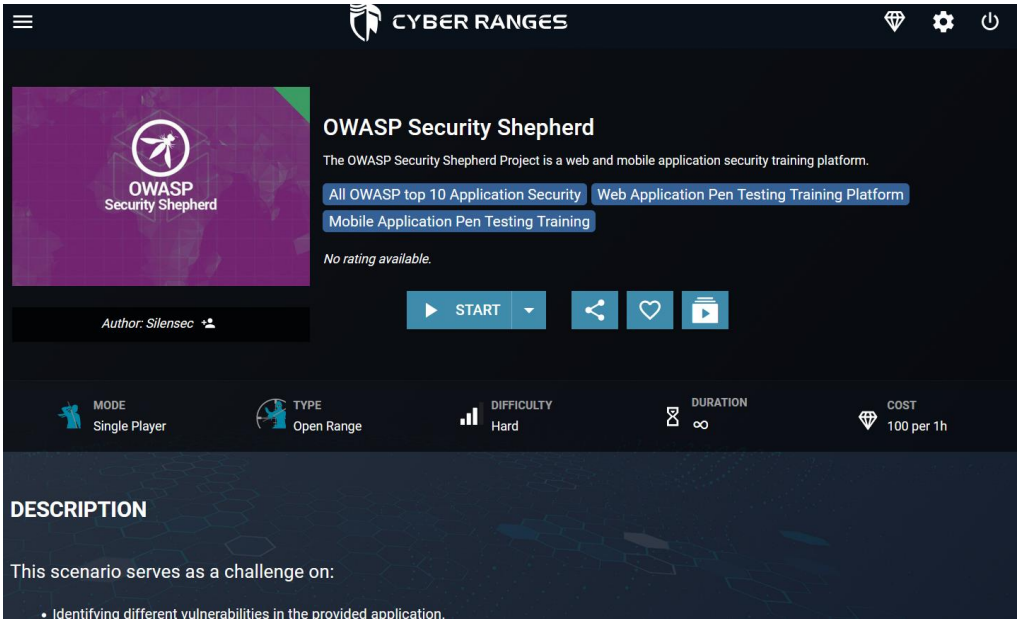


Figure 4. Scenario information

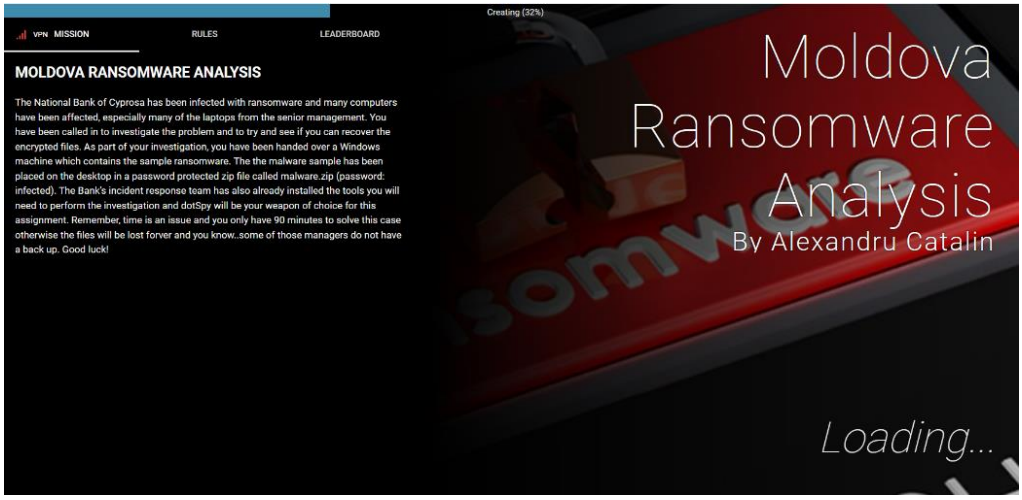


Figure 5. Scenario loading

7 Conclusion & Future Work

This paper described the initial approach and findings for the evaluation of the UX for experiential learning in the context of the CyberRanges web application for the µDevOps project. We initially presented the challenges of the microservices architectures that the µDevOps aims to address. We then introduced the case study of CyberRanges web application and the need for user experience evaluation in the context of microservices architectures. Initial findings suggest several improvements for critical functionalities of the application including the library of scenarios, the scenario information screen, and the scenario loading interface. Suggestions aim at dealing with important usability issues and improving the user experience. Since there are several commercial applications that offer cyber ranges and similar cybersecurity trainings, the user experience may play key role in the competition for determining which groups of solutions takes a significant part of the market share.

Future research can further dive into other approaches to UX evaluation that could be especially suitable for scientific objectives of the µDevOps project. Focusing on the first scientific objective, collecting real-time UX metrics could help to dynamically characterize the operational context of the microservice-based application, in terms of usage profile. This information could be used to improve the SQA processes by focusing testing efforts on those specific features in which more user experience issues are found. The integration of this data with artificial intelligence and machine learning techniques could help to create user models that predict UX problems, as well as to generate in-vivo test based on UX collected data and found issues. Another interesting problem in terms of UI and UX is to automatically create and manage replicas of corporate apps for CyberRanges exercises and other cybersecurity training.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 871342

References

Aderaldo, C. M., Mendonça, N. C., Pahl, C., & Jamshidi, P. (2017, May). Benchmark requirements for microservices architecture research. In Proceedings of the 1st International Workshop on Establishing the Community-Wide Infrastructure for Architecture-Based Software Engineering (pp. 8-13). IEEE Press.

Aries Security. (2020). What is a Cyber Range? A Definitive Guide and Definition. https://www.ariessecurity.com/what-is-a-cyber-range-a-definitive-guide-and-definition/ [Accessed on 24th May 2022]

Di Francesco, P., Malavolta, I., Lago, P. Architecting with Microservices: a Systematic Mapping Study, Journal of Systemsand Software, 2019.

Düllmann, Thomas F., and André van Hoorn. Model-driven Generation of Microservice Architectures for Benchmarking Performance and Resilience Engineering Approaches. Proceedings of the 8th ACM/SPEC on International Conference on Performance Engineering Companion. ACM, 2017.

Thomas F. Düllmann, Robert Heinrich, Andre van Hoorn, Teerat Pitakrat, Jürgen Walter, and Felix Willnecker. Caspa: A platform for comparability of architecture-based software performance engineering approaches. In Proceedings of the 2017 IEEE International Conference on Software Architecture (ICSA 2017), 2017. IEEE Internation Conference on Software Architecture Workshops. 2017.

D. Farley and J. Humble, Continuous Delivery: Reliable software releases through Build, Test and Deployment Automation . Addison-Wesley Professional, 2010.

Granchelli, G., Cardarelli, M., Di Francesco, P., Malavolta, I., Iovino, L., Di Salle, A. MicroART: A Software Architecture Recovery Tool forMaintaining Microservice-based Systems. In Software Architecture Workshops (ICSAW), 2017 IEEE International Conference on (pp. 298-302). IEEE.

Hasselbring, Wilhelm, and Steinacker. Microservice architectures for scalability, agility and reliability in E-commerce. Software Architecture Workshops (ICSAW), 2017 IEEE International Conference on. IEEE, 2017.ç

Heorhiadi, V., Rajagopalan, S., Jamjoom, H., Reiter, M. K., & Sekar, V. (2016, June). Gremlin: systematic resilience testing of microservices. IEEE 36th International Conference on Distributed Computing Systems (ICDCS), 2016 (pp. 57-66).

K. Holger. Sustaining runtime performance while incrementally modernizing transactional monolithic software towards microservices. Proceedings of the 7th ACM/SPEC on International Conference on Performance Engineering. ACM, 2016.

Baskaran Jambunathan and Y Kalpana. Multi cloud deployment with containers. International Journal of Engineering and Technology (IJET), 8(1):421–428, 2016.

C. Jones, O. Bonsignour, The Economics of Software Quality, FINANCIAL TIMES/PRENTICE HALL, 2011.

Bogner Justus, and Alfred Zimmermann. "Towards integrating microservices with adaptable enterprise architecture." Enterprise Distributed Object Computing Workshop (EDOCW), 2016 IEEE 20th International. IEEE, 2016.

Linthicum, David S. "Practical use of microservices in moving workloads to the cloud." IEEE Cloud Computing 3.5 (2016): 6-9.

Rahman Mazedur, Jerry Gao. A reusable automated acceptance testing architecture for microservices in behavior-driven development. IEEE Symposium on Service-Oriented System Engineering (SOSE), IEEE, 2015.

Manar Mazkatli and Anne Koziolek. Continuous integration of performance model. In The 9th ACM/SPEC on International Conference on Performance Engineering Companion, Berlin, Germany, 2018, ICPE '18 Companion. ACM, New York, NY, USA. 2018.

Alshuqayran Nuha, Nour Ali, and Roger Evans. A systematic mapping study in microservice architecture. 9th International Conference on Service-Oriented Computing and Applications (SOCA). IEEE, 2016.

O'Connor Rory V., Peter Elger, and Paul M. Clarke. Continuous software engineering —A microservices architecture perspective. Journal of Software: Evolution and Process 29.11 (2017).

Hassan Sara, Nour Ali, and Rami Bahsoon. "Microservice Ambients: An Architectural Meta-modelling Approach for Microservice Granularity." Software Architecture (ICSA), 2017 IEEE International Conference on. IEEE, 2017.

Gerald Schermann, Dominik Schöni, Philipp Leitner, and Harald C. Gall. Bifrost: Supporting Continuous Deployment with Automated Enactment of Multi-Phase Live Testing Strategies. In Proceedings of the 17th International Middleware Conference (Middleware). 2016. ACM, New York, NY, USA.

Schmidt, Douglas C. "Model-driven engineering." COMPUTER-IEEE COMPUTER SOCIETY-39.2 (2006): 25.

Seelam, S. R., Dettori, P., Westerink, P., & Yang, B. B. (2015, March). Polyglot Application Auto Scaling Service for Platform as a Service Cloud. In Cloud Engineering (IC2E), 2015 IEEE International Conference on (pp. 84-91). IEEE.

A. van Hoorn, M. Rohr, W. Hasselbring, J. Waller, J. Ehlers, S. Frey, and D. Kieselhorst, “Continuous monitoring of software services: Design and application of the kieker framework,” 2009.

Viennot, N., Lécuyer, M., Bell, J., Geambasu, R., & Nieh, J. (2015, April). Synapse: a microservices architecture for heterogeneous-database web applications. In Proceedings of the Tenth European Conference on Computer Systems (p. 21). ACM.

Vögele, C., van Hoorn, A., Schulz, E., Hasselbring, W., & Krcmar, H. (2016). WESSBAS: Extraction of probabilistic workload specifications for load testing and performance prediction—A model-driven approach for session-based application systems. Software & Systems Modeling, 1-35

Yablonski, J. (2020). Laws of UX. O'Reilly Media.

Yahia, E. B. H., Réveillère, L., Bromberg, Y. D., Chevalier, R., & Cadot, A. Medley: An event-driven lightweight platform for service composition. In International Conference on Web Engineering (pp. 3-20). Springer, 2016.

Sun Yuqiong, Susanta Nanda, and Trent Jaeger. "Security-as-a-service for microservices-based cloud applications." Cloud Computing Technology and Science (CloudCom), 2015 IEEE 7th International Conference on. IEEE, 2015.

Threat Modeling Methods in the Medical Device Industry: An Integrative Literature Review

Nadica Hrgarek Lechner, Vjeran Strahonja, Zlatko Stapić
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{nhrgarek, vjeran.strahonja, zstapic}@foi.hr

Abstract. *Threat modeling is a structured information generating process that begins early in the development life cycle to identify potential security threats to a system being built. Threat modeling for medical devices is a relatively new area of research. The medical device industry recognizes the benefits of performing threat modeling throughout the product life cycle by adopting threat modeling to systematically uncover threats and proactively develop secure medical devices. This paper presents the results of an integrative literature review of 26 relevant studies. A total of 32 threat modeling methods and approaches in the medical device industry were identified and systematized.*

Keywords. attack trees, cybersecurity, integrative literature review, medical devices, medical device industry, STRIDE, threat model, threat modeling, threat modeling methods

1 Introduction

The early beginnings of the study of threats were not in science but industry. In the late 1990s, Microsoft employees Loren Kohnfelder and Praerit Garg published an internal paper that described a taxonomy of threats to software (Kohnfelder, 2022). They named it the S.T.R.I.D.E. security threat model to address six major categories of threats: Spoofing of user identity, Tampering with data, Repudiability, Information disclosure (privacy breach), Denial of Service (D.o.S.), and Elevation of privilege (Kohnfelder & Garg, 1999). Threat modeling has become a crucial element of the Microsoft Security Development Lifecycle (SDL). Microsoft has released the Threat Modeling Tool ("Microsoft Threat Modeling Tool," 2020) as a free application for Windows to guide the development teams through the threat modeling process using STRIDE.

While threat modeling is commonly used in software development projects as part of a secure development lifecycle (Howard & Lipner, 2006), it is a relatively new topic within the medical device industry. Medical device manufacturers need to design and deliver safe and secure products, and have a long experience with the safety risk management process required by ISO 14971 standard (*ISO 14971: Medical devices – Application of risk management to medical devices*, 2019). However, security risk management in the medical domain is a relatively new topic that was addressed by the FDA (Food and Drug Administration), a federal agency of the United States responsible for protecting public health. FDA (2014) issued the first guidance that specifically addresses the management of cybersecurity in medical devices.

New draft guidance (FDA, 2022, p. 10) recommends threat modeling to be performed as part of the security risk assessment to inform and support the risk analysis activities. In the context of medical devices, threat modeling identifies threats that could adversely impact the safety and security of a medical device (The MITRE Corporation & MDIC, 2021, p. 48). According to the technical information report (*AAMI TIR57: Principles for medical device security–Risk management*, 2019), the first activity of the security risk management process is security risk assessment consisting of security risk analysis and security risk evaluation.

Several cybersecurity guidances (FDA, 2016), (FDA, 2018), (FDA, 2022), (IMDRF, 2020), (MDCG, 2020), (NMPA, 2022), (SFDA, 2019), (TGA, 2021), ("IT Security Guideline for Medical Devices," 2021), standard (*IEC 81001-5-1: Health software and health IT systems safety, effectiveness and security – Part 5-1: Security – Activities in the product life cycle*, 2021), and other publications (*AAMI TIR57: Principles for medical device security–Risk management*, 2019), (HSCC, 2019), (The MITRE Corporation & MDIC, 2021) for the medical device industry emphasize the importance of threat modeling which can be used as part of security risk assessments.

A threat model should be the output of a systematic approach (Ray, 2021, p. 141) for identifying assets, vulnerabilities, and threats. Unstructured brainstorming sessions and other non-systematic approaches for identifying potential threats and

vulnerabilities are not acceptable for medical devices because they cannot produce an exhaustive list of threats.

Threat modeling can be applied to software, devices, systems, networks, distributed systems, and business processes (*IEC 81001-5-1: Health software and health IT systems safety, effectiveness and security – Part 5-1: Security – Activities in the product life cycle*, 2021). According to Shostack (2014, pp. xxiii-xxiv), there are many reasons to use threat modeling: finding security bugs early, understanding security requirements, engineering and delivering better products, and addressing issues other techniques won't find. In the context of software development, Howard & Lipner (2006, p. 102) list the following benefits of threat modeling: contributes to the risk management process because threats to software and infrastructure are risks to the user and environment deploying the software, uncovers threats to the system before the system is committed to code, revalidates the architecture and design by having the development team go over the design again, forces development staff to look at the design from a different viewpoint– that of security and privacy, helps clarify the selection of appropriate countermeasures for the application and environment, contributes to the Attack Surface Reduction (ASR) process for the software, helps guide the code review process, and guides the penetration testing process. When done right, threat modeling can assist cybersecurity professionals, developers, and subject matter experts during the security risk assessment of a medical device.

FDA (2022, p. 11) recommends that premarket submissions include threat modeling documentation and does not prescribe any specific methodology or method for threat modeling. There are many threat modeling methods, whose comprehensive summary is given by Shevchenko et al. (2018), citing STRIDE and associated derivations, PASTA (Process for Attack Simulation and Threat Analysis), attack trees, LINDUN (Linkability, Identifiability, Non-repudiation, Detectability, Disclosure of information, Unawareness, Noncompliance), CVSS (Common Vulnerability Scoring System), persona non grata, security cards, hTMM (Hybrid Threat Modeling Method), Quantitative TMM (Threat Modeling Method), Trike, VAST (Visual, Agile, and Simple Threat) modeling, and OCTAVE (Operationally Critical Threat, Asset, and Vulnerability Evaluation). According to (Shevchenko et al., 2018), threat modeling methods are used to create an abstraction of the system, profiles of potential attackers, and a catalog of potential threats that may arise. Three main approaches in threat modeling that help highlight threats that may be present in a system are: system-centric approach, attacker-centric approach, and asset-centric approach (Tarandach & Coles, 2020).

Having the importance of threat modeling explained, the question that arises is what methods and approaches could be used in threat modeling, particularly in the medical device industry. Thus, the purpose of this paper is to present an integrative review of published literature on the different threat modeling methods and approaches in the medical device industry.

The rest of the paper is organized as follows. The research methodology is described in section 2. Section 3 presents the results from the conducted integrative literature review and a discussion of the findings. The final section of the paper restates the research problem, summarizes the findings, discusses the implications, and provides recommendations for future research directions.

2 Methodology

An integrative literature review is a form of research that reviews, critiques, and synthesizes representative literature on a topic in an integrated way such that new frameworks and perspectives on the topic are generated (Torraco, 2005, p. 356). In cases where a research question requires a more creative collection of data, an integrative review approach can be useful when the purpose of the review is not to cover all articles ever published on the topic but rather to combine perspectives and insights from different fields or research traditions (Snyder, 2019).

We followed the integrative literature review methodology proposed by Whittemore & Knafl (2005). In their methodology, the integrative literature review consists of five stages: problem identification, literature search, data evaluation, data analysis, and presentation. The integrative literature review was chosen because it includes research articles, books, and other published texts (Snyder, 2019).

2.1 Problem Identification

Malamas et al. (2021) provide an overview of IoMT (Internet of Medical Things) threat models found in the scientific literature. Vakhter et al. (2022) proposed a domain-specific qualitative-quantitative threat model for miniaturized wireless biomedical devices and compared their proposed model to four other threat models. To the best of the authors' knowledge, a review and synthesis of different threat modeling methods and approaches in the medical device industry does not exist in published literature.

We aim to answer the following research question: Which threat modeling methods and approaches can be used in the medical device industry? To answer the research question, we conducted an integrative literature review.

To get an overview of the research topic, a basic Topic search for the exact phrase "threat modeling" was performed in the Clarivate Analytics' Web of Science Core Collection citation database and returned 356 publications. The search range was from 1955 to August 16th, 2022. The search results were refined by

the publication date range as listed in Table 1. Web of Science Core Collection was chosen because every journal and book covered by Web of Science Core Collection is assigned to at least one or more Web of Science categories.

Table 1. Distribution of publications by publication date range

Publication date range	Number of publications	% of 356
2005–2010	14	3.93
2011–2015	73	20.51
2016–2020	189	53.09
2021–16/08/2022	80	22.47
Total	**356**	**100.00**

First publications on threat modeling were published between 2005 and 2010. A significant increase in the number of published publications was noted between 2016 and 2020. This increase may be related to cybersecurity threats which continue to grow year over year. In the time between January 2021 and August 16th, 2022, 80 publications were published and the number is expected to increase over the coming years.

The search results were analyzed according to the Web of Science categories. The top 25 of the total 63 categories are presented and sorted in descending order according to the number of occurrences in Table 2.

Table 2. Distribution of publications by the top 25 of the total 63 Web of Science categories

No.	Web of Science categories	Number of publications	% of 356
1	Computer Science Information Systems	149	41.85
2	Computer Science Theory Methods	132	37.08
3	Engineering Electrical Electronic	91	25.56
4	Computer Science Software Engineering	82	23.03
5	Telecommunications	66	18.54
6	Computer Science Interdisciplinary Applications	29	8.15
7	Computer Science Hardware Architecture	23	6.46
8	Computer Science Artificial Intelligence	22	6.18
9	Engineering Multidisciplinary	11	3.09
10	Automation Control Systems	7	1.97
11	Multidisciplinary Sciences	7	1.97
12	Engineering Industrial	6	1.68
13	Chemistry Analytical	5	1.40
14	Instruments Instrumentation	5	1.40
15	Transportation Science Technology	5	1.40
16	Computer Science Cybernetics	4	1.12
17	Operations Research Management Science	4	1.12
18	Physics Applied	4	1.12
19	Business	3	0.84
20	Construction Building Technology	3	0.84
21	Education Educational Research	3	0.84
22	Education Scientific Disciplines	3	0.84
23	Engineering Civil	3	0.84
24	Engineering Mechanical	3	0.84
25	Medical Informatics	3	0.84

As expected, the first eight ranked categories are related to IT in general, but it is worth mentioning that medical informatics is still not recognized as an important category and only 3 publications are assigned to that category.

2.2 Literature Search

The following types of literature written in English from 2010 to 2022 were included in the integrative review: articles published in conference materials and academic journals, books, guidances, standards, and other professional publications about medical device cybersecurity. The search range was from 2010 because at that time medical device manufacturers increasingly started considering adding or expanding connectivity options for medical devices. From a regulatory point of view, guidance that addresses the management of cybersecurity in medical devices throughout the premarket phase (FDA, 2014) was first published in October 2014. In May 2015, the FDA issued a first safety communication related to reported security vulnerabilities in Hospira's LifeCare PCA3 and PCA5 Infusion Pump Systems ("LifeCare PCA3 and PCA5 Infusion Pump Systems by Hospira: FDA Safety Communication - Security Vulnerabilities," 2015). The search of the literature in electronic databases was performed on April 18, 2022.

The EBSCO Discovery Service (EDS) search engine was used to search for specific terms in full text to identify potential articles: *"((threat modeling) OR (threat modelling) OR (threat model)) AND ((medical device) OR (medical devices) OR (medical IoT devices) OR (medical device software))"*. The search was conducted using the Boolean/Phrase search mode. In addition, the following expanders were applied: full text and equivalent subjects. The search returned 4 studies and 1 duplicate was excluded.

Additional articles were collected using the PubMed and IEEE Xplore search engines. The PubMed database was chosen because it supports the search and retrieval of biomedical and life sciences literature. IEEE Xplore provides full-text access to the technical literature in engineering and technology.

PubMed was searched using the following query: *"((threat modeling) OR (threat modelling) OR (threat model)) AND ((medical device) OR (medical devices) OR (medical IoT devices) OR (medical device software))"* with parameter [Title/Abstract]. The search returned 1 study.

IEEE Xplore advanced search was used to search for search terms *"threat modeling" OR "threat modelling" OR "threat model" AND "medical device"*

OR "medical devices" OR "medical IoT devices" OR "medical device software" in all metadata. The search resulted in a literature set consisting of 87 studies published in the IEEE Xplore digital library from 2010 to 2022.

Grey literature was found by conducting Google searches for documents published on the Internet from January 2010 to May 2022. Google search was performed on May 13th, 2022. The search strategy included the following queries: *threat modeling medical devices*, *threat modeling medical devices filetype:pdf*. Since it is impossible to screen all retrieved results from Google searches, we relied on relevancy ranking within the Google search engine to bring the most relevant results to the top of the list, and limited our search to the first 10 pages of results of each search query. The search results were reviewed using the title and short text underneath to identify the studies that are relevant to the research topic. The websites of organizations issuing standards for medical devices (i.e., IEC, AAMI) were also searched and 2 publications were found. In addition, 2 books about medical devices were screened and found to be eligible for inclusion in the integrative literature review.

2.3 Data Evaluation

The results of the database searches and grey literature searches were exported to an Excel spreadsheet and 2 duplicates were excluded through the 'Remove Duplicates' function. The abstracts and/or full text of all studies were manually reviewed to determine if a study is related to medical devices and threat modeling and if threat modeling methods/approaches are explicitly specified.

A total of 26 relevant studies were included in the present review, as shown in Fig. 1. A variable named "n" stands for the number of studies. The included studies are summarized in Table 3. Primary literature sources relevant to the research topic included 7 journal articles, 8 conference papers, 3 magazine articles, and 1 early access article. 7 out of 26 included studies represent grey literature including other types of non-journal literature such as books, playbooks, guidances, whitepapers, standards, and technical information reports.

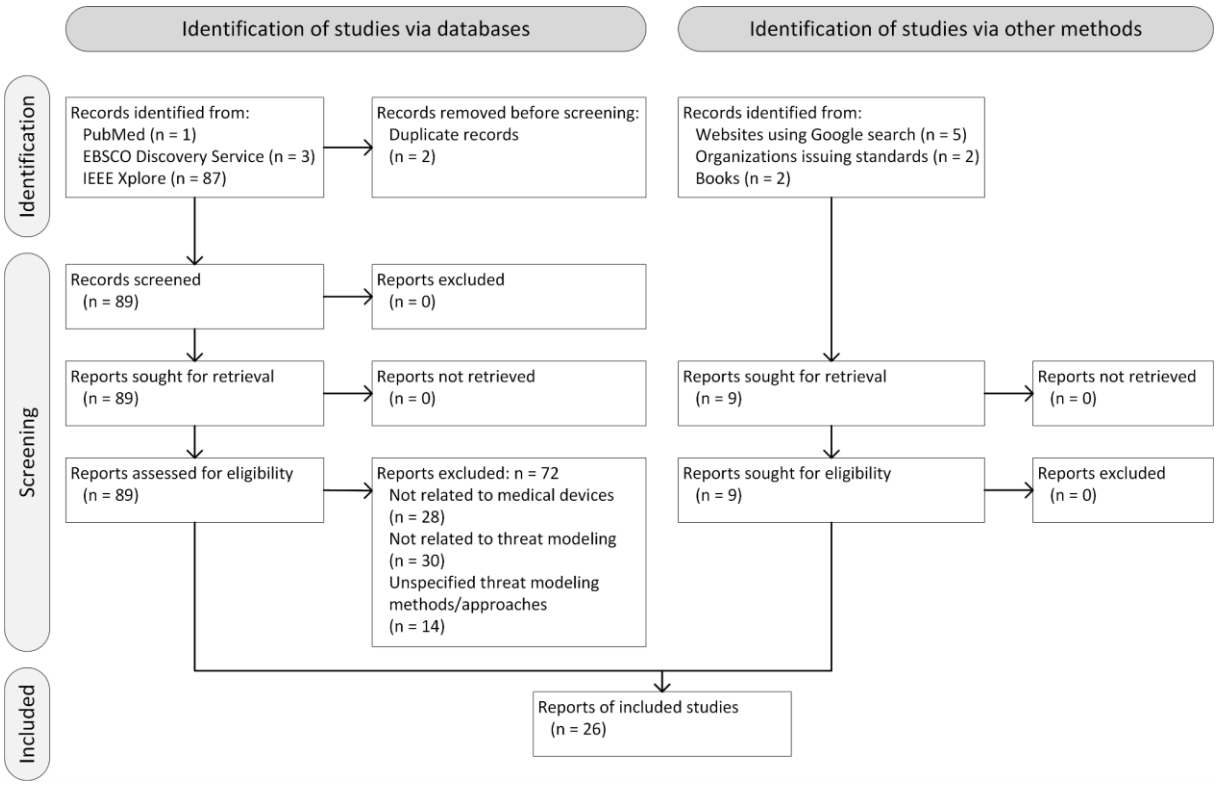


Figure 1. Flow diagram of studies screened through the review process

2.4 Data Analysis

The included studies were categorized according to the study type, device type, and threat modeling methods/approaches. It was interesting to see that some studies have focused on a particular medical device type such as an implantable pacemaker or similar, while the vast majority of studies have focused on a broader range of devices referred to as "medical devices" or a little bit more specific but still general

enough such as “interoperable medical devices” or “telecare medical information systems”. The comprehensive results displaying studies, their types, targeted devices, and identified threat modeling methods or approaches are presented in Table 3.

Table 3. Studies included in the integrative review

Source	Study	Study type	Device type	Threat modeling methods/approaches
PubMed	(Ibrahim et al., 2020)	Journal article	Implantable pacemaker	Attack graphs
EBSCO Discovery Service	(Malamas et al., 2021)	Journal article	Internet of Medical Things (IoMT)	Adversarial model Asset-based Attack paths Attack/defense tree HMG IS1 Logic decision diagram STRIDE STRIDE/DREAD
IEEE Xplore	(Xu et al., 2016)	Conference paper	Interoperable medical devices	Attack trees
	(Roy et al., 2018)	Journal article	E-healthcare systems	Dolev–Yao threat model
	(Kim et al., 2020)	Journal article	Medical devices	Attack trees
	(Manikandan & Sathyadevan, 2021)	Conference paper	Medical Implant Communication Systems (MICS) network	Attack trees
	(Almohri et al., 2017)	Conference paper	Medical cyber physical systems	Attack trees Attacker-centric model System-centric model
	(Lei & Chuang, 2019)	Journal article	Telecare medical information systems	Adversarial model
	(Shen et al., 2019)	Magazine article	Medical image retrieval for MIoT (Medical Internet of Things)	STRIDE derivation
	(Liu et al., 2020)	Journal article	Crowdsourcing IoT (Internet of Things)	Dolev-Yao threat model
	(Venkatasubramanian et al., 2012)	Magazine article	Interoperable medical devices	Attack classes
	(Vasserman et al., 2012)	Magazine article	Interoperable medical devices	Attack-consequences model with attack scenarios
	(Alsuwaidi et al., 2020)	Conference paper	Medical devices	Ways of attacks
	(Vakhter et al., 2022)	Early access article	Miniaturized wireless biomedical devices	Domain-specific qualitative-quantitative threat model
	(Cagnazzo et al., 2018)	Conference paper	Mobile health systems	DREAD STRIDE
	(Atamli & Martin, 2014)	Conference paper	IoT (smart healthcare system use case)	STRIDE derivation
	(Ould-Yahia et al., 2018)	Conference paper	e-Health in IoT-cloud environment	Adversary model
Google search	(The MITRE Corporation & MDIC, 2021)	Playbook	Medical devices	ATT&CK framework Attack trees Cyber Attack Lifecycle DREAD NIST SP 800-30 Appendices D-I Rubric for applying CVSS to Medical Devices STRIDE
	(Medcrypt, 2020)	Whitepaper	Medical devices	Attack trees CVSS PASTA Rubric for applying CVSS to Medical Devices

Source	Study	Study type	Device type	Threat modeling methods/approaches
				STRIDE
	(Seifert & Reza, 2016)	Journal article	Cyber-physical systems for healthcare	STRIDE DREAD
	(TGA, 2021)	Guidance	Medical devices	ATT&CK framework
	(Seale et al., 2018)	Conference paper	Networked medical devices	CVE (Common Vulnerabilities and Exposures) CVSS CWE (Common Weakness Enumeration) NVD (National Vulnerability Database) STRIDE
Organizations issuing standards	(*AAMI TIR57: Principles for medical device security–Risk management*, 2019)	Technical information report	Medical devices	Attack trees CVSS Security risk assessment approaches (threat-oriented, asset/impact-oriented, vulnerability-oriented)
	(*IEC 81001-5-1: Health software and health IT systems safety, effectiveness and security – Part 5-1: Security – Activities in the product life cycle*, 2021)	Standard	Health software, health IT systems	Attack-defense trees CAPEC dictionary of known patterns of attack CWE/SANS Top 25 Most Dangerous Software Errors CWSS (Common Weakness Scoring System) DREAD List known potential vulnerabilities OCTAVE OWASP Top 10 STRIDE Trike VAST
Book	(Ray, 2021)	Book	Medical devices	Attack tree CBOM (Cybersecurity Bill of Materials) CVSS Rubric for applying CVSS to Medical Devices STRIDE
	(Wirth et al., 2020)	Book	Medical devices	STRIDE

2.5 Presentation

The results of the integrative review are presented in Fig. 2. These results indicate that most methods have not stepped out of scientific laboratories into a wider professional application. Similar to other areas, in addition to the inherent quality of the method, it is crucial whether there are computer tools based on the method.

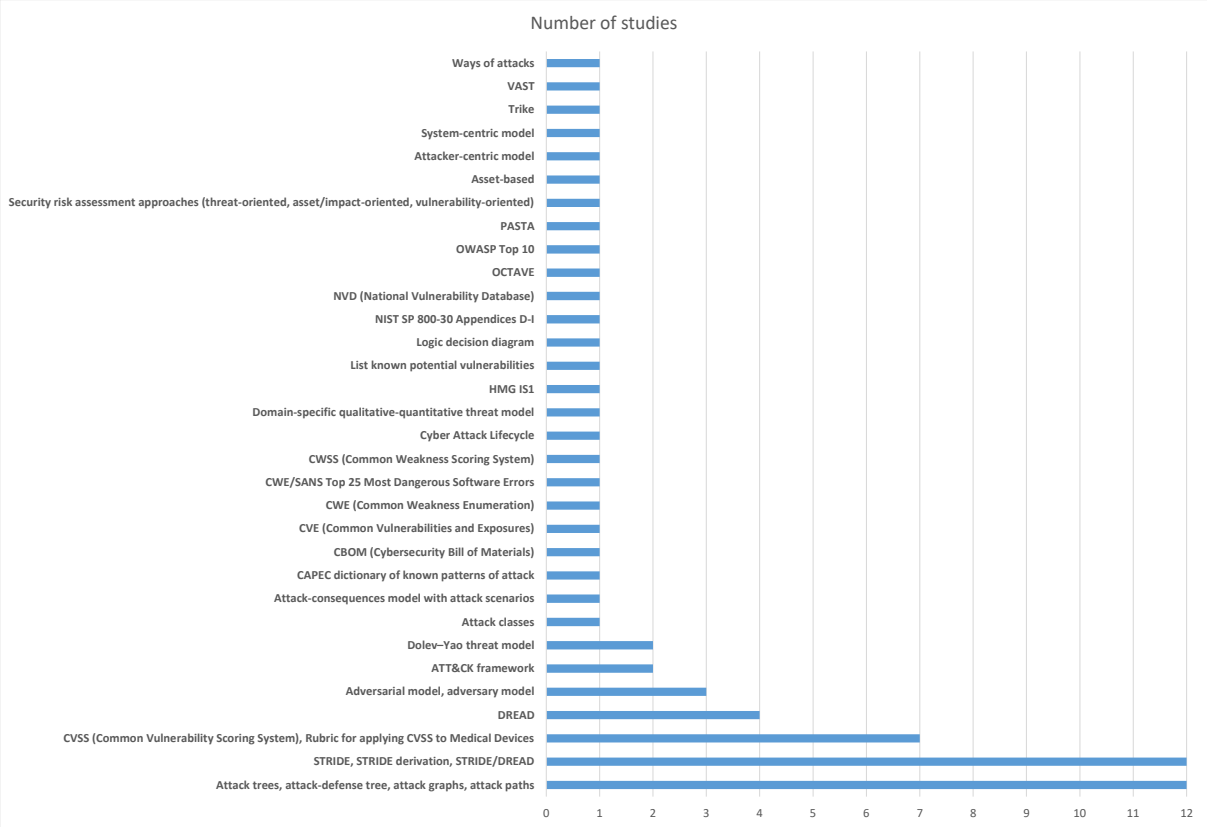


Figure 2. Number of studies per identified threat modeling method/approach

3 Discussion on Results

Our research identified a total of 32 threat modeling methods and approaches in the medical device industry. As illustrated in Fig. 2, the integrative literature review identified that 75% of threat modeling methods and approaches are not reported to be used outside the laboratory of the researchers presenting it. Although hurtful, the results also show the importance of this topic being actively researched over the course of past years.

Attack trees and the STRIDE threat taxonomy, including their derivations, appear to be the most used threat modeling methods. CVSS, Rubric for applying CVSS to Medical Devices and DREAD are the most used approaches for vulnerability scoring and to support the calculation of security risks. Threat catalogs based on NVD, CVE, CWE, ATT&CK framework and CAPEC dictionary of known patterns of attack can be used to derive further attack vectors to which a medical device may be subjected. Other approaches such as the OWASP Top 10, CWE/SANS TOP 25 Most Dangerous Software Errors, and NIST SP 800-30 Appendices D-I can be used to support threat modeling and security risk assessments.

Two studies (Roy et al., 2018), (Liu et al., 2020) included the Dolev-Yao threat model (Dolev & Yao, 1983) that can be used to analyze the security of public key protocols against saboteurs. The Dolev-Yao threat model was not mentioned in grey literature. It should be noted that grey literature provided examples of other threat modeling methods which were not found in the primary literature. Two studies (Almohri et al., 2017), (*AAMI TIR57: Principles for medical device security–Risk management*, 2019) also included approaches to threat modeling such as attacker-centric, asset-centric, and threat-centric. Five studies mentioned the DREAD (Damage potential, Reproducibility, Exploitability, Affected users, Discoverability) risk assessment technique. In the past, Microsoft used DREAD ratings to calculate security risk (Howard & Lipner, 2006). Three studies mentioned the technical paper (The MITRE Corporation, 2020) about a rubric for applying CVSS to medical devices.

3.1 Research Limitations

There may be some possible limitations in the present research. The first is a literature gap due to the lack of previous research on the topic. The second limitation concerns the manual review of studies to identify if a study is related to medical devices and threat modeling, and which threat modeling methods/approaches are specified. The third limitation is related to data collection. The research included studies presented in one language and there is no guarantee that all relevant grey literature was retrieved from the search.

4 Conclusions

Threat modeling is an integral part of a secure medical device product development life cycle. This paper presents the results of an integrative literature review of published studies on the different threat modeling methods in the medical device industry. Although further investigations are needed, the present research contributes to the existing literature on the topic of threat modeling methods and approaches in the medical device industry.

This paper may assist security professionals, threat modeling experts, security risk managers, software/firmware developers, subject matter experts, and other stakeholders who participate in threat modeling activities of medical devices to choose a specific threat modeling method and embrace threat modeling throughout the medical device life cycle.

The studies were collected using the following search engines: PubMed, EDS, and IEEE Xplore. Future research could conduct literature searches in other academic research databases such as Web of Science, Scopus, SpringerLink, SAGE journals, or ScienceDirect.

Further investigation should explore the advantages and disadvantages of identified threat modeling methods and approaches and which threat modeling tools are recommended for medical device threat modeling. More research is needed to explore if threat modeling is carried out not just for new medical devices, but also on legacy devices. Future research should also investigate if the TARA (Threat Agent Risk Assessment) methodology (Rosenquist, 2009) can be adapted for the medical device industry.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions that significantly improved this paper.

References

AAMI TIR57: Principles for medical device security–Risk management. (2019).

Almohri, H., Cheng, L., Yao, D., & Alemzadeh, H. (2017). On Threat Modeling and Mitigation of Medical Cyber-Physical Systems. *Proceedings of the 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, Philadelphia, (pp. 114–119). IEEE. doi:10.1109/CHASE.2017.69

Alsuwaidi, A., Hassan, A., Alkhatri, F., Ali, H., Qbea'H, M., & Alrabaee, S. (2020). Security Vulnerabilities Detected in Medical Devices. *Proceedings of the 2020 12th Annual Undergraduate Research Conference on Applied Computing (URC)*, Dubai, (pp. 1–6). IEEE. doi:10.1109/URC49805.2020.9099192

Atamli, A. W., & Martin, A. (2014). Threat-Based Security Analysis for the Internet of Things. *Proceedings of the 2014 International Workshop on Secure Internet of Things*, Wroclaw, (pp. 35–43). IEEE. doi:10.1109/SIoT.2014.10

Cagnazzo, M., Hertlein, M., Holz, T., & Pohlmann, N. (2018). Threat modeling for mobile health systems. *Proceedings of the 2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, Barcelona, (pp. 314–319). IEEE. doi:10.1109/WCNCW.2018.8369033

Dolev, D., & Yao, A. C. (1983). On the Security of Public Key Protocols. *IEEE Transactions on Information Theory*, 29(2), 198–208. doi:10.1109/TIT.1983.1056650

FDA. (2014). Content of Premarket Submissions for Management of Cybersecurity in Medical Devices – Guidance for Industry and Food and Drug Administration Staff. Retrieved from https://www.fda.gov/media/86174/download

FDA. (2016). Postmarket Management of Cybersecurity in Medical Devices – Guidance for Industry and Food and Drug Administration Staff. Retrieved from https://www.fda.gov/media/95862/download

FDA. (2018). Cybersecurity in Medical Devices: Quality System Considerations and Content of Premarket Submissions – Draft Guidance for Industry and Food and Drug Administration Staff. Retrieved from https://www.fda.gov/media/119933/download

FDA. (2022). Cybersecurity in Medical Devices: Quality System Considerations and Content of Premarket Submissions – Draft Guidance for Industry and Food and Drug Administration Staff. Retrieved from https://www.fda.gov/media/119933/download

Howard, M., & Lipner, S. (2006). *The Security Development Lifecycle. SDL: A Process for Developing Demonstrably More Secure Software*. Redmond: Microsoft Press.

HSCC. (2019). Medical Device and Health IT Joint Security Plan. Retrieved from https://healthsectorcouncil.org/wp-content/uploads/2019/02/HSCC-MEDTECH-JSP-v1.2.pdf

Ibrahim, M., Alsheikh, A., & Matar, A. (2020). Attack Graph Modeling for Implantable Pacemaker. *Biosensors*, 10(2), 14. doi:10.3390/bios10020014

IEC 81001-5-1: Health software and health IT systems safety, effectiveness and security – Part 5-1: Security – Activities in the product life cycle. (2021).

IMDRF. (2020). Principles and Practices for Medical Device Cybersecurity. Retrieved from https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-200318-pp-mdc-n60.pdf

ISO 14971: Medical devices – Application of risk management to medical devices. (2019).

IT Security Guideline for Medical Devices. (2021). Retrieved from https://github.com/johner-institut/it-security-guideline/blob/master/Guideline-IT-Security_EN.md

Kim, D., Choi, J., & Han, K. (2020). Medical Device Safety Management Using Cybersecurity Risk Analysis. *IEEE Access*, 8, 115370–115382. doi:10.1109/ACCESS.2020.3003032

Kohnfelder, L. (2022). *Designing Secure Software: A Guide for Developers*. San Francisco: No Starch Press.

Kohnfelder, L., & Garg, P. (1999). The threats to our products. Retrieved from https://adam.shostack.org/microsoft/The-Threats-To-Our-Products.docx

Lei, C.-L., & Chuang, Y.-H. (2019). Privacy Protection for Telecare Medicine Information Systems with Multiple Servers Using a Biometric-based Authenticated Key Agreement Scheme. *IEEE Access*, 7, 186480–186490. doi:10.1109/ACCESS.2019.2958830

LifeCare PCA3 and PCA5 Infusion Pump Systems by Hospira: FDA Safety Communication - Security Vulnerabilities. (2015). Retrieved from https://wayback.archive-it.org/7993/20170112164109/http:/www.fda.gov/Safety/MedWatch/SafetyInformation/SafetyAlertsforHumanMedicalProducts/ucm446828.htm

Liu, W., Wang, X., & Peng, W. (2020). Secure Remote Multi-Factor Authentication Scheme Based on Chaotic Map Zero-Knowledge Proof for Crowdsourcing Internet of Things. *IEEE Access*, 8, 8754–8767. doi:10.1109/ACCESS.2019.2962912

Malamas, V., Chantzis, F., Dasaklis, T. K., Stergiopoulos, G., Kotzanikolaou, P., & Douligeris, C. (2021). Risk Assessment Methodologies for the Internet of Medical Things: A Survey and Comparative Appraisal. *IEEE Access*, 9, 40049–40075. doi:10.1109/ACCESS.2021.3064682

Manikandan, R., & Sathyadevan, S. (2021). Medical Implant Communication Systems (MICS) Threat Modelling. *Proceedings of the 2021 2nd International Conference on Secure Cyber Computing and Communications (ICSCCC)*, Jalandhar, (pp. 518–523). IEEE. doi:10.1109/ICSCCC51823.2021.9478155

MDCG. (2020). MDCG 2019-16 Rev.1 Guidance on Cybersecurity for medical devices. Retrieved from https://ec.europa.eu/health/document/download/b23b362f-8a56-434c-922a-5b3ca4d0a7a1_en?filename=md_cybersecurity_en.pdf

Medcrypt. (2020). *Medical device threat modeling* [White paper]. Retrieved from https://www.medcrypt.co/whitepaper_resources/WP9_Medical%20Device%20Threat%20Modeling_v4.1.pdf

Microsoft Threat Modeling Tool. (2020). Retrieved from https://docs.microsoft.com/en-us/azure/security/develop/threat-modeling-tool

NMPA. (2022). Guidelines for Registration Review of Medical Device Cybersecurity (Revised Edition 2022).

Ould-Yahia, Y., Bouzefrane, S., & Boucheneb, H. (2018). Towards privacy and ownership preserving of outsourced health data in IoT-cloud context. *Proceedings of the 2018 International Symposium on Programming and Systems (ISPS)*, Algiers, (pp. 1–6), doi:10.1109/ISPS.2018.8379018

Ray, A. (2021). *Cybersecurity for Connected Medical Devices*. Eastbourne: Academic Press.

Rosenquist M. (2009). *Prioritizing Information Security Risks with Threat Agent Risk Assessment* [White paper]. Retrieved from https://media10.connectedsocialmedia.com/intel/10/5725/Intel_IT_Business_Value_Prioritizing_Info_Security_Risks_with_TARA.pdf

Roy, S., Chatterjee, S., Das, A. K., Chattopadhyay, S., Kumari, S., & Jo, M. (2018). Chaotic Map-Based Anonymous User Authentication Scheme With User Biometrics and Fuzzy Extractor for Crowdsourcing Internet of Things. *IEEE Internet of Things Journal*, 5(4), 2884–2895. doi:10.1109/JIOT.2017.2714179

Seale, K. A., McDonald, J. T., & Glisson, W. B. (2018). MedDevRisk: Risk Analysis Methodology for Networked Medical Devices. *Proceedings of the 51st Hawaii International Conference on System Sciences (HICSS)*, Hawaii, (pp. 3271–3280). doi:10.24251/HICSS.2018.414

Seifert, D., & Reza, H. (2016). A Security Analysis of Cyber-Physical Systems Architecture for Healthcare. *Computers*, 5(4), 27. doi:10.3390/computers5040027

SFDA. (2019). Guidance to Post-Market Cybersecurity of Medical Devices. Retrieved from https://beta.sfda.gov.sa/sites/default/files/2021-01/MDS-G37.pdf

Shen, M., Deng, Y., Zhu, L., Du, X., & Guizani, N. (2019). Privacy-Preserving Image Retrieval for Medical IoT Systems: A Blockchain-Based Approach. *IEEE Network*, 33(5), 27–33. doi:10.1109/MNET.001.1800503

Shevchenko, N., Chick, T., A., O'Riordan, P., Scanlon, T. P., & Woody, C. (2018). Threat Modeling: A Summary of Available Methods. Retrieved from https://resources.sei.cmu.edu/asset_files/WhitePaper/2018_019_001_524597.pdf

Shostack, A. (2014) *Threat Modeling: Designing for Security*. Indianapolis: John Wiley & Sons.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. doi:10.1016/j.jbusres.2019.07.039

Tarandach, I., & Coles, M. J. (2020). *Threat Modeling: A Practical Guide for Development Teams*. Sebastopol: O'Reilly Media.

TGA. (2021). Medical device cyber security guidance for industry. Retrieved from https://www.tga.gov.au/sites/default/files/medical-device-cyber-security-guidance-industry.pdf

The MITRE Corporation & MDIC. (2021). Playbook for Threat Modeling Medical Devices. Retrieved from https://www.mitre.org/sites/default/files/publications/Playbook-for-Threat-Modeling-Medical-Devices.pdf

The MITRE Corporation. (2020). Rubric for applying CVSS to Medical Devices. Retrieved from https://www.mitre.org/sites/default/files/publications/pr-18-2208-rubric-for-applying-cvss-to-medical-devices.pdf

Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4(3), 356–367. doi:10.1177/1534484305278283

Vakhter, V., Soysal, B., Schaumont, P., & Guler, U. (2021). Security for Emerging Miniaturized Wireless Biomedical Devices: Threat Modeling with Application to Case Studies. *IEEE Internet of Things Journal*. Advanced online publication. doi:10.1109/JIOT.2022.3144130

Vakhter, V., Soysal, B., Schaumont, P., & Guler, U. (2022). Threat Modeling and Risk Analysis for Miniaturized Wireless Biomedical Devices. *IEEE Internet of Things Journal*. Advanced online publication. doi:10.1109/JIOT.2022.3144130

Vasserman, E. Y., Venkatasubramanian, K. K., Sokolsky, O., & Lee, I. (2012). Security and Interoperable-Medical-Device Systems, Part 2: Failures, Consequences, and Classification. *IEEE Security & Privacy*, 10(6), 70–73. doi:10.1109/MSP.2012.153

Venkatasubramanian, K. K., Vasserman, E. Y., Sokolsky, O., & Lee, I. (2012). Security and Interoperable-Medical-Device Systems, Part 1. *IEEE Security & Privacy*, 10(5), 61–63. doi:10.1109/MSP.2012.128

Whittemore, R., & Knafl, K. (2005). The integrative review: updated methodology. *Journal of Advanced Nursing*, 52(5), 546–553. doi:10.1111/j.1365-2648.2005.03621.x

Wirth, A., Gates, C., & Smith, J. (2020). *Medical Device Cybersecurity for Engineers and Manufacturers*. Norwood: Artech House.

Xu, J., Venkatasubramanian, K. K., & Sfyrla, V. (2016). A methodology for systematic attack trees generation for interoperable medical devices. *Proceedings of the 2016 Annual IEEE Systems Conference (SysCon)*, Orlando, (pp. 1–7). IEEE. doi:10.1109/SYSCON.2016.7490632

An Overview of Modern Cross-platform Mobile Development Frameworks

Jelica Stanojević, Uroš Šošević, Miroslav Minović, Miloš Milovanović
Faculty of Organizational Sciences
University of Belgrade
Jove Ilića 154, 11010 Belgrade, Serbia
{jelica.stanojevic, uros.sosevic, miroslav.minovic, milos.milovanovic}@fon.bg.ac.rs

Abstract. *Interest and usage of Cross-platform mobile development frameworks (CMDF) is growing over the years. This paper outlines trends in the field of mobile development, specifically in CMDFs area which is rapidly evolving with new emerging approaches for creating mobile applications. The most popular cross-platform mobile frameworks today are Flutter, React Native, Cordova, Ionic and Xamarin, respectively. Mentioned frameworks cover both predominant mobile operating systems (Android and iOS) and also provide support for desktop and web applications. The aim of this paper is to shed light on currently (2022) most popular CMDFs in order for academia to follow the rapid industry development in this field.*

Keywords. Mobile development, Cross-platform, Flutter, React Native, Cordova, Ionic, Xamarin

1 Introduction

In dynamic environments where the proper timing of application market entry is of great importance building applications with a Write Once, Run Anywhere mobile application tools (Shah, Sinha, & Mishra, 2019) could be beneficial in contrast to building native apps targeting each platform separately which will increase time, effort and cost in order to satisfy predominant operating systems. This is especially important in the case of early stage startups where the efforts are focused towards building the Most Viable Product in a short time span. In such situations where time, effort, cost and cross-platform compatibility are important, developers tend to choose CMDF approach rather than native approach for developing applications (Shah et al., 2019). On the other hand, developers need to understand if and which CMDF could meet their needs.

An analysis from 2015 shows that 3.73% out of about 12000 mobile applications available on Google Play Store, were created using CMDF hybrid approach (Malavolta, Ruberto, Soru, & Terragni, 2015). A study from 2022, which was done on a dataset consisting of about 650 thousand mobile applications, shows a higher value of 8,67%. This could indicate a growth in interest in the hybrid development approach (Biørn-Hansen, Grønli, Majchrzak, Kaindl, & Ghinea, 2022). Regardless of the approach, all mobile applications made with CMDFs occupy 15% of the dataset which was analyzed by (Biørn-Hansen et al., 2022).

According to the worldwide survey conducted by JetBrains, and which included over 32000 software developers, predominant cross-platform mobile frameworks in 2021 were Flutter[1], React Native[2], Cordova[3], Ionic[4] and Xamarin[5] (JetBrains, 2022). These results are in line with Google Trends search queries for these technologies over the past few years, specifically 2020-2022 ("Google Trends Comparison (React Native vs. Flutter vs. Ionic vs. Xamarin vs. Cordova)," 2022).

Android and iOS represent the two most dominant mobile phone operating systems with a total market share of about 70% and 25% respectively (Laricchia, 2022). Mentioned frameworks cover both predominant mobile operating systems and also give a solution for desktop and web applications.

There aren't many research papers that include and compare the majority of the top 5 most popular CMDFs nowadays which are previously mentioned in this paper. Deciding from various options of CMDF with different underlying paradigms that offer different functionalities and performance based on use cases isn't always straight-forward. That is why it is important to conduct their in-depth analysis and comparison in order to include more important parameters in making this decision.

2 Methodology

The aim of this paper is to shed light on currently (2022) most popular CMDFs in order for academia to follow the rapid industry development in this field. Explanation of

[1] https://flutter.dev/
[2] https://reactnative.dev/
[3] https://cordova.apache.org/
[4] https://ionicframework.com/
[5] https://dotnet.microsoft.com/en-us/apps/xamarin

different types of applications that CMDFs produce will be given and mostly used frameworks will be categorized on this basis. Different possible taxonomies will be mentioned. Main characteristics of predominant CMDFs will be described. Lastly, discussion regarding already mentioned CMDFs based on existing research in this area will be realized and conclusions will be made. Google Scholar was the main database used in research of existing work on this topic. Papers included in the Related Work section are relatively new Journals, Magazines and Conference Proceedings published during 2018-2022 with a focus on review and comparison of modern CMDFs.

3 Cross-platform Mobile Development Frameworks

Beside native approach in which Android apps are created using Java or Kotlin and iOS apps using Swift or Objective-C programming language, one of the approaches for developing mobile applications is using cross-platform mobile development frameworks. Types of applications produced by CMDF could be grouped into following categories (Xanthopoulos & Xinogalos, 2013):

- Web Apps,
- Hybrid Apps,
- Interpreted Apps and
- Generated Apps.

The new taxonomy presented by Nunkesser in 2018, expands previously noted groups, categorizing mobile applications as:

- Pandemic, which are subdivided as
    - Web App
    - Hybrid Web App - Cordova and Ionic
    - Hybrid Bridged App - React Native and Flutter
    - System Language App,
- Endemic (native apps) and
- Ecdemic/Foreign Language App (like Xamarin), which can be
    - Interpreted App,
    - Generated App,
    - VM App.

Hybrid Web Apps is a match for Hybrid Apps in (Xanthopoulos & Xinogalos, 2013) categorization and Ionic and Cordova belong to this category without any disagreement in academic papers or industry. That isn't the case for the rest of the top 5 CMDFs. This is mainly because the term hybrid, in industry and academia, is used for approaches that utilize web technology but without building a strictly defined hybrid app (Martinez, 2019; Singh & Shobha, 2021; Vishal & Kushwaha, 2018).

In 2018, two predominant approaches were Hybrid and Interpreted (Biørn-Hansen & Ghinea, 2018), which can now be appended with Cross-compiled, Widget-based approach like Flutter.

M-site is another approach for creating mobile apps which results with web applications, applications that are created using widespread web technologies (e.g. HTML, CSS and JavaScript) and users access them through a web browser (Zohud & Zein, 2021). Since the main goal of this paper is to research CMDF that produce mobile applications which can be downloaded and installed on mobile devices from Google Play Store or App Store, and interact with native devices APIs, this type of apps isn't going to be in the focus of this paper.

CMDF which produces hybrid types of applications, leverages the common feature of different mobile operating systems, the mobile browser accessible from the native code (Charland & Leroux, 2011). Hybrid applications are built based on web technologies, HTML, CSS and JavaScript, which are browser-supported languages, and are wrapped as native apps inside a special native container (Que, Guo, & Zhu, 2016; Zohud & Zein, 2021) called WebView[6] in Android OS and WKWebView[7] in iOS. The name hybrid comes from a combination of native and web applications features (Que et al., 2016).

The WebView, which deals with user interface and gives better access to device capabilities, and the Plugins, which are intended to deal with native device features (e.g. storage, camera etc.) could be considered as main characteristics of hybrid apps CMDFs (Shah et al., 2019; Zohud & Zein, 2021). Plugins bind device APIs to the application, which makes it easy to be invoked simply by using JavaScript that provides developers with a generic API for each feature to bridge all the service requests from the web-based code to the corresponding platform API, abstracting away device specific APIs (Pinto & Coutinho, 2018; Shah et al., 2019)**.** Examples of Hybrid CMDF that are currently popular (2022) are Cordova and Ionic.

Logic in interpreted applications is implemented in a platform-independent way, while their native code is automatically generated to implement the user interface with platform-specific native UI elements (Nunkesser, 2018; Xanthopoulos & Xinogalos, 2013). The native features are provided by an abstract layer that interprets the code on runtime across different platforms to access the native APIs (Latif, Lakhrissi, Nfaoui, & Es-Sbai, 2017). These types of apps provide the look and feel of native applications because of native user interfaces (S.Thakare, Shirodkar, Parween, & Parween, 2014). On the downside, new platform-specific features, like new user interface features of a new android version, can be available to apps only when supported by the development environment so there is the complete dependence on the software development environment (S.Thakare et al., 2014). React Native is nowadays (2022) an actual CMDF that produces interpreted apps (Biørn-Hansen & Ghinea, 2018; Huber, Demetz, & Felderer, 2020; Shah et al., 2019).

[6] https://developer.android.com/reference/android/webkit/WebView

[7] https://developer.apple.com/documentation/webkit/wkwebview

As for generated applications, which are also called cross-compiled (S.Thakare et al., 2014), their code, also written in common programming languages, is compiled to a specific native code for each targeted platform (Latif et al., 2017; Xanthopoulos & Xinogalos, 2013). Because of the generated native code these types of apps achieve high overall performance (S.Thakare et al., 2014). Xamarin is currently (2022) one of the most used CMDFs which create this type of apps (Latif, Lakhrissi, Nfaoui, & Es-Sbai, 2016). Flutter is a cross-compiled framework (Biørn-Hansen, Rieger, Grønli, Majchrzak, & Ghinea, 2020) but also belongs to a special category of Widget based apps (Shah et al., 2019).

In next chapters a brief overview of each mentioned which represents also the most popular CMDFs is going to be given.

3.1 Cordova

Apache Cordova is an open-source version of PhoneGap which was acquired by Adobe in 2011. Cordova is a hybrid CMDF so its main characteristics are a container, WebView, for running the code (HTML/CSS/JavaScript) inside a native wrapper which is device specific and Plugins which are dealing with the native device features. Plugins bind device APIs to the application, which makes it easy to be invoked simply by using JavaScript. (Shah et al., 2019)

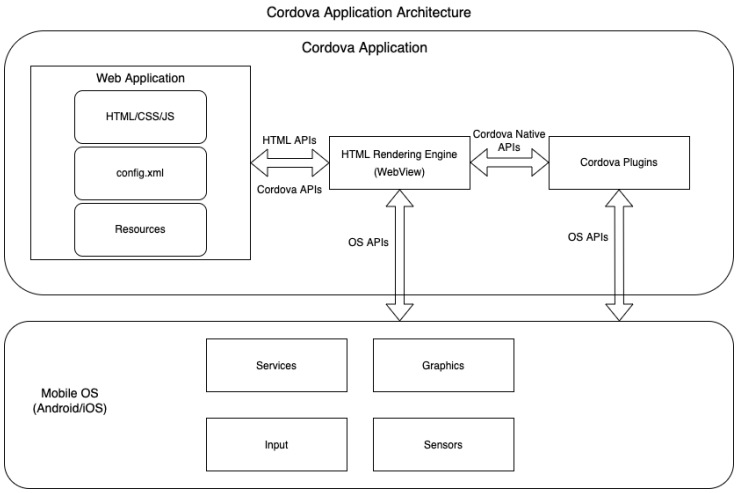


Figure 1. Cordova Application Architecture (The Apache Software Foundation, n.d.)

3.2 Ionic

As previously mentioned, Ionic represents an example of Hybrid App CMDF. First version of Ionic Framework was released in 2013 and was built on top of AngularJS[8], which is now deprecated (from January 2022), and Apache Cordova. In 2016 release of version 2 of Angular framework with drastic changes also caused version 2 of Ionic framework. Back then, Ionic components, which eased development of mobile UI, were built as Angular directives. That meant that Ionic framework
is coupled with Angular and with the rise of other Single Page Application (SPA) frameworks like React[9] and Vue[10] that became a problem. With adoption of Web Components standard which represents a suite of different technologies allowing you to create reusable custom elements - with their functionality encapsulated away from the rest of your code - and utilize them in your web apps (MDN Contributors, 2022) Ionic released version 4 in 2019 with components based on this standard. That way, Ionic was decoupled from the Angular framework, and Ionic components can be used with other SPA frameworks. Ionic team even created a tool, named Stencil[11], for creating components. Ionic 6 is the currently active version (from December 2021).

Beside components, another important part of Hybrid CMDF like Ionic, are tools like Cordova and Capacitor that help bridge the gap between web apps and native device features. The Ionic team developed Capacitor in 2018, and while it has similarities with Cordova, it is using new modern APIs which were not available when Cordova was released (Lynch, n.d.) and is considered as Ionic's official native runtime (Ionic team, n.d.).

[8] https://angularjs.org/
[9] https://reactjs.org/
[10] https://vuejs.org/
[11] https://stenciljs.com/

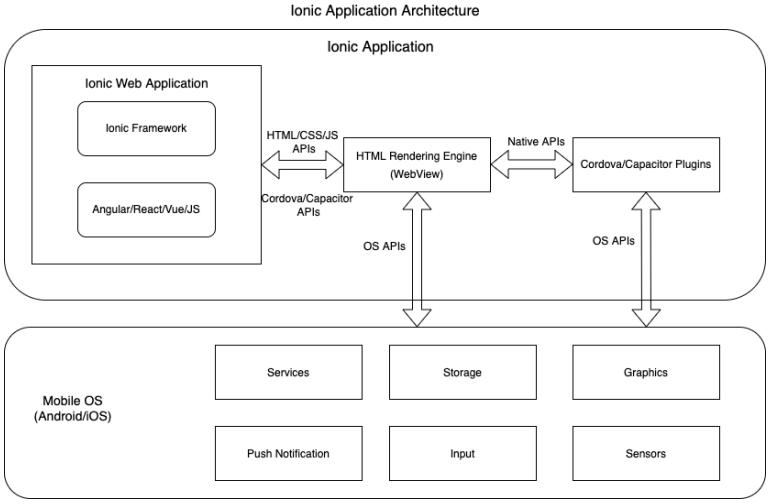


Figure 2. Ionic Application Architecture (Singh & Shobha, 2021)

3.3 React Native

React Native is nowadays the most popular Interpreted Apps CMDF. React Native started at an internal Meta Hackathon and was released in March 2015 by Meta as an open-source framework. Same as Ionic, React Native will feel familiar to web developers because they can create mobile apps using JavaScript (Pinto & Coutinho, 2018) so their learning curve is eased.

React Native is using a JavaScript engine, JavaScriptCore, in order to interpret (hence the categorization) deployed application source code directly to the mobile device (Shah et al., 2019). At runtime, React Native creates the corresponding Android and iOS views for React Native components. Because React Native components are backed by the same views as Android and iOS, React Native apps look, feel, and perform like any other apps. (Meta Platforms, 2022)

The bridge layer allows to invoke platform specific APIs, Swift or Objective-C APIs for iOS and Java or Cotlin APIs for Android, needed for rendering components of the native view controller (Shah et al., 2019). Also, if an interpreted app requires access to a native feature (e.g. camera), it will pass the request down through the respective bridge handling such requests, and return a result up through the same bridge where needed, e.g. to return a value from the native-side to the app-side (Biørn-Hansen & Ghinea, 2018).

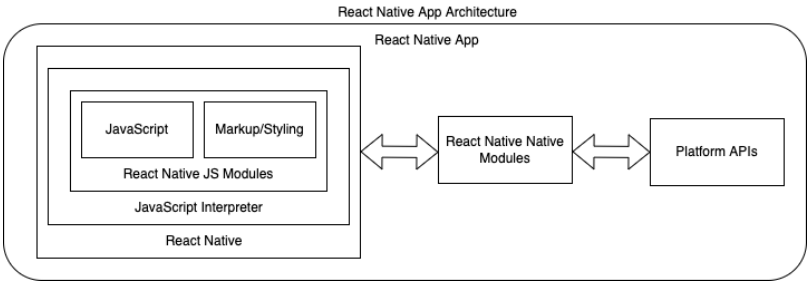


Figure 3. React Native Application Architecture (Biørn-Hansen & Ghinea, 2018)

First announced in 2018, recently, in 2022, React Native rolled out new architecture with the new native TurboModule system for interaction with the native side and the Fabric Renderer, new rendering system, meaning the bridge is splitted in two parts (Sciandra, 2019).

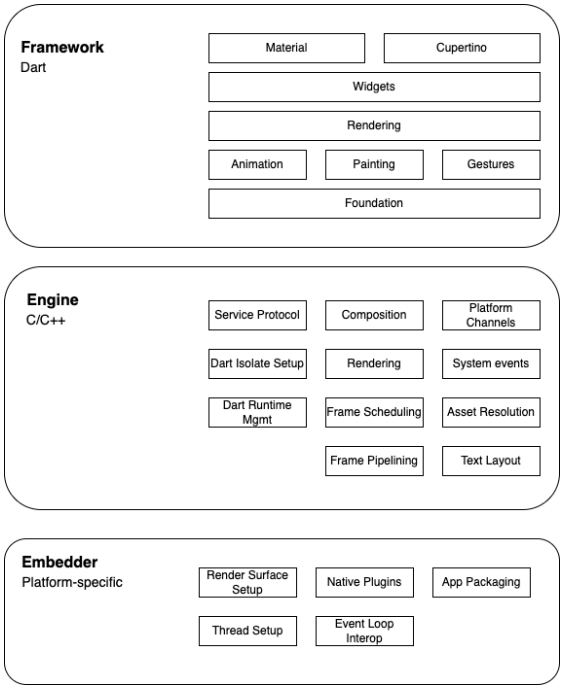


Figure 4. Flutter architectural overview (Flutter, n.d.)

3.4 Flutter

Flutter became the most popular CMDF lately. It was created by Google and the first version was released in 2017. Flutter applications are made using Dart programming language[12] which is similar to JavaScript.

Flutter framework includes a rich set of platform, layout, and foundational libraries for animation, painting, gestures and design like Material and Cupertino libraries which offer comprehensive sets of controls to implement the Material or iOS design languages. (Flutter, n.d.; Shah et al., 2019) It uses Dart for creating components using the Skia 2D graphics engine which is used for rendering the application UI (Shah et al., 2019). A platform-specific embedder, written in appropriate platform language, provides an entrypoint; coordinates with the underlying operating system for access to services like rendering surfaces, accessibility, and input; and manages the message event loop. Using the embedder, Flutter code can be integrated into an existing application as a module, or the code may be the entire content of the application. (Flutter, n.d.)

3.5 Xamarin

Example of Generated apps CMDF is Xamarin, developed by Microsoft since 2016, which uses C# in order to create applications for various devices.

Xamarin.Forms, which was introduced in Xamarin 3 in 2014, represents an open-source UI framework that allows developers to create user interfaces in XAML with code-behind in C# which will be rendered as native controls on each platform (Johnson, Britch, & Dunn, 2021). The process of accessing native utilities is simplified by abstraction, a library Xamarin.Essentials, that provides cross-platform APIs for native device features (Justin Johnson, David Britch, Craig Dunn, & Hemant Arya, 2021)**.**

Microsoft's evolution of Xamarin.Forms, .NET Multi-platform App UI[13] (.NET MAUI), is a framework for building modern, multi-platform, natively compiled iOS, Android, macOS, and Windows apps using C# and XAML in a single codebase. Conversion from Xamarin.Forms to .NET MAUI is basically complete in 2022. Xamarin support will continue through November 2023 (Ramel, 2022).

[12] https://dart.dev/

[13] https://dotnet.microsoft.com/en-us/apps/maui

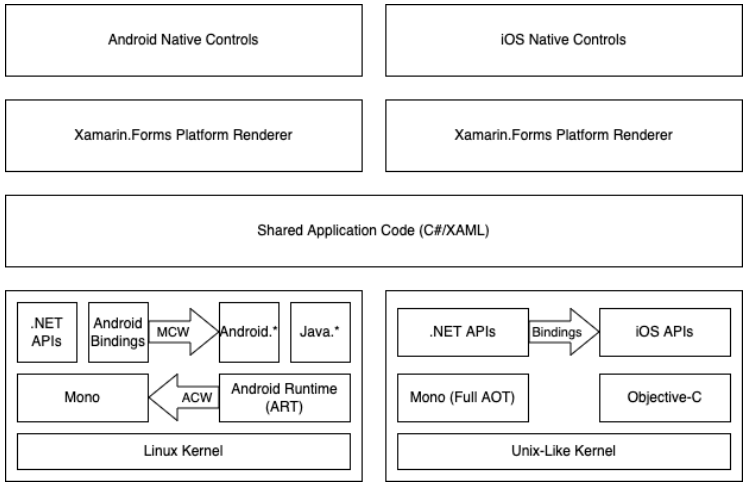


Figure 5. Xamarin.Forms Architecture (Johnson et al., 2021)

4 Related work

As for cross-platform approaches in general it can be concluded that developers need to compromise on the user experience which is of highest quality on native devices (Shah et al., 2019), but it's important to notice that it is improving in the newer releases of CMDF and that performance of mobile devices is constantly improving. For example, based on a study which was carried out by Huber, Demetz, & Felderer (2020) React Native and Android Native apps in terms of UI fluidness are comparable which results in better user experience in React Native cross-platform solution, but also comes with a consequence of high load on CPU and main memory. On the other hand Ionic has insignificant better performance compared to React Native concerning mentioned aspects, but missing the benefits of a low janky frame rate and low GPU memory usage which results in performance degradation perceived by the user (Huber et al., 2020).

In a research paper by Biørn-Hansen, Grønli, & Ghinea (2019) device hardware impact and penalties caused by transitions and animations (in-app page navigation, opening side menu and complex animation - Lottie animation[14]) were studied and reports are made on following metrics: frames per second (FPS), CPU usage, device memory usage and GPU memory usage. Eight different mobile apps were created: native iOS and Android, and also Ionic, Xamarin and React Native for iOS and Android. Even though the performance of running complex animation in Native apps is much better in
contrast to Ionic Hybrid App, Biørn-Hansen, Grønli, & Ghinea (2019), from users' perspective, couldn't observe the difference between these apps. This indicates that users' experience should be taken into account while measuring performances for certain parameters (in this case FPS).

Mobile apps created using CMDFs, specifically Ionic/Capacitor and React Native, compared to natively developed applications put a higher load on CPU usage, main memory and GPU memory consumption (Huber et al., 2020). Except in the use case of building high-end gaming applications, CMDFs usually prevail over the disadvantages that come with their usage (Shah et al., 2019).

When it comes to implementation and development of communication bridges for using native device features in Hybrid (Ionic-Cordova) and Interpreted (React Native) approaches Biørn-Hansen & Ghinea (2018) conclude that these aren't complex nor tedious programming tasks in both cases. However, the Hybrid-based Ionic-Cordova app proved to be five times faster than the Interpreted app built with React Native for the functionality of fetching images stored on native devices (Biørn-Hansen & Ghinea, 2018).

On the other hand, in comparative study (Quazi & Sinha, 2018) of Ionic, React Native and PhoneGap, React Native proved to be the finest CMDF based on execution time to perform the hardware and database related operations. React Native is also the most popular from the industrial point of view based on multiple case-study (Zohud & Zein, 2021) carried out in four different software development companies in Palestine in 2021, even though Ionic is more used than React Native since its older.

Shah et al. (2019) conclude that Flutter is overall the best choice among all CMDF which combines the advantages of development tools from other approaches.

Differences regarding performances for CMDF on different platforms, iOS and Android, exist. Cordova apps exhibit significantly different patterns of rendering time, which can be attributed to the differences in the underlying HTML rendering engines and the JavaScript engines of the respective platforms, and UI response on the iOS and Android

[14] https://lottiefiles.com/72-favourite-app-icon

platform. (Jia, Ebone, & Tan, 2018) This is also in line with study of Biørn-Hansen, Grønli, & Ghinea (2019) in which Android and iOS are found to differ greatly in terms of memory consumption, CPU usage and rendered FPS for both the native and cross-platform apps.

There are some papers regarding industry perspective on CMDFs and their usage, but this area represents the potential for further exploration due to the rapid pace of development in the field, in order for academia to stay relevant with industry standards in mobile development (Biørn-Hansen, Grønli, Ghinea, & Alouneh, 2019). Based on a survey conducted in 2016 (Biørn-Hansen, Grønli, Ghinea, et al., 2019) which was targeting industry practitioners, the most popular frameworks were Phone Gap (which is now deprecated), Ionic and React Native, both in hobby and professional projects. Biørn-Hansen et al. (2019) concluded that industry is moving towards using React Native and Ionic which is in line with today's statistics (JetBrains, 2022). Obtained conclusion in another, already mentioned study (Zohud & Zein, 2021), carried out over industry practitioners, is that developers experience is the most influential factor in the development process. This could indicate that web developers who are familiar with SPA framework like Angular would prefer to work with Ionic, rather than learn new CMDF from scratch like Flutter. This is also the case with React Native since it uses JavaScript, and on the other hand for Flutter, Dart, its own development programming language must be learned (Işitan & Koklu, 2020).

Conclusion of Biørn-Hansen et al. (2020) is that certain cross-platform frameworks can perform equally well or even better than native on certain metrics. In this case metrics were measured for tasks of creating contact, reading from file system, accessing location information and usage of accelerometer for Ionic, React Native, NativeScript, Flutter, MAML/MD_2 and Native Android. No single framework scores best across all features in the study, nor do they state that a silver bullet solution exists. This indicates that well-defined technical requirements and specifications are very important before choosing between CMDF (or not choosing CMDF) since bad decisions can potentially lead to underperforming apps (Biørn-Hansen et al., 2020).

Based on literature study it can be observed that there aren't many studies which include the comparison of the majority of beforehand mentioned frameworks, especially Flutter as newest vs. older ones.

5 Conclusion

The Flutter and React Native ecosystems are in its rise and are constantly improving. On the other hand, Ionic, as representative of hybrid mobile applications CMDF, is going to stay relevant in the future because of the possibility that web developers with experience in SPA frameworks can quickly adapt to the Ionic framework and start building mobile applications. As for Xamarin.Forms, it is yet to be seen how its evolution .NET MAUI is going to be accepted.

The further analysis of mentioned, most recent and currently most popular CMDFs based on different aspects and criteria could be of great use when it comes to decision making of choosing between different approaches and frameworks. Since there is constant improvement of mobile devices and also CMDFs, the interesting topic for further research would be measuring real users' experience of using apps written in mentioned CMDFs (and native languages) in various use cases along with performance measurement which can affect it. This could help to draw a conclusion if the higher performance parameters are very noticeable and therefore crucial in decision making.

The fast-paced development of new versions and emerging of new CMDF and also staying updated with technical advances is a challenge for both industry and academia. Continuous research in this field is especially important due to the rapid pace at which cross-platform frameworks are being released, updated and deprecated. (Biørn-Hansen, Grønli, & Ghinea, 2019)

References

Biørn-Hansen, A., & Ghinea, G. (2018). *Bridging the Gap: Investigating Device-Feature Exposure in Cross-Platform Development*. 8.

Biørn-Hansen, A., Grønli, T.-M., & Ghinea, G. (2019). Animations in Cross-Platform Mobile Applications: An Evaluation of Tools, Metrics and Performance. *Sensors*, *19*(9), 2081. https://doi.org/10.3390/s19092081

Biørn-Hansen, A., Grønli, T.-M., Ghinea, G., & Alouneh, S. (2019). An Empirical Study of Cross-Platform Mobile Development in Industry. *Wireless Communications and Mobile Computing*, *2019*, 1–12. https://doi.org/10.1155/2019/5743892

Biørn-Hansen, A., Grønli, T.-M., Majchrzak, T. A., Kaindl, H., & Ghinea, G. (2022). *The Use of Cross-Platform Frameworks for Google Play Store Apps*. Presented at the Hawaii International Conference on System Sciences. https://doi.org/10.24251/HICSS.2022.934

Biørn-Hansen, A., Rieger, C., Grønli, T.-M., Majchrzak, T. A., & Ghinea, G. (2020). An empirical investigation of performance overhead in cross-platform mobile development

frameworks. *Empirical Software Engineering*, *25*(4), 2997–3040. https://doi.org/10.1007/s10664-020-09827-6

Charland, A., & Leroux, B. (2011). Mobile application development: Web vs. native. *Communications of the ACM*, *54*(5), 49–53. https://doi.org/10.1145/1941487.1941504

Flutter. (n.d.). Flutter architectural overview. Retrieved June 6, 2022, from https://docs.flutter.dev/resources/architectural-overview

Google Trends comparison (React Native vs. Flutter vs. Ionic vs. Xamarin vs. Cordova). (2022, June 9). Retrieved June 9, 2022, from Google Trends website: https://tinyurl.com/CMDFComparison

Huber, S., Demetz, L., & Felderer, M. (2020). Analysing the Performance of Mobile Cross-platform Development Approaches Using UI Interaction Scenarios. In M. van Sinderen & L. A. Maciaszek (Eds.), *Software Technologies* (pp. 40–57). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-52991-8_3

Ionic team. (n.d.). Cordova Community Plugins. Retrieved June 6, 2022, from https://ionicframework.com/docs/native/community#capacitor-support

Işitan, M., & Koklu, M. (2020). Comparison and Evaluation of Cross Platform Mobile Application Development Tools. *International Journal of Applied Mathematics Electronics and Computers*, 273–281. https://doi.org/10.18100/ijamec.832673

JetBrains. (2022, February 21). Cross-platform mobile frameworks used by software developers worldwide from 2019 to 2021. Retrieved May 30, 2022, from Statista website: https://www.statista.com/statistics/869224/worldwide-software-developer-working-hours/

Jia, X., Ebone, A., & Tan, Y. (2018). A performance evaluation of cross-platform mobile application development approaches. *Proceedings of the 5th International Conference on Mobile Software Engineering and Systems*, 92–93. Gothenburg Sweden: ACM. https://doi.org/10.1145/3197231.3197252

Johnson, J., Britch, D., & Dunn, C. (2021, July 8). What is Xamarin.Forms? Retrieved June 6, 2022, from https://docs.microsoft.com/en-us/xamarin/get-started/what-is-xamarin-forms

Justin Johnson, David Britch, Craig Dunn, & Hemant Arya. (2021, July 8). What is Xamarin? Retrieved June 6, 2022, from https://docs.microsoft.com/en-us/xamarin/get-started/what-is-xamarin

Laricchia, F. (2022, February 7). Mobile operating systems' market share worldwide from January 2012 to January 2022. Retrieved May 28, 2022, from Statista website: https://www.statista.com/statistics/272698/global-market-share-held-by-mobile-operating-systems-since-2009/

Latif, M., Lakhrissi, Y., Nfaoui, E. H., & Es-Sbai, N. (2016). *Cross platform approach for mobile aplication development: A survey*. 5.

Latif, M., Lakhrissi, Y., Nfaoui, E. H., & Es-Sbai, N. (2017). Review of mobile cross platform and research orientations. *2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS)*, 1–4. Fez, Morocco: IEEE. https://doi.org/10.1109/WITS.2017.7934674

Lynch, M. (n.d.). Capacitor vs Cordova: Hybrid Mobile App Development. Retrieved June 6, 2022, from https://ionic.io/resources/articles/capacitor-vs-cordova-modern-hybrid-app-development

Malavolta, I., Ruberto, S., Soru, T., & Terragni, V. (2015). Hybrid Mobile Apps in the Google Play Store: An Exploratory Investigation. *2015 2nd ACM International Conference on Mobile Software Engineering and Systems*, 56–59. Florence, Italy: IEEE. https://doi.org/10.1109/MobileSoft.2015.15

Martinez, M. (2019). Two Datasets of Questions and Answers for Studying the Development of Cross-Platform Mobile Applications using Xamarin Framework. *2019 IEEE/ACM 6th International Conference on Mobile Software Engineering and Systems (MOBILESoft)*, 162–173. Montreal, QC, Canada: IEEE. https://doi.org/10.1109/MOBILESoft.2019.00032

MDN Contributors. (2022, April 27). Web Components. Retrieved June 6, 2022, from https://developer.mozilla.org/en-US/docs/Web/Web_Components

Meta Platforms. (2022, March 30). Core Components and Native Components. Retrieved

June 10, 2022, from https://reactnative.dev/docs/intro-react-native-components

Nunkesser, R. (2018). *Beyond Web/Native/Hybrid: A New Taxonomy for Mobile App Development*. 5.

Pinto, C. M., & Coutinho, C. (2018). *From Native to Cross-platform Hybrid Development*. 669–676. IEEE. https://doi.org/10.1109/IS.2018.8710545

Quazi, F. U. R., & Sinha, N. (2018). Android-Platform Based Determination of Fastest CrossPlatform Framework. *International Journal of Computer Science and Mobile Computing*, *7*(9), 1–12.

Que, P., Guo, X., & Zhu, M. (2016). A Comprehensive Comparison between Hybrid and Native App Paradigms. *2016 8th International Conference on Computational Intelligence and Communication Networks (CICN)*, 611–614. Tehri, India: IEEE. https://doi.org/10.1109/CICN.2016.125

Ramel, D. (2022, May 24). .NET MAUI Reaches General Availability, Replacing Xamarin.Forms. Retrieved June 10, 2022, from https://visualstudiomagazine.com/articles/2022/05/24/net-maui-ga.aspx

Sciandra, L. (2019, April 9). The New React Native Architecture Explained: Part Three. Retrieved June 10, 2022, from https://formidable.com/blog/2019/fabric-turbomodules-part-3/

Shah, K., Sinha, H., & Mishra, P. (2019). Analysis of Cross-Platform Mobile App Development Tools. *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*, 1–7. Bombay, India: IEEE. https://doi.org/10.1109/I2CT45611.2019.9033872

Singh, M., & Shobha, G. (2021). Comparative Analysis of Hybrid Mobile App Development Frameworks. *International Journal of Soft Computing and Engineering*, *10*(6), 21–26. https://doi.org/10.35940/ijsce.F3518.0710621

S.Thakare, B., Shirodkar, D., Parween, N., & Parween, S. (2014). State of Art Approaches to Build Cross Platform Mobile Application. *International Journal of Computer Applications*, *107*(20), 22–23. https://doi.org/10.5120/18868-0389

The Apache Software Foundation. (n.d.). Apache Cordova Documentation—Overview. Retrieved June 17, 2022, from https://cordova.apache.org/docs/en/11.x/guide/overview/index.html

Vishal, K., & Kushwaha, A. S. (2018). Mobile Application Development Research Based on Xamarin Platform. *2018 4th International Conference on Computing Sciences (ICCS)*, 115–118. Jalandhar: IEEE. https://doi.org/10.1109/ICCS.2018.00027

Xanthopoulos, S., & Xinogalos, S. (2013). A comparative analysis of cross-platform development approaches for mobile applications. *Proceedings of the 6th Balkan Conference in Informatics on - BCI '13*, 213. Thessaloniki, Greece: ACM Press. https://doi.org/10.1145/2490257.2490292

Zohud, T., & Zein, S. (2021). Cross-Platform Mobile App Development in Industry: A Multiple Case-Study. *International Journal of Computing*, 46–54. https://doi.org/10.47839/ijc.20.1.2091

Software Quality Assessment Standards and Metrics: A Systematic Literature Review

Žana Zekić, Zlatko Stapić
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{zzekic, zstapic}@foi.hr

Abstract. *The paper presents results of performing a systematic literature review process aiming to answer which ISO standards and metrics are used for software quality assessment and which are their area of application. The review was performed on 696 original scientific studies. After applying inclusion criteria, 32 papers were filtered that dealt with this topic. Results of performed review were 8 identified software quality standards with different areas of application. Also, 62 metrics were identified for software quality assessment with corresponding software quality attributes. These metrics are divided into two groups for product quality assessment and for quality in use assessment. In the context of product quality 42 metrics were identified and 20 metrics for quality in use assessment.*

Keywords. Software Quality, Standards, Metrics, Systematic literature review, SLR, Software development

1 Introduction

According to ISO/IEC 25000, software quality is the capability of software product to satisfy stated and implied needs when used under specified condition (Hakim et al., 2016).

Software quality is an important aspect of the development process but in today's dynamic environment many development teams do not have time to focus on software quality. Because of that, at the end of project teams need to devote time to redo some parts of applications due to quality issues (Kasims, 2018).

Thus, to identify software quality assessment standards and metrics that could help development teams to achieve better quality, systematic literature review will be performed. The aim of this research is to find scientific studies that present software quality standards and the area of their application. Also, goal of this research is to find software quality evaluation process and metrics.

There are existing systematic literature reviews that focus on software quality standards and metrics. For example, Rehman & Khan's (2012) research identify different quality attributes that affect software quality and corresponding quality metrics. They identify different quality attributes and metrics. Research performed by Wedyan & Abufakher (Wedyan & Abufakher, 2019) studies the impact of design patterns on quality attributes and the evaluation of quality attributes. They identified 50 different papers with belonging quality attributes and metrics. Another study (Kaur, 2020) focuses on the relationship between code smells and software quality attributes. Their conclusion is that different code smells have the opposite effect on different software quality attributes.

However, to our knowledge, there are no reviews that show a complete overview of software quality ISO standards, metrics, and software quality assessment processes that could help a development team to evaluate and improve their product.

This paper presents the results of systematic literature review on the topic of software quality assessment process and quality standards and metrics. Focus is on ISO standards, which describe the best practices, and they are internationally agreed by experts. Paper is organized as follows: research methodology stating research questions is presented in the second section, third section provides research results in which are identified software quality assessment standards and metrics for an evaluation of different quality attributes which are then analysed and discussed in section four, and finally, section five represents the conclusion of the study.

2 Research methodology

Research methodology that was used in this research is Systematic literature review (SLR). SLR allows to identify, evaluate, and interpret available research relevant to the stated research question. It follows a strict research protocol which makes it replicable and repeatable by other researchers. According to

Kitchenham and Chartes (2007) Systematic literature review has three phases: planning the review, conducting the review, and reporting the review which are designed in a such a way to make it objectivistic and bias free.

2.1. Planning the review

Planning is the first phase of the literature review. In this phase, research questions are stated and a review protocol is created and validated (Brereton et al., 2007).

To find all relevant studies, two research questions are stated. **RQ1 -** What standards are used for software quality assessment and what are their areas of application? **RQ2 -** What metrics are used for measuring software quality attributes?

Digital databases used in this search have been recommended by study (Brereton et al., 2007) as the most relevant databases in software engineering. *Table 1.* shows used digital databases and search string.

Table 1. Digital databases and search string

Digital databases	Search string
IEEExplore	("software quality") AND ("ISO") AND ("standard" OR "model")
ACM Digital library	
Scopus	
Web of Science	

The review process had *four phases*, namely: (1) searching digital databases by a search string; (2) filtering studies by inclusion criteria applied to abstracts; (3) filtering studies by inclusion criteria applied to introduction and conclusion and (4) full-text review and removing duplicates.

Defined *inclusion criteria* were as follows: (1) study focuses on ISO software quality standards and quality attributes metrics; (2) study is an original scientific work; (3) study is published after 2017; (4) study is written in English and (5) study is published in scientific conference or journal.

Furthermore, for study to pass to the data extraction phase it needs to meet all defined criteria.

2.2. Conducting the review

The four-phase process of filtering the studies is presented in *Table 2.* In the first phase, 696 studies were found. In the next phase, 158 papers are filtered. After filtering studies by applying inclusion criteria on introduction and conclusion, 77 studies were filtered. In the end, 32 papers met the defined criteria and were used in the data extraction phase. Studies were analysed to answer research questions. Title, authors, type of the study, and major results and conclusions were extracted from the studies. Furthermore, from examples of software quality assessment process defined in filtered studies ISO standards with the corresponding area of application, quality attributes and metrics will be extracted.

Table 2. Review phases

Source	Phase 1	Phase 2	Phase 3	Phase 4
IEE Explore	128	35	23	15
ACM Digital library	210	28	7	1
Scopus	193	60	27	7
Web of Science	165	35	20	9
TOTAL	**696**	**158**	**77**	**32**

3 Results

Results of the conducted research are shown by stated research questions.

Quality standards are used for software quality assessment. However, in different areas of software engineering, different standards are used. In this study, we identified eight quality standards and their area of application (**RQ1**). Results are presented in *Table 3.* Identified areas of application represent different areas of software engineering. Some areas are being repeated in the table because one standard can be used in many areas of software engineering.

Table 3. Software quality standards and their area of application

Standard	Area of application	References
ISO/IEC 25000	• Web applications	Wiese et al. (2021)
ISO/IEC 25010	• Mobile games • Mobile applications • Information system • Web application security • E-books • Ecommerce • Internet of things	Alves, Albino, et al. (2016) Alves, Savaris, et al. (2016) Anggraini et al. (2019, p. 25010) Hakim et al. (2016) Hasanah et al. (2020) Haslinda et al. (2015) Hovorushchenko & Pomorova (2016, p. 2011) Idri, Bachiri, & Fernández-Alemán (2016) Idri, Bachiri, Luis Fernandez-Aleman, et al. (2016) Idri et al. (2017)

		Manglapuz & Lacatan (2019) Peters & Aggrey (2020) Rahmi Dewi et al. (2020) Sekarini et al. (2020) Tambotoh et al. (2017) Trichkova-Kashamova (2021) Trisnadoli et al (2015) Zulfa et al. (2020)
ISO/IEC 25012	• Data security	Verdugo & Rodriguez (2020)
ISO/IEC 25022	• Mobile games • Mobile applications	Sulla-Torres et al. (2020) Trisnadoli et al. (2015)
ISO/IEC 25023	• Web application security	Aziz et al. (2018)
ISO 25062	• Mobile applications	Moumane, Idri, & Nafil (2016) Moumane, Idri, & Abran (2016)
ISO/IEC 9126	• Mobile games • Information systems • E-learning • Web applications	Barros et al. (2015) Rochimah et al. (2015) Trisnadoli et al. (2015) Dzulfiqar et al. (2018) Beckhauser et al. (2018) Molnar et al. (2019)
ISO 9241	• Mobile applications	Moumane, Idri, & Nafil (2016) Moumane, Idri, & Abran (2016)

Related to second research question (**RQ2**), the review showed metrics for the assessment of quality attributes. We identified 20 different quality attributes and belonging metrics. In total, 62 different metrics are reported. Metrics are divided into two tables depending on the type of quality. *Table 4.* shows the quality attributes and their metrics for *products quality measurement*, while *Table 5.* shows the quality attributes and their metrics for *quality in use assessment*. For product quality evaluation, eight software quality attributes are identified with 42 belonging metrics.

Table 4. Product quality assessment

Quality attributes	Metrics	References
Functional Suitability	• Questionary • Expert evaluation • System compatibility with the required level of precision • The suitability of the function with its purpose • Conformity of usage function with usage procedure	Haslinda et al. (2015) Hasanah et al. (2020) Idri et al. (2017) Trichkova-Kashamova (2021)
Reliability	• Questionary • Expert evaluation	Haslinda et al. (2015) Idri et al. (2017) Trichkova-Kashamova (2021)
Security	• Questionary • Access controllability • Data Encryption • Strength of cryptographic algorithm • Data Integrity Conformance • Internal Data Corruption Prevention • Validity of Array Accesses • Utilization of Digital • Signature Access • System Log Retention Conformance • Authentication Protocol Conformance • Authentication protocol measures • Establishment of authentication rules measures • User Health and Safety • Expert evaluation	Haslinda et al. (2015) Soad et al. (2016) Aziz et al. (2018) Sekarini et al. (2020) Hakim et al. (2016) Trisnadoli et al. (2015) Trichkova-Kashamova (2021)
Performance	• Questionary • Expert evaluation	Haslinda et al. (2015) Idri et al. (2017)

		Trichkova-Kashamova (2021)
Maintainability	• Questionary • Code analysis without running • Expert evaluation • Size • Complexity • Coupling • Cohesion • Encapsulation	Haslinda et al. (2015) Rahmi Dewi et al. (2020) Trichkova-Kashamova (2021) Wiese et al. (2021)
Compatibility	• Questionary • Expert evaluation	Haslinda et al. (2015) Trichkova-Kashamova (2021)
Portability	• Questionary • Adaptability of data structures • Organizational environment adaptability • Porting user friendliness • System software environmental adaptability • Ease of setup re-try Installation effort Installation flexibility • Continued use of data • Function inclusiveness • Expert Evaluation	Kaur (2020) Idri, Bachiri, Luis Fernandez-Aleman, et al. (2016) Trichkova-Kashamova (2021)
Operability	• Questionary • Input validity Checking • User operation • Cancellability • User operation Undoability • Customizability • Physical • Accessibility • Operation status monitoring capability • Operational Consistency • Message Clarity • Interface element clarity • Operational error recoverability	Idri et al. (2017) Rochimah et al. (2015)

Quality attributes and their metrics for *quality in use assessment* are presented in the following table. For this type of quality, 12 different attributes and 20 metrics are reported.

Table 5. Quality in use assessment

Quality attributes	Metrics	References
Usability	• Goal completeness • Error Frequency • Goal Time Efficiency • Satisfaction Scale • Questionary • System Usability Scale • Expert evaluation	Trisnadoli et al. (2015) Haslinda et al. (2015, p. 25010) Soad et al. (2016) Alves, Albino, et al. (2016) Trichkova-Kashamova (2021)
Efficiency	• Questionary • Expert evaluation	(Haslinda et al., 2015) (Idri et al., 2017) Trichkova-Kashamova (2021)
Flexibility	• Flexible Context of Use	Trisnadoli et al. (2015)
Performance	• Questionary • Expert evaluation	Haslinda et al. (2015) Idri et al. (2017) Trichkova-Kashamova (2021)
Operability	• Questionary • Input validity Checking • User operation Cancellability • User operation Undoability • Customizability • Physical Accessibility • Operation status monitoring capability Operational • Consistency Message Clarity • Interface element clarity	Idri et al. (2017) Rochimah et al. (2015)

	• Operational error recoverability	
Functionality	• Questionary	Soad et al. (2016)
Educational (Pedagogical)	• Questionary	Soad et al. (2016)
Support	• Questionary	Soad et al. (2016)
Learnability	• Completeness of user documentation and/or help facility • Questionary	Rochimah et al. (2015) Dzulfiqar et al. (2018)
Understandability	• Completeness of description • Demonstration capability • Evident Functions • Function understandability • Questionary	Rochimah et al. (2015) Dzulfiqar et al. (2018)
Attractiveness	• Questionary	Dzulfiqar et al. (2018)
Sociocultural (Communication)	• Questionary	Soad et al. (2016)

4 Discussion

Software quality standards define quality model which contains quality characteristics and sub-characteristics and enable to evaluate software quality. A lot of scientific studies put focus on a software quality standard. Some of them use different standards to measure software quality (Rahmi Dewi et al., 2020, p. 25010). The most popular and most widely used framework for software quality is ISO/IEC 25000. It is a series of standards, also known as SQuaRE series of standards (System and Software Quality Requirements and Evaluation). This series contains the following divisions: ISO/IEC 2502n - Quality Measurement Division; ISO/IEC 2503n - Quality Requirements Division; ISO/IEC 2504n - Quality Evaluation Division (Hakim et al., 2016). According to this framework software quality is the capability of a software product to enable the behaviour of a system to satisfy stated and implied needs when used under specified conditions (Hakim et al., 2016).

One of the most popular quality standards, ISO/IEC 25010, defines two aspects of quality: product quality and quality in use. Product quality describes properties of a software product, and they are usually available during the development process. On the other hand, quality in use describes the impact that the product has on users (Idri et al., 2017). A quality in use refers not only to the technological view, but also to pedagogical, socio-cultural, and socio-economic aspects. For example, pedagogical characteristics refer to intends to address the issues related to teaching and learning, and communication refers to the ability to provide communication and interaction among users. (Soad et al., 2016). A product quality contains eight attributes: Functional suitability, Reliability, Performance efficiency, Operability, Security, Compatibility, Maintainability and Transferability. These attributes are divided into sub-attributes that can be measured. On the other hand, quality in use contains five attributes also divided into sub-attributes.

However, in this paper we combined attributes and sub-attributes identified in reviewed papers for systematic presentation of software quality metrics. These papers used attributes and sub-attributes defined by ISO standards.

With ISO/IEC 25010 one of the most used standards is ISO/IEC 9126. This is a generic quality model which contains six attributes and 27 sub-attributes. This standard also defines metrics for sub-attributes measurements. Studies show that is most used in the following areas: mobile games development, information systems, and e-learning (Trisnadoli et al., 2015).

Furthermore, above-mentioned software quality standards define metrics for software quality assessment. Metrics are defined by software quality attributes and sub-attributes. Different types of quality require use of different metrics. Product quality contains the following quality attributes: functional suitability, reliability, security, performance, maintainability, portability, and operability (Trichkova-Kashamova, 2021). For these quality attributes in most cases parameters are clearly defined and results of measurements can be easily compared. For example, in the context of security, authentication protocol measures are calculated by the function F, mathematical formula is F = A/B where A is number of provided authentication protocols (e.g., user id/password or IC card) and B is number of required authentication protocols in the specification (Hakim et al., 2016). On the other hand, quality in use contains the following quality attributes: usability, efficiency, flexibility, operability, functionality, pedagogical, support, learnability, understandability, attractiveness, and communication (Trichkova-Kashamova, 2021). In this case, for quality measuring is necessary to collect data from users. For this purpose, usually a questionary is used to collect data and additional analysis needs to be provided to rate quality in the context of use (Soad et al., 2016).

Nevertheless, one of the most used ways to evaluate software quality is expert evaluation. The implementation of this approach starts with the evaluation of the system by experts in the field of

information technology and users of the system. A set of indicators is described, and they are presented in the survey. Collected data is analysed. Decisions are taken about the functional development of the system after an inspection of the obtained values (Trichkova-Kashamova, 2021).

5 Conclusion

The paper provides the presentation of most used software quality standards and their area of application. Also, the paper presents metrics for software quality assessment grouped by type of software quality.

Systematic literature review was performed on 696 original studies. After applying inclusion criteria, 32 papers were filtered that dealt with this topic. The result of the performed review was 8 identified software quality standards with different areas of application. Also, we identified 62 metrics for software quality assessment for measuring different software quality attributes. These metrics are divided into two groups for product quality assessment and for quality in use assessment. Furthermore, 42 metrics were identified in the context of product quality and 20 for quality in use assessment.

This study showed different standards and metrics that could help development teams to improve software quality during the development process. However, usage of these standards and metrics in practice could be expensive and complicated. For that reason, additional research is necessary to identify appropriate tools for faster software quality assessment. Identifying tools for software quality assessment will enable development team to easily use them in development process and improve quality in iteration base.

Acknowledgment

This work has been supported in part by Croatian Science Foundation under the project IP-2019-04-4864.

References

Alves, J. M., Albino, D. B. L., Resener, M. C., Zannin, M., Savaris, A., von Wangenheim, C. G., & von Wangenheim, A. (2016). Quality Evaluation of Poison Control Information Systems: A Case Study of the DATATOX System. 2016 IEEE 29th International Symposium on Computer-Based Medical Systems (CBMS), 30–35. https://doi.org/10.1109/CBMS.2016.53

Alves, J. M., Savaris, A., Von Wangenheim, C. G., & Von Wangenheim, A. (2016). Software quality evaluation of the laboratory information system used in the Santa Catarina state integrated telemedicine and telehealth system. 2016-August, 76–81. https://doi.org/10.1109/CBMS.2016.51

Anggraini, N., Putra, M. J. D., & Hakiem, N. (2019). Development of an Islamic Higher Education Institution Tracer Study Information System and It's Performance Analysis using ISO/IEC 25010. https://doi.org/10.1109/CITSM47753.2019.8965356

Aziz, M. N., Sapta, I. M., & Rochimah, S. (2018). Security Characteristic Evaluation Based on ISO/IEC 25023 Quality Model, Case Study: Laboratory Management Information System. 332–336. https://doi.org/10.1109/EECCIS.2018.8692982

Barros, P. R. M., Cazella, S. C., & Flores, C. D. (2015). Analyzing softwares in medical education focusing on quality standards. 2015-July, 292–297. Scopus. https://doi.org/10.1109/CBMS.2015.20

Beckhauser, E., von Wangenheim, A., Savaris, A., & Krechel, D. (2018). ISO Compliant Evaluation of a Teledermatology Structured Report System for the Brazilian Public Healthcare. In J. Hollmen, C. McGregor, P. Soda, & B. Kane (Eds.), 2018 31st Ieee International Symposium on Computer-Based Medical Systems (cbms 2018) (pp. 135–140). Ieee. https://doi.org/10.1109/CBMS.2018.00031

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. Journal of Systems and Software, 80(4), 571–583. https://doi.org/10.1016/j.jss.2006.07.009

Dzulfiqar, M. D., Khairani, D., & Wardhani, L. K. (2018). The Development of University Website using User Centered Design Method with ISO 9126 Standard. 2018 6th International Conference on Cyber and IT Service Management (CITSM), 1–4. https://doi.org/10.1109/CITSM.2018.8674325

Hakim, H., Sellami, A., & Ben Abdallah, H. (2016). Evaluating Security in Web Application Design Using Functional and Structural Size Measurements. 2016 Joint Conference of the International Workshop on Software Measurement and the International Conference on Software Process and Product Measurement (IWSM-MENSURA), 182–190. https://doi.org/10.1109/IWSM-Mensura.2016.036

Hasanah, N. A., Atikah, L., & Rochimah, S. (2020). Functional Suitability Measurement Based on ISO/IEC 25010 for e-Commerce Website. 2020 7th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), 70–75. https://doi.org/10.1109/ICITACEE50144.2020.9239194

Haslinda, H., Fahmy, F., Sukinah, S., Roslina, R., Fariha, F., Suhana, S., Azliza, A., & Shiratuddin, N. (2015). Evaluation of e-Book applications using ISO 25010. 114–118. https://doi.org/10.1109/ISTMET.2015.7359012

Hovorushchenko, T., & Pomorova, O. (2016). Evaluation of mutual influences of software quality characteristics based ISO 25010:2011. 80–83. https://doi.org/10.1109/STC-CSIT.2016.7589874

Idri, A., Bachiri, M., & Fernández-Alemán, J. L. (2016). A Framework for Evaluating the Software Product Quality of Pregnancy Monitoring Mobile Personal Health Records. Journal of Medical Systems, 40(3), 1–17. https://doi.org/10.1007/s10916-015-0415-z

Idri, A., Bachiri, M., Fernández-Aleman, J. L., & Toval, A. (2017). ISO/IEC 25010 Based Evaluation of Free Mobile Personal Health Records for Pregnancy Monitoring. 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC), 1, 262–267. https://doi.org/10.1109/COMPSAC.2017.159

Idri, A., Bachiri, M., Luis Fernandez-Aleman, J., & Toval, A. (2016). Experiment Design of Free Pregnancy Monitoring Mobile Personal Health Records Quality Evaluation. 2016 Ieee 18th International Conference on E-Health Networking, Applications and Services (Healthcom), 466–471. https://www.webofscience.com/wos/woscc/full-record/WOS:000391459700084

Idri, A., Sardi, L., & Fernández-Alemán, J. L. (2018). Quality evaluation of gamified blood donation apps using ISO/IEC 25010 standard. 5, 607–614. https://doi.org/10.5220/0006724806070614

Kasims, G. (2018). Applying Lean to Improve Quality in Software Development Projects. Proceedings of the 2nd International Conference on Business and Information Management, 130–134. https://doi.org/10.1145/3278252.3278254

Kaur, A. (2020). A Systematic Literature Review on Empirical Analysis of the Relationship Between Code Smells and Software Quality Attributes. Archives of Computational Methods in Engineering, 27(4), 1267–1296. https://doi.org/10.1007/s11831-019-09348-6

Kitchenham, B., & Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering.

Manglapuz, S. J. R., & Lacatan, L. L. (2019). Academic management android application for student performance analytics: A comprehensive evaluation using ISO 25010:2011. International Journal of Innovative Technology and Exploring Engineering, 8(12), 5085–5089. Scopus. https://doi.org/10.35940/ijitee.L2735.1081219

Molnar, A.-J., Neamțu, A., & Motogna, S. (2019). Longitudinal evaluation of software quality metrics in open-source applications. 80–91. https://doi.org/10.5220/0007725600800091

Moumane, K., Idri, A., & Abran, A. (2016). Usability evaluation of mobile applications using ISO 9241 and ISO 25062 standards. SpringerPlus, 5(1). https://doi.org/10.1186/s40064-016-2171-z

Moumane, K., Idri, A., & Nafil, K. (2016). An Empirical Evaluation of Mobile Software Usability Using ISO 9126 and QoS DiffServ Model. In H. Fujita & G. A. Papadopoulos (Eds.), New Trends in Software Methodologies, Tools and Techniques (Vol. 286, pp. 177–188). Ios Press. https://doi.org/10.3233/978-1-61499-674-3-177

Peters, E., & Aggrey, G. K. (2020). An ISO 25010 based quality model for ERP systems. Advances in Science, Technology and Engineering Systems, 5(2), 578–583. https://doi.org/10.25046/aj050272

Rahmi Dewi, M., Ngaliah, N., & Rochimah, S. (2020). Maintainability Measurement and Evaluation of myITS Mobile Application Using ISO 25010 Quality Standard. 2020 International Seminar on Application for Technology of Information and Communication (ISemantic), 530–536. https://doi.org/10.1109/iSemantic50169.2020.9234283

Rehman, Sr., & Khan, S. U. (2012). Swot Analysis Of Software Quality Metrics For Global Software Development: A Systematic Literature Review Protocol. IOSR Journal of Computer Engineering, 1, 1–7. https://doi.org/10.9790/0661-0210107

Rochimah, S., Rahmani, H. I., & Yuhana, U. L. (2015). Usability characteristic evaluation on administration module of Academic Information System using ISO/IEC 9126 quality model. 363–368. https://doi.org/10.1109/ISITIA.2015.7220007

Sekarini, D., Alfiani, F. S., & Rochimah, S. (2020). Security Characteristic Evaluation of New Students Admission Information System Based on ISO/IEC 25010 Quality Standard. 2020 12th International Conference on Information Technology and Electrical Engineering (ICITEE), 120–124. https://doi.org/10.1109/ICITEE49829.2020.9271756

Soad, G. W., Duarte Filho, N. F., & Barbosa, E. F. (2016). Quality evaluation of mobile learning applications. 2016 IEEE Frontiers in Education Conference (FIE), 1–8. https://doi.org/10.1109/FIE.2016.7757540

Sulla-Torres, J., Gutierrez-Quintanilla, A., Pinto-Rodriguez, H., Gomez-Campos, R., & Cossio-Bolanos, M. (2020). Quality in Use of an Android-based Mobile Application for Calculation of Bone Mineral Density with the Standard ISO/IEC

25022. International Journal of Advanced Computer Science and Applications, 11(8), 158–163.

Tambotoh, J. J. C., Isa, S. M., Gaol, F. L., Soewito, B., & Warnars, H. L. H. S. (2017). Software quality model for Internet of Things governance. https://doi.org/10.1109/ICODSE.2016.7936138

Trichkova-Kashamova, E. (2021). Applying the ISO/IEC 25010 Quality Models to an Assessment Approach for Information Systems. 2021 12th National Conference with International Participation (ELECTRONICA), 1–4. https://doi.org/10.1109/ELECTRONICA52725.2021.9513662

Trisnadoli, A., Hendradjaya, B., & Danar Sunindyo, W. (2015). A proposal of quality model for mobile games. 377–381. https://doi.org/10.1109/ICEEI.2015.7352530

Verdugo, J., & Rodriguez, M. (2020). Assessing data cybersecurity using ISO/IEC 25012. Software Quality Journal, 28(3), 965–985. https://doi.org/10.1007/s11219-019-09494-x

Wedyan, F., & Abufakher, S. (2019). Impact of Design Patterns on Software Quality: A Systematic Literature Review. In IET Software (Vol. 14). https://doi.org/10.1049/iet-sen.2018.5446

Wiese, L., Wiese, I., & Lietz, K. (2021). Software Quality Assessment of a Web Application for Biomedical Data Analysis. 25th International Database Engineering & Applications Symposium, 84–93. https://doi.org/10.1145/3472163.3472172

Zulfa, F., Munawaroh, H., & Rochimah, S. (2020). Portability Characteristics Evaluation of MyITS Mobile using ISO/IEC 25010 Quality Standard. 2020 International Seminar on Application for Technology of Information and Communication (ISemantic), 537–542. https://doi.org/10.1109/iSemantic50169.2020.9234241

Exploring the Access to the Static Array Elements via Indices and via Pointers — the Introductory C++ Case

Robert Logožar, Matija Mikac
University North, Croatia
Dpt. of Electrical Engineering
104. brigade 3, 42 000 Varaždin, Croatia
{robert.logozar, matija.mikac}@unin.hr

Danijel Radošević
University of Zagreb
Faculty of Organization and Informatics
Pavlinska 2, 42 000 Varaždin, Croatia
darados@foi.unizg.hr

Abstract. *We revisit the old but formally still undecided debate on the time efficiency of accessing the elements of 1D arrays via indices versus accessing them via pointers. To analyze that, we have programmed benchmarks of minimal complexity in the C++ language and inspected the machine code of their 32-bit compilation in the x86 assembly language. Before the performance study, we have briefly compared a few methods used for the execution time measurements. There is no advantage in the use of pointers over indices except for some benchmarks and array (data) types, while for the others, the exact opposite may be true. The parallel aim was to provide a ground for the possible further analysis and measurements of this kind on different computers and platforms, and different languages.*

Keywords. Static arrays, pointers, C/C++, accessing the array elements, time measurement and efficiency.

1 Introduction

The array is the simplest and ubiquitous data structure that resembles the organization of the computer main memory itself. As such, it is unavoidable in computer programming and all general-purpose high-level languages implement it one way or the other. The simplest way to access the array elements — which follow the mathematical notation of vectors and matrices — is by "subscripting" the arrays with their *indices*. In Pascal and the C language, that became possible also via *pointers*, which are the addresses of the defined data types.

The fathers of the C language, B. Kernighan and D. Ritchie, in their C language "bible," devoted the whole chapter 5 to the pointers and arrays and their relation (Kernighan & Ritchie, 1978, 1988, known as K&R). In §5.3 they say: "*Any operation that can be achieved by array subscripting can also be done with pointers. The pointer version will in general be faster but, at least to the uninitiated, somewhat harder to understand.*"

This claim must be echoing in the minds of many computer scientists and practical programmers who have read the valuable classical literature and care to write the most efficient code. Namely, if we follow the implicit suggestion, should we insist on accessing the array elements by using pointers, and if so, what are the time efficiency improvements? Kernighan and Ritchie did not support their statement with any explanation or proof. They have even relativized it in the further elaboration on the array element access in the rest of that chapter and the textbook. Besides that, when trying to see if someone else tried to corroborate or dispute this thesis, we could not find any systematic work or firm results on this topic from other authors.

To investigate this subject properly, one should get a deeper insight into it and reach the final verdict only after the concrete time measurements. That is, for the appropriately tailored benchmarks of the two approaches, we must measure their *benchmark execution times*, which we shall abbreviate as BETs.

We did some simple, preliminary time measurements roughly a decade ago. They showed that the access to the elements of one-dimensional arrays of some integer types (`short int`, `int`) via pointers was roughly 15% to 20% faster than the access via indices. We did not pursue the measurements more systematically nor did we investigate the causes of this behavior at that time. We left that — seemingly trivial research — for some "future work," which finally continues with this paper. Thus, our first aim is to provide an introduction for a detailed analysis of the provided benchmarks and their execution time measurements. At the same time, we shall expose our first results obtained by using a modern integrated developing environment (IDE) on a relatively contemporary computer and today's common operating system (OS), but leaving the many details from this shorter version of our report.

Concerning the outline of this paper, in the next, section 2, we shortly expose the basics of the array data structure and the pointer data type and their implementation in the C/C++ languages. Section 3 presents and explains our benchmarks and exposes their assembly language code. Section 4 discusses the methods, program setup and prerequisites for the time measurements. There, we present the results of our BET measurements and discuss them. Section 5 concludes this paper and opens several new directions and topics for future work.

2 Arrays and Pointers

Before elaborating on our benchmark details, we shall briefly outline the basics of the arrays and pointers, with an emphasis on their implementation in C/C++.

2.1 Arrays

One-dimensional (1D) arrays serve for storing the series of n equal a_i elements, where the nonnegative integer index i spans through n consecutive values. For example, with $i = 1, 2, \dots, n$, the elements a_i could be interpreted as the components of vector $\boldsymbol{a}$ in an n-dimensional space. Following this mathematical notation, the standard syntax for accessing the array elements in high-level programming languages requires stating the array name and the element index. Based on that, the compilers ensure the run-time calculation of i-th array element memory address as:

$$A(i) = A_0 + l_T \times (i - i_0), \quad i = i_0,\ i_0 + 1, \dots, i_0 + n - 1. \qquad (1a)$$

Here, A_0 is the address of the starting element, the one with the index i_0, and l_T is the size of the element data type (T) in the number of bytes (B)—here, of course, in their original meaning of the basic memory location.

2.2 C/C++ Arrays and Pointers

The C and C++ languages fix the starting index to $i_0 = 0$, so that its value needs not to be subtracted from i. This leads to the simplest possible formula and the fastest possible address calculation:

$$A(i) = A_0 + l_T \times \mathrm{i}, \quad i = 0, 1, \dots, n - 1. \qquad (1b)$$

The value of the i-th element is the contents of the $A(i)$ address expressed as a certain data type.

As hinted in the introduction (§1), the element addresses and their contents can be operated also by using the variables of the *pointer data type*, or *pointers*. They hold the addresses with assigned data types and can be shortly described as *typed addresses*.

Altogether, this leads to the three possible ways of accessing the array elements, which we immediately write in the C/C++ languages (K&R 1988, Stroustrup 1997), in the same way as they appear in the benchmarks' source code in Listing 1:

1) by using the element index: `iX[i]`;
2) by using the pointer arithmetic and dereferencing the obtained pointer: `*(pI0 + i)`;
3) by incrementing the pointer: `*(++pI)`.

The first two are of the *random access* kind because their elements can be accessed directly via its index, regardless of the previously accessed element. The third access applies only to the passage through successive array elements. To enable the comparison of the time efficiency of all three access types, here we restrict the consideration only to such passage through the arrays, which is quite common in practice.

The above general deliberation is applicable to both *static* and *dynamic* arrays, but in the further text, we restrict our deliberation to the former.

3 Our Benchmarks

In this section, we describe and analyze our benchmark routines. They are the core of our three C++ test programs, created in MS Visual Studio, Community Version 2019 (further on VS-CV 2019). The slight variations between them served to investigate the many peculiarities on which the execution times depend (§4.3). Because of the repetitive code for each data type, the programs built up to a bit more than 1000 lines each.

3.1 Benchmark Source Code

Listing 1 shows *our standard benchmarks* for the `int` type of 1D arrays. In each version of our test program, there are benchmark loops similar to those in Listing 1 but for the arrays of all the six standard data types — four integer and two floating-point types:

- `char` (1B), `short int` = `short` (2B), `int` (4B) (in Listing 1), and `long long int` = abbr. `llint` (8B);
- `float` (4B), and `double` (8B).

To avoid the user stack overflow, the very large arrays, like ours, must be declared as `static` or as global.

Listing 1. Our standard benchmarks for accessing the elements of a (large) C/C++ `int` array by: 1) indices, 2) pointer arithmetic, 3) incrementing the pointer.

```
// Compiler Optimizations OFF!
#define intMid     1111
#define intLrg 22222222
typedef unsigned int uint;
// Defining the (static) int array:
const uint cuiN = 20000000;
uint uiN1 = cuiN;
static int iX[cuiN] = {0, };
iLVal = intMid, iRVal = intLrg;
// Pointers to the 0-th and last el.:
int *pI0 = iX, *pI1 = iX + uiN1;

 // A. Storing into array elements
// 1) Via index: iX[uI]:
for (uint uI = 0; uI < uiN1; ++uI)
     iX[uI] = iRVal;
// 2) Via pointer arithmetic: *(pI0 + uI).
for (uint uI = 0; uI < uiN1; ++uI)
   *(pI0 + uI) = iRVal;
// 3) Via icrementing the pointer: *(++pI).
for (int* pI = iX; pI < pI1; ++pI)
   *pI = iRVal;
// B. Retrieving from array elements
// 1) Via index: iX[uI]:
for (uint uI = 0; uI < uiN1; ++uI)
     iLVal = iX[uI];
// 2) Via pointer arithmetic: *(pI0 + uI).
for (uint uI = 0; uI < uiN1; ++uI)
   iLVal = *(pI0 + uI);
// 3) Via icrementing the pointer: *(++pI).
for (int* pI = iX; pI < pI1; ++pI)
   iLVal = *pI;
```

For each of the three observed and implemented access methods, there are two benchmarks:

A. for storing a value in the `uI`-th array element, and
B. for retrieving (fetching) a value from the `uI`-th array element into a variable.

To investigate primarily the observed influence of different element-accessing mechanisms, we have kept the A and B types of operations as simple as possible. In A, instead of storing the quasi-random numbers into the array elements that serve as *l*-values, we assigned them the value of the unchanging `iRVal` variable. In the action B, one and only one *l*-value, the variable `iLVal`, is assigned the values of the array elements.

In the release version of the program, the compiler could and would optimize both actions if set so, especially B, because its final outcome is the single assignment: `iLVal = iX[uiN1 - 1]`. That is why these benchmarks must run without any optimization.

The objection holds that such circumstances can be considered very artificial and that the arrays could have been filled up in some other ways. However, if our *r*- or *l*-values were more complex with the aim to prevent the extreme optimization efficiency, the net effect would have been the same as now. Furthermore, the investigation of the assembly language code (in the next subsection), would not be possible in VS-CV 2019, as it is for the non-optimized, debug version. Therefore, in this introductory investigation, we stick to the proposed benchmarks, with elementary actions.

We shall discuss other details of the concrete program implementation as needed.

3.2 Benchmark Machine Code Disassembled

When running the programs in the debug mode, the VS integrated developing environment (IDE) enables the in-place presentation of the Intel assembly language instructions in symbolic form, after disassembling the machine code of the debug version of the program (Intel 2022-1). The x86 (32-bit) and x64 (64-bit) versions of assembly languages are available, but in this paper, we focus on the still-standard 32-bit version. A reader more interested in this topic can find a concise review of the x86 assembly language in (Evans 2022), whereas the exhaustive reference is in (Intel 2022-2). For our six benchmark routines (A/B.1, 2, 3) with the array of `int` data type, this is shown in Listing 2.

3.2.1 Compilation of the `for`-loops (`int` arrays)

Of the six `for`-loops from our benchmarks in Listing 1, the first two of the A and B types, A/B.1 and A/B.2, have the standard `for`-loop heads, with (rising) indices. Because of the same source form, their compilation results in a series of eight machine instructions, which are *equivalent* to the A.1 `for`-loop head. I.e., they differ from each other only in the possible use of different alternative registers: `ECX` instead of `EAX`, `EDX` instead of `ECX`, and `EAX` instead of `ECX`. Additionally, (`jae`) jumps to the: i) loop *conditions* (`cmp` instruction), ii) *exits* (behind the last, `jmp` instruction in the loop body), and iii) *loop-expressions*, i.e., the index increments, starting at the second `mov` instruction.

Listing 2. Intel x86 assembly language code of the crucial parts of our benchmarks from Listing 1. The order of the code snippets is rearranged (see §4.2.1).

```
 // A. Storing into array elements
// 1) Via index: iX[uI]
    for (uint uI = 0; uI < uiN1; ++uI)
00C738B8  mov   dword ptr [ebp-7E8h],0
00C738C2  jmp   main+20E3h (0C738D3h)
00C738C4  mov   eax,dword ptr [ebp-7E8h]
00C738CA  add   eax,1
00C738CD  mov   dword ptr [ebp-7E8h],eax
00C738D3  mov   ecx,dword ptr [ebp-7E8h]
00C738D9  cmp   ecx,dword ptr [ebp-77Ch]
00C738DF  jae main+2106h (0C738F6h)
        iX[uI] = iRVal;
00C738E1  mov   edx,dword ptr [ebp-7E8h]
00C738E7  mov   eax,dword ptr [ebp-738h]
00C738ED  mov dword ptr iX (0C826A8h)[edx*4],eax
00C738F4  jmp   main+20D4h (0C738C4h)
// 2) Via pointer arithmetic: *(pI0 + uI)
 for (uint uI = 0; uI < uiN1; ++uI) // As in A.1.
        *(pI0 + uI) = iRVal;
00C73B5B  mov   eax,dword ptr [ebp-7F8h]
00C73B61  mov   ecx,dword ptr [ebp-76Ch]
00C73B67  mov   edx,dword ptr [ebp-738h]
00C73B6D  mov   dword ptr [ecx+eax*4],edx
00C73B70  jmp   main+234Eh (0C73B3Eh)
// 3) Via incrementing the pointer: *(++pI)
    for (int* pI = iX; pI < pI1; ++pI)
00C73DD1  mov   dword ptr [ebp-808h], offset iX
                                    (0C826A8h)
00C73DDB  jmp   main+25FCh (0C73DECh)
00C73DDD  mov   edx,dword ptr [ebp-808h]
00C73DE3  add   edx,4
00C73DE6  mov   dword ptr [ebp-808h],edx
00C73DEC  mov   eax,dword ptr [ebp-808h]
00C73DF2  cmp   eax,dword ptr [ebp-770h]
00C73DF8  jae   main+261Ah (0C73E0Ah)
       *pI = iRVal;
00C73DFA  mov   ecx,dword ptr [ebp-808h]
00C73E00  mov   edx,dword ptr [ebp-738h]
00C73E06  mov   dword ptr [ecx],edx
00C73E08  jmp   main+25EDh (0C73DDDh)
 // B. Retrieving from array elements
// 1) Via index: iX[uI]
 for (uint uI = 0; uI < uiN1; ++uI) // As in A.1.
        iLVal = iX[uI];
00C73A09  mov   eax,dword ptr [ebp-7F0h]
00C73A0F  mov ecx,dword ptr iX (0C826A8h)[eax*4]
00C73A16  mov   dword ptr [ebp-784h],ecx
00C73A1C  jmp   main+21FCh (0C739ECh)
// 2) Via pointer arithmetic: *(pI0 + uI)
 for (uint uI = 0; uI < uiN1; ++uI) // As in A.1.
        iLval = *(pI0 + uI);
00C73C9B  mov   edx,dword ptr [ebp-800h]
00C73CA1  mov   eax,dword ptr [ebp-76Ch]
00C73CA7  mov   ecx,dword ptr [eax+edx*4]
00C73CAA  mov   dword ptr [ebp-784h],ecx
00C73CB0  jmp   main+248Eh (0C73C7Eh)
// 3) Via incrementing the pointer: *(++pI)
for (int* pI = iX; pI < pI1; ++pI) // As in A.3.
        iLval = *pI;
00C73F3D  mov   ecx,dword ptr [ebp-810h]
00C73F43  mov   edx,dword ptr [ecx]
00C73F45  mov   dword ptr [ebp-784h],edx
00C73F4B  jmp   main+2730h (0C73F20h)
```

Summarily, all these for-loops consist of eight (8) machine instructions in the loop head and one (1) additional at the end of the loop body, totaling nine (9) instructions, which are stored in altogether 43B (cf. type int benchmark A.1, the for-loop control part).

We leave the in-detail explanation and analysis of the disassembled machine code for the extended version of this report. Here we just pay attention to a slight alteration of the for-loop conditions that results in different compilation results and execution times.

If the index upper limit in the for-loop conditions is a (variable declared as) constant, as in the following for-loop head:

```
for (uint uI = 0; uI < cuiN; ++uI),
```

then the compiler could and would omit the second mov instruction after add (at the address 00C738D3h) and prepare the cmp instruction as

```
cmp dword ptr [ebp-7E8h], 1312D00h.
```

Here, the constant value of cuiN = 1312D00h = 20 000 000h is stored in the instruction itself. This is just an example of how the x86 set of instructions is not *orthogonal*, i.e., how many operations are not allowed with an arbitrary addressing mode.

In our old benchmarks, we wrote the for-loops in just the above way and their loop heads were translated into the following seven (7) instructions, with their lengths in bytes in parenthesis:

- mov(10), jmp(2), mov(6), add(3), mov(6), cmp(10), jae(2), plus jmp(2) at the end of the loop body,
  - totaling eight (8) instructions in 41B.

This is one (1) instruction and 2B less than in our standard benchmarks. However, whether such loops will run faster or not, is still to be determined (§4.3.3). Nevertheless, they do not resemble the general case in which the upper limit can and often will be a non-constant value, so we have abandoned them.

To get back to the present state of our benchmark loops, in Listing 2 we observe that the for-loop in A.3 benchmark is very similar to the previous ones, and the one in B.3 is equivalent. They have the same number of instructions (8 + 1), with the same lengths as the loops in the first two benchmarks. Furthermore, all instructions in the A.3 for-loop are of the same type as those in A.1, with just a few minor differences.

Overall, the differences in manipulating the pointer from manipulating the indices are only subtle. The compilation of all for-loops in our standard benchmarks results in the same number of instructions, with the same operations, and the same or very similar addressing modes, resulting in their same lengths. A thorough analysis confirmed — what could have been expected in the first place — that this holds for the loops of the benchmarks with other five array (data) types as well. From this conclusion, one could expect them to execute within the nearly same time, but whether this is so, we still have to see.

3.2.2 Compilation of the array-element-accessing statements (int arrays)

In the for-loop bodies of all the presented benchmarks, there is just one C/C++ statement. First, we focus on the compilation of this statement in the three benchmarks of type A, in which a *right*-value provided by a variable iRVal is stored into the uI-th array element of the integer array iX.

In the benchmarks A.1 and A.3 for the int array, this is accomplished by three (3) and in the benchmark A.2 by four (4) mov instructions. Therefore, we compare A.1 and A.3 cases first. In A.1, the three mov instructions together are 19B long, and in A.3 (only) 14B. The first two instructions in both benchmarks are not only of the same length but also of the exact same type. In A.1 (A.3) benchmark:

- the first mov instruction places the value of the index uI (pointer pI) into EDX (ECX);
- the second mov instruction, in both A.1 and A.3, places the value of the local variable iRVal = intLrg = 22 222 222d = 0153158Eh, stored at the address A(iRval) = EBP – 738h, into EAX (EDX).

The difference is only in the third mov instruction. In A.1, it moves the value of EAX (= iRVal) to the address of the uI-th array element, which is formed by the index addressing (eq. 1b). In this case:

```
A(iX[uI]) = iX + 4*uI = 00C826A8h + 4*EDX.   (2)
```

It requires a 7B-mov instruction, which is 1B longer than the previous two. Namely, besides the information of the source operand (here EAX), it must store the information of the index register (EDX), the length of the (integer) array element (4) and — as the longest data — the 32-bit address of the iX array.

In A.3, the third mov instruction in the loop body is much simpler and because of that 5B shorter. Namely, the address of the uI-th array element is already prepared in the iP pointer (which was moved into ECX in the first mov instruction). So, the value of EDX is simply moved to the address shown by iP.

To summarize briefly, in that third mov instruction, we note a clear simplification in the benchmark A.3, comparing it to the benchmark A.1. However, having in mind that this 32-bit code will be executed on the 64-bit platform (and particularly, on the 64-bit processor!), it remains to see if the shorter and simpler last mov instruction will bring some speed benefits.

As for the loop body of the benchmark A.2, the four instructions are 6B, 6B, 6B, and 3B long, in total 21B, i.e., one instruction and 2B (7B) more than in A1 (A3).

In B.1, the three mov instructions are of the same type as those in A.1 and have the same lengths, but they are placed in a different order and with different source and destination operands. The first mov places the value of uI (now with A(uI) = EBP – 7F0h) into EAX, the second stores the value of iX[uI] into ECX, and the last one moves the value from ECX into the iLVAl variable, A(iLVal) = EBP – 784h.

The analogous situation is in B.3. Again, the instructions there are of the same type as those in A.3, but permuted. First, the pointer `pI` value is moved to `ECX` [`A(pI) = EBP – 810h`]. The second `mov` retrieves the contents from the address in `ECX`, i.e., the value of `*pI`, and stores it into `EDX`, making this crucial fetch of the array element very effective. Finally, the third `mov` places the `EDX` value into `iLVal`.

Similarly, in B.2, the four `mov` instructions are of the same type as those in A.2. The first two of them do the same thing as those in A.2. The third, the shortest one (3B), stores the value of `*(pI0 + iRVal)` into `ECX`, and the last one moves this value from `ECX` to `iLVal`. Judging solely by the number of instructions and their lengths, this solution is longer but — as suggested previously — its time efficiency may depend on other factors, as well.

As for the benchmarks with the other array (data) types, here we just summarize our findings after the analysis of all six (6) benchmarks for every of the six (6) standard data types. The implementation of the single C/C++ statement within the loop body is more diverse than of the loop control, as we have already shown for the `int` arrays. The observation that the type-2 benchmarks (A.2 and B.2) have one instruction more and the type-3 benchmarks (A.3 and B.3) have the shortest total length of their instructions holds also for the other array data types. Besides that general remark, while the number of instructions for the `short`- and `int`-type arrays is the same, the former has the net length that is systematically 2B longer. These instructions use the 16-bit `X`-type registers (`AX`, `CX`, `DX`). The situation levels up for the `char` type. These benchmarks have the instructions of the same or faintly shorter length than those for the `int` type, and they use the lower byte of the `X` registers (`AL`, `CL`).

The benchmarks with the `llint` arrays have systematically two (2) instructions more than the corresponding ones of `int` type and are because of that considerably longer. In the x86 mode, the VS-CV 2019 compiler produces the x86, 32-bit machine code, without using the available 64-bit (`R`) registers, so that moving this type of data required the engagement of two 32-bit registers instead of one, and the use of the standard `dword` (`double word = 32-bit word`) `mov` instructions. In our concrete example, the value of the `llint` variable `llRVal` is stored into the two registers: `ECX:EDX`, of which the left (right) holds the more (less) significant half of the 64-bit value.

The remaining two types of arrays are of the floating-point type. To access the array elements of the `float` (`double`) type, the compiler uses `movss` (`movsd`) instructions to move the 32(64)-bit contents to the lowest (lower) portion of the 128-bit `xxm` registers (in our case it was the `xmm0` register (Intel 2022-2). The assembly code analysis showed that the numbers of these instructions are equal to the numbers of instructions in the corresponding benchmarks for the `int` arrays and that their length is equal for both the 32-bit `float` and 64-bit `double` `float` type.

3.2.3 A Glimpse to the x64 Compilation

In VS-CV 2019, the default compilation option is for the x86 set of machine instructions, as a still de-facto standard for many sorts of applications. In addition, there is also the x64 compilation, which translates the C++ source code into the Intel's x64 set of 64-bit machine instructions.

In this brief overview, we just summarize that the structure of the `for`-loops in our benchmarks remained the same after the x64 compilation: there are eight (8) instructions in the loop heads and one (1) unconditional jump at the end of the loop bodies. That is, there are nine (9) instructions with the total length of 42B, 1B less than in our standard, x86 version loop. This holds for all `for`-loops with indices. The `for`-loops with the incrementing pointers are organized somewhat differently, so that they have in total ten (10) instructions and a much longer length of 55B.

As for the assignment statement in the loop body, it is compiled to four (4) instructions: `mov`(7), `lea`(7), `mov`(4), `mov`(3), totaling 23B, and having one (1) instruction and 4B more than the x86 version. Here all but the first instruction work with the `R`-type of registers for preparing the operand addresses. As for the assignment statements for the `int`-type benchmarks, their *r*-values and the array elements are manipulated in accordance to their type, that is, as the 32-bit memory and register values. The same instructions, but ordered differently, are in the loop bodies of the B-type benchmarks. The situation is very similar for the shorter (integer) types, for which the loop bodies have one instruction more than in the x86 version. Likewise, for the `float` type, the loop body of A.1 benchmark is realized by four (4) instructions: `mov`(7), `lea`(7), `movss`(6), `movss`(5), totaling 25B, and having one (1) instruction and 2B more than the x86 version. Generally, for the types equal to or shorter than 32-bits, the use of x64 architecture does not improve the structure of the compiled machine code.

The benefits of the 64-bit architecture should — if anywhere — become obvious for the 64-bit data types. Really, the loop body of the `llint` A.1 benchmark is realized by (only) four (4) instructions: `mov`(7), `lea`(7), `mov`(8), `mov`(4), totaling 28B, which is one (1) instruction and 4B less than in the x86 machine code. However, for the type `double` of A.1 benchmark, the instructions in the loop body are: `mov`(7), `lea`(7), `movsd`(9), `movsd`(5), totaling 28B, which is one (1) instruction and 5B more than in the x86 version.

3.2.4 Importance of the Benchmark Analysis

The benchmarks written for this purpose were not intended to perform some standard operations (though the A-type benchmarks do perform the initialization of the array elements to the same value), but to be as simple as possible and thus eliminate all unnecessary consumption of the processor time that is not connected with the purpose of this testing. In such reduction, one must pay attention to the generality of the written program code. Otherwise, it may easily

happen that the obtained benchmarks favor some of the testing options before the others.

For instance, in the early versions of our A-type benchmarks, instead of the `iRVal` variable, we have used a numerical constant as a value to be stored in the array element. This enabled the use of the immediate addressing mode for that operand, which resulted in one `mov` instruction less than in our standard benchmarks (in Listing 2). The numbers and the lengths of those instructions (in parenthesis) are as follows:

A.1 – 2 instr.: `mov`(6), `mov`(11), in total 17B;

A.2 – 3 instr.: `mov`(6), `mov`(6), `mov`(7), in total 19B;

A.3 – 2 instr.: `mov`(6), `mov`(6), in total 12B.

Another example of a pitfall in the benchmark code was already commented in §3.2.1.

In conclusion, when writing benchmarks, one should pay great attention to their generality and follow all the rules of good programming practices. Besides that, before applying the newly created benchmarks, it is good to check their assembly language form.

4 Execution Time Measurements

In this section, we shall briefly present the methods used for our BET (benchmark execution time) measurements and the prerequisites needed to achieve consistent and precise results. Then we present these results and comment on them.

4.1 Time Measurement Methods

There are several ways to measure the elapsed time of certain program parts in C++. In our preliminary and motivational measurements (mentioned in sec. 1), we have used the MS Windows `SYSTEMTIME` `struct`, available in VS IDE after including the `windows.h` header file (details in MS 2022-1, example of application in Listing 3). With its smallest time division being a millisecond (ms), this method is useful when the measured times approach the order of one second (s). This was the case in our early measurements, where we measured the passage through the `short` and `int` arrays with `700 0000h = 117 440 512d` elements, i.e., five times larger than now, and on slower computers. In our present programs, this method is deserted.

More precise time measurements should get more accurate "time stamps" from the hardware timers, which rely directly on the processor's time cycles (MS 2012-2). Such are the methods (ii) and (iii) in Listing 3. In (ii), the `HRTimer` class uses the member function `QueryPerformanceCounter` to do the job. We used this method for the time measurements in (Logozar 2012-1/ 2012-2), overestimating its precision to 10ns. The third time measurement method tested in this work used the `high_resolution_clock` class from the C++ `std::chrono` library (C++ reference, 2022).

In a special, short test C++ program, we have compared the results of the three time-measuring methods on the A-type benchmarks 1, 2, and 3, for the `short` and `int` types. Method (i) gives only roughly correct

Listing 3. Examples of the three time measurement methods in C++, applied to the A.1 benchmark for `int` array: i) `_SYSTEMTIME` structures, ii) `CHRTimer` class, and iii) the `high_resolution_clock` class.

```
// Declarations and definitions as in Listing 1.
//    ...       ...
// Variables for the elapsed times in ms.
float fDltTmSYS, fDltTmHRT, fDltTmHRC;

// i) Time mesurement by SYSTEMTIME (SYS)
#include "windows.h"
// _SYSTEMTIME structures and ptrs. to struc.:
_SYSTEMTIME sT1, sT2, *pST1 = &sT1, *pST2 = &sT2;
GetSystemTime(pST1);       // Stopwatch on.
for (uint uI = 0; uI < cN; uI++)
     iX[uI] = intLrg;
GetSystemTime(pST2);     // Stopwatch off.
fDltTmSYS =
  (float)(pST2->wSecond - pST1->wSecond)*1000 +
  (float)(pST2->wMilliseconds - pST1>wMilliseconds);

// ii) Time mesur. by CHRTimer class (HRT)
#include "HRTimer.h" // High Resolution Time.
CHRTimer hrTimer;   // CHRTimer object.
hrTimer.StartTimer();      // Stopwatch on.
for (uint uI = 0; uI < cN; uI++)
     iX[uI] = intLrg;
fDltTmHRT = (float)hrTimer.StopTimer()*1.e3f;
                          // Stopwatch off.
// iii) By high_resolution_clock class (HRC)
#include <chrono>  // XYZ
using namespace std;
using namespace chrono;
duration<float, milli> durTDlt;
auto t1 = high_resolution_clock::now();
auto t2 = high_resolution_clock::now();
t1 = high_resolution_clock::now();// On.
for (uint uI = 0; uI < cN; uI++)
     iX[uI] = intLrg;
t2 = high_resolution_clock::now();// Off.
fDltTmHRC = (durTDlt = t2 - t1).count();
```

individual results and nearly correct averages, which is not bad regarding its above-stated deficiencies. Methods (ii) and (iii) both give very consistent results of satisfying accuracy for our measurements. Surprisingly, though, their results differ for the order of magnitude of 10μs, which is much more than it should be if they are related to the processor's time cycles (C++ reference, 2022). Anyhow, because the accuracy of this method is more than satisfactory for our case, we have used it for the measurements in this paper.

4.2 Time Measurement Setup

To achieve consistent and accurate time measurements, a programmer should comply with some program- and system-wise conditions. In our approach, we tend to measure the "average best results," i.e., the optimal benchmark execution times for the given computer.

4.2.1 Benchmark Execution Order

In the test programs, we have placed the benchmarks for each of the six array types in an order different from the one shown in Listing 1. In that order, there are also the *pre-runs* (explained in the next section), as follows:

Repeat for the array element access type $k = 1, 2, 3$:

A. Pre-run A.k, then **A.k with BET** measurements;

B. Pre-run B.k, then **B.k with BET** measurements;

End repeat.

In this way, we have simulated a somewhat more realistic situation, in which the A- and B-type benchmarks exchange first, and then there is a change in the type of array element access, according to the enumeration given in §2.2.

4.2.2 Pre-runs

The measured BETs can significantly depend on the momentary state of the memory system of today's computers with multitasking OSs. If the multilevel caches are already optimally filled with the data used in the benchmarks, the execution times will also be optimal, i.e., close to the shortest possible. If this condition is not fulfilled, the measured times can increase severely, making the results prone to erratic changes. To deal with this problem, here we have used the benchmark *pre-runs*. They execute the benchmark fully or partly before measuring its execution time, in the same or very similar way (Logozar 2012-1 and 2). Here, the pre-runs executed the benchmarks by assigning the array elements different number values. Obviously, because the longest BETs will mostly happen at the execution of the first benchmarks, these two approaches produce similar results.

In our C++ programs, the user can control the pre-runs by switching them on and off for the A- and B-types of benchmarks separately, which come one after the other for each array (data) type. The user can also change the number of iterations, i.e., she can specify the number of elements the pre-runs will access.

4.2.3 Computer and OS specifications

To make the BET measurement results complete, it is essential to record the specifications of the computer hardware and OS. In this paper, we mostly present the measurements that are made on one computer with the following specifications:

Processor:	Intel® Core™ i7-8700K CPU @ 3.70 GHz, with 6 cores.
Installed RAM:	32.0GB
System type:	64-bit OS, x64-based processor.
OS:	Windows 10 Pro Education (build 19044.1766).

4.2.4 Measurement Preconditions

Our aim is to measure the execution times of our benchmarks only, as "pure" as possible. To achieve that, we want to eliminate all side effects that could interfere with the computer hardware and OS performance. For that matter, there are a few things to consider.

The first is which program version — debug or release, and in which way — from the IDE or by starting the executable code file — to run. The measurements showed that the influence of both of these options in our VS-CV 2019 is rather minor. The comparison of running the programs directly from the exe-files vs. starting them from the VS-VC19 IDE and leaving the IDE on during the test program executions gives similar results. The former approach does give a bit shorter BETs (0.3% to 1.0%) but its true advantage is greater stability. Namely, the latter approach is prone to sporadic upsurges of BETs up to 10%.[1]

Thus, as a precaution to achieve better precision, our choice for the final measurements was as follows:

- to run the release versions (without compiler optimizations);
- to start the exe-files, with all other applications turned off.

Furthermore, for today's standard multitasking OSs, with the otherwise "standard" execution environment, we take the following precautionary procedure:

- turn off all user applications;
- disable Ethernet and WLAN;
- check the task manager for the possibly demanding background processes and try to turn them off;
- *run the test program with benchmarks*.

4.3 Results of the BET Measurements

With all the preconditions being fulfilled, we run our benchmark test programs. Once the user enters the required input parameters, the remaining free run of one program on our computer lasts approximately 25s. It performs a benchmark pre-run and a BET measurement, normally both through the arrays with 20×10^6 elements, does that for six (6) different benchmarks, repeat the whole action ten (10) times to get the averages and repeat the same thing for six (6) different data types of array. This totals to the $6 \times 10 \times 6 = 360$ pre-run and the same number of measured iterations, resulting in 14.4×10^9 assignment operations on the array elements. If both the A- and B-type benchmark pre-runs are turned off, this number is halved. Each program lists 6×10 BETs for the six array (data) types, with their averages and standard deviations.

4.3.1 BETs of Our Standard Benchmarks

The typical results of the BET measurements for our six standard benchmarks (from Listing 1), extracted from a single run of (test) Program 1, are in Table 1.

For the A- and B-type assignments and the three different methods of accessing the array elements, six average $\overline{\Delta t}_{BET}$ times and (corrected) standard deviations are calculated for the set of 10 non-successive BET measurements for each array type.[2] In the extra columns for the A/B.2 and A/B.3 benchmarks, which

[1] Although we did not practice it, the VS-VC19 IDE can be turned off after launching the desired (debug or release) version of the test program, and before proceeding with the program, i.e., before entering the few required input values.

[2] With 20×10^6 iterations performed in $\approx$ 30ms, one loop iteration takes 1.5ns, or 5.55 clock cycles (CC) of our processor, executing 12 – 13 machine instructions. This means that the working-core's pipeline has an average throughput of 2.25 instructions per CC.

access the array elements by using pointers, we can track the relative difference of their BETs comparing to the A/B.1 BETs.

The measured times and relations between them are similar for the A- and B-type benchmarks marked with the same numbers. A/B.1 is only slightly slower than A/B.3, meaning that the successive access to the array elements by incrementing the pointers did not significantly shorten the execution, although the lengths of their loop bodies are shorter by roughly a quarter. The BETs for A/B.3 benchmarks are shorter than those for A/B.1, but only for the amounts comparable to the measurement standard deviations.

Quite a surprise is that the execution is fastest for the A/B.2 benchmarks (except for the `llint`), although the number of the instructions in their loop body is greater by one-third than those in A/B.1 and 3. In addition, their total length is roughly 10% (30%) longer than in A.3 (B.3). This example clearly shows how the number of instructions and their lengths do not directly dictate their execution times. The decrease is quite significant for the `short` and `int` integers, and for both floating-point types, ranging from roughly 30% to more than 40%. The big exception for the `llint` type (+57%) is caused by the simultaneous lack of the comparable improvement for its A/B.2 benchmarks and the sudden 38% (34%) decrease of its BETs for A.1 & 3 (B.1 &3), resulting in the great raise of the relative values for A/B.2.

For the `char` type, the BETs for A/B.2 benchmarks are just a little faster than for A/B.1 & 3. Furthermore, they are $\approx 45\%$ longer than for all other, longer types, except the `llint`, which is quite surprising. Namely, one would expect the BETs for this data type to be similar to the (other) integers, but not slower.

From above and by looking at the absolute values of those times, we come to the following conclusion.

For our standard benchmarks (Listing 1), the fastest way to access the array elements of type:

- `char`, `short`, `int`, `float`, and `double` is by using **pointer arithmetic** [`*(p0 + i)`], as in A/B.2, (though for `char` it is only slightly better);
- `llint` type is by using
  – **pointer incrementing** [`*(pI)`], as in A/B.3, or just a little bit slower by using
  – **indices** (`Arr[i]`), as in A/B.1.

4.3.2 BETs of the Modified Benchmarks

In our Program 2, we have modified the benchmarks from Listing 1 by replacing the assignment (=) operator with the compound addition-assignment (+=) operator. For these benchmarks, Table 2 contains the average values of their typical BET measurement set.

The times of the A- and B-type benchmarks are still close to each other, but not as much and as consistently as for the set of our standard benchmarks. In addition, there are a few greater discrepancies, as in the following cases: for `int` – cf. A.1 vs. B.1 and A.3 vs. B.3; for `char` – cf. A.2 vs. B.2. A big difference is also that now A.2 and B.2 are not the best access methods as they were for the standard benchmarks (except for `llint`). On the contrary, those are now either worst or close to that, with `float`-B.2 being the only exception. As above, we resume this as follows:

For our modified benchmarks (`'='` → `'+='`), the fastest way to access the array elements of

- all types except `int` is by **pointer incrementing** [`*(pI)`], as in A/B.3, (`int`-A.3 losing over `int`-A.1 for mere 0.1%), or just a little bit worse by using
  – **indices** (`Arr[i]`), as in A/B.1.

An interesting observation is that these benchmarks — which not only assign the *r*-values but also add them to the *l*-values — execute faster than the assignment-only benchmarks in the following cases:

- A/B.1 for `char`, `short`, and B.1 alone also for `int`, where the speed-up is in the range from roughly 20% to more than 30% (the best for `int`-B1);
- A/B.3 for `char` and `short`, and B.3 alone also for `int`, with the speed-up from 20% to 25%.

On the other hand, we see that the execution times for the A.2 benchmarks lag from those of our standard benchmarks (in Table 1) for the first three integer array types, especially so for the `char` and `short`, and then also for the floating-point types. B.2 is better for char, but then greatly lags for the same types as A.2.

In our third test program, the benchmarks were the same as in Listing 1, except that in the A-types the array elements are assigned numerical constants, as it was discussed §3.2.4. The aim was to investigate the influence of this less-then-general case. The table with BETs for these benchmarks are not shown here, but will be just briefly discussed. Since the B-type benchmarks are the same as in Program 1, one can check that their BETs follow those from Table 1 within the standard deviation values. As for the A-type benchmarks, the BETs for A.1 and A.3 are very close to each other and rather short for all integer types. In opposite to that, the times of the integer versions of A.2 are approx. 55% to 72% longer(!) than those of A.1 and A.3, and slightly less so than of the A.2 types of our standard benchmarks. For the floating-point types, the situation is reversed, with A.2 BETs being very short.

Summarily, the shorter loop body greatly improved the performance of the A.1 and A.3 benchmarks with the arrays of the first three integer types. For the A.2 benchmark, the opposite happens for the arrays of the first two integer types.

4.3.3 BETs of Our Older Benchmarks

In the older version of our benchmarks, now marked as Program 3-old, there is another use of a constant: in the upper limit of the conditions in the `for`-loops with running indices (benchmarks A/B.1 & 2). We have discussed that in §3.2.1, showing that this kind of `for`-loop compiled into one instruction and 2B less than the `for`-loop of our standard benchmarks, hinting at the possible speed-ups. To resume, in these A-type benchmarks, there are constants in both the

Table 1. Average $\overline{\Delta t}_{BET}$ (Benchmark Execution Times), for the benchmarks from Listing 1. The benchmarks run in the order stated in §4.2.1, in the release version of the program, on the computer specified in 4.2.3. $\overline{\Delta t}_{BET}$ and the corresponding standard deviations are calculated for 10 (nonconsecutive) measurements and expressed in ms (milliseconds). For the benchmarks A/B.2 and A/B.3, the second column gives the relative Δrel difference of their averages from the benchmark A/B.1 Δt_{BET}. Those are marked in green (red) if $\leq -10\%$ ($\geq +10\%$).

20×10^6 *elements*	*A:* `Arr. El. = r-Value`, $(\overline{\Delta t}_{BET} \pm s_{dev})$/ms					*B:* `l-Value = Arr. El.`, $(\overline{\Delta t}_{BET} \pm s_{dev})$/ms				
Array Type	A.1 `Arr[i]`	A.2 `*(p0+i)`	Δrel to A.1	A.3 `*(++p)`	Δrel *to* A.1	B.1 `Arr[i]`	B.2 `*(p0+i)`	Δrel *to* B.1	B.3 `*(++p)`	Δrel *to* B.1
`char`	35.50 ±0.45	33.68 ±0.43	−5.1%	34.90 ±0.47	−1.7%	36.96 ±0.32	35.36 ±0.37	−4.3%	35.95 ±0.29	−2.7%
`short`	35.45 ±0.45	23.89 ±0.43	−32.6%	35.05 ±0.48	−1.1%	36.78 ±0.36	22.63 ±0.26	−38.5%	35.90 ±0.26	−2.4%
`int`	38.45 ±0.50	22.65 ±0.34	−41.1%	37.81 ±0.34	−1.7%	36.78 ±0.53	23.28 ±0.23	−36.7%	35.96 ±0.32	−2.2%
`llint`	23.47 ±0.38	36.92 ±0.41	+57.3%	23.34 ±0.22	−0.6%	24.16 ±0.22	34.42 ±0.47	+42.4%	23.84 ±0.48	−1.4%
`float`	39.02 ±0.48	22.42 ±0.28	−42.5%	37.85 ±0.32	−3.0%	36.88 ±0.37	23.42 ±0.31	−36.5%	36.09 ±0.40	−2.1%
`double`	38.93 ±0.28	23.04 ±0.15	−40.8%	38.28 ±0.52	−1.7%	37.35 ±0.20	25.46 ±0.36	−31.8%	36.20 ±0.08	−3.1%

Table 2. A set of the BET measurements for slightly altered benchmarks (Prog. 2): in the benchmarks from Listing 1, the assignment operator (`=`) is replaced by the compound addition-assignment operator (`+=`).

20×10^6 *elements*	*A:* `Arr. El. = r-Value`, $(\overline{\Delta t}_{BET} \pm s_{dev})$/ms					*B:* `l-Value = Arr. El.`, $(\overline{\Delta t}_{BET} \pm s_{dev})$/ms				
Array Type	A.1 `Arr[i]`	A.2 `*(p0+i)`	Δrel to A.1	A.3 `*(++p)`	Δrel *to* A.1	B.1 `Arr[i]`	B.2 `*(p0+i)`	Δrel *to* B.1	B.3 `*(++p)`	Δrel *to* B.1
`char`	28.02 ±0.58	37.24 ±0.32	+32.9%	27.64 ±0.41	−1.4%	27.74 ±0.29	29.58 ±0.39	+6.7%	27.17 ±0.32	−2.1%
`short`	28.12 ±0.36	36.21 ±0.63	+28.8%	27.26 ±0.27	−3.1%	27.92 ±0.24	34.19 ±0.30	+22.5%	27.22 ±0.29	−2.5%
`int`	38.62 ±0.34	40.43 ±0.40	+4.7%	38.67 ±0.39	+0.1%	24.78 ±0.29	35.15 ±0.15	+41.8%	24.17 ±0.10	−2.5%
`llint`	35.21 ±0.38	36.01 ±0.40	+2.3%	35.02 ±0.17	−0.5%	30.28 ±0.29	35.00 ±0.35	+15.6%	29.44 ±0.23	−2.8%
`float`	39.07 ±0.66	40.95 ±0.61	+4.8%	38.49 ±0.65	−1.5%	40.53 ±0.35	39.55 ±0.74	−2.4%	39.36 ±0.37	−2.9%
`double`	39.61 ±0.82	41.55 ±0.86	+4.9%	39.14 ±0.80	−1.2%	40.83 ±0.79	40.16 ±1.08	−1.7%	39.61 ±0.58	−3.0%

array-element assignment statements and in the `for`-loop conditions, while in B.1 & 2, only the latter applies. In fact, in the early version of the program, there were only A.1 and A.3 benchmarks, for only the `short` and `int` types, but we have upgraded it to include also all the other benchmarks and array types as in the present test programs.

In Table 3, we compare the BETs of these benchmarks to the corresponding execution times for our standard benchmarks. Despite the expectations to see only improvements, we can see all three cases: decreases, invariabilities, and increases in the execution times, which do not necessarily follow the decrease in the number and length of the machine instructions. For instance, besides the usual unexpected behavior of the `llint` array elements in A(B).1 benchmarks, with the delay of 41% (27%), we also see the great decreases of BETs of A/B.2 benchmarks for the first two integers and both floating-point types. On the other hand, there are also quite significant improvements in A.3 for the first three integer types, etc.

By observing solely the old benchmark's execution times, it can be deduced that the A.3 benchmark BETs for *all* integer types (that is, now *including* the `llint` type, which has otherwise been a notorious exception) are from around 30% up to 35% shorter than for A.1s! This result roughly confirms our early, much less systematic findings, obtained on the same (old) benchmarks, but only on the arrays of the `short` and `int` types, and mentioned in sec. 1.

Table 3. Relative differences of the BETs for our *old* benchmarks (as in our preliminary research), now in Program 3-old, to the BETs of our standard benchmarks (from Table 1).

$\overline{\Delta t}_{BET,rel}$ *(Prog. 3-old / Prog. 1)/ %*						
Arr. typ.	A.1	A.2	A.3	B.1	B.2	B.3
`char`	−4.7%	−29.3%	−36.9%	−13.6%	−2.1%	+0.6%
`short`	−5.0%	+37.3%	−38.0%	−13.3%	+47.2%	+0.1%
`int`	−12.2%	+46.1%	−36.9%	−12.6%	+43.6%	0.0%
`llint`	+40.6%	−14.2%	−0.3%	+26.9%	−10.3%	−0.6%
`float`	−14.6%	+68.0%	−0.2%	−13.7%	+42.6%	0.2%
`double`	−16.4%	+55.1%	−0.5%	−13.1%	+34.2%	+1.2%

Table 4. Relative differences of the BETs for our standard benchmarks (Program 1) compiled to the x64 machine code from the BETs of the same benchmarks compiled as x86 code, in Table 1.

$\overline{\Delta t}_{BET,rel}$ *(Prog. 1 x64 / Prog. 1 x86)/%*						
Arr. typ.	A.1	A.2	A.3	B.1	B.2	B.3
`char`	−11.9%	+23.1%	+32.9%	−14.8%	−6.2%	−11.0%
`short`	−12.1%	+70.1%	+23.5%	−14.4%	+46.9%	−10.8%
`int`	−17.7%	+68.6%	+1.0%	−14.0%	+44.8%	−10.8%
`llint`	+35.9%	−2.3%	+53.0%	+34.9%	−0.2%	+37.4%
`float`	−18.2%	+72.4%	+2.9%	−14.8%	+44.8%	−11.4%
`double`	−18.2%	+62.2%	−5.4%	−14.3%	+33.4%	−9.2%

4.3.4 A Brief Overview of the BETs With x64 Compilation and on Other Computers

The 64-bit, x64 machine code did not result in much faster executions (cf. §3.2.3). By inspecting Table 4, we see the noticeable improvements from our standard benchmarks (Prog. 1) only for the A/B.1 and B.3 BETs, again except for the `llint` array type. Here, this bad behavior of the 64-bit integer type is particularly odd, because one would expect that the x64 compilation would improve at least the executions for the compatible, 64-bit data types. In addition, the A.3, and even more A/B.2 benchmarks show much worse behavior than the corresponding x86 code.

In Program 2, the 64-bit compilation produces faster execution than the 32-bit compilation for all A-benchmarks, but not for the first two integer types (not shown in a comparison table). The improvements are in the range from around 10% up to 25%. The B-type benchmarks' BETs are from 10% shorter to 20% longer. In Program 3, A.1 (B.1) benchmarks' BETs show slow-downs of about 40% (40% to 60%!) for all integer types, and diverse behavior for the other benchmarks and array types.

Finally, for our old benchmarks (Program 3-old), the x64 machine code shows quite uniform BETs: for the A.1 benchmarks they are in the approximate range from 31ms to 34ms, for A.2 from 32ms to 37ms, and for all the B-type benchmarks from 32ms to 34ms. Comparing them to the BETs of the corresponding x86 versions, it turns out they are significantly faster (from 10% to 12%) only for the B.3 benchmarks (except for the `llint` type!). Furthermore, since the BETs for the A.3 benchmarks are in this case similar to the other values, and in the x86 version those were much better, their relative lags are quite large: for `char` 55%, `short` 53%, `int` 40%, and `llint` 38%. Because of that, in the x64 version of the benchmarks, the A.3-type of access is no longer significantly faster than A.1, as it was in their x86 version.

We have repeated the same measurements on a few other computers with the same, Wintel platform. One of them has Intel® Core™ i3-6100U CPU @ 2.30GHz, with (only) two cores and relatively low energy consumption, and is running on the Windows 10 OS. It executes our benchmarks for roughly twice the time of the much more powerful computer described in §4.2.3. The BETs obtained on it follow the trends described in the previous text for our all test programs, with a few smaller discrepancies for certain benchmarks and array types. It also resembles the results of our old benchmarks, i.e., show the decrease of the BETs when using the pointer incrementing (A.3) over the use of the array indices (A.1), though slightly less than in our case: 25% for `char`, 32% for `short` and `int`, and 15% for `llint`.

The 64-bit compilations on this computer behave similarly to those on our primary computer, i.e., there are minor speed-ups in some cases, but also slow-downs in the other. In addition, regarding the Program 3-old, there are no improvements in A.3 over A.1 benchmarks, same as on the primary computer.

5 Conclusion

The first intention of this research seemed to be quite simple: to investigate whether one can improve the time efficiency of a program code by replacing the standard, index-based access to 1D-array elements, with access to them via pointers. However, this paper shows that the answer to this question is neither unique nor simple, and even less easy to explain.

Here we upgraded the benchmarks from our ad-hoc measurements from a decade ago, in which the array-element assignment statements were kept as simple as possible and had only two types of accesses: via indices (No. 1), and via the incremented pointer (now No. 3). In their `for`-loop bodies, there was a single assignment statement with the array elements as the *l*-values. From the arrays of the `short` and `int` types only, these benchmarks were expanded to all six standard numerical array (data) types). Then, we included also the benchmarks with the mirror symmetrical assignments, in which the array elements are the *r*-values. Thus, the two types of benchmarks were formed and named A and B (Listing 1). Both of them are simple and because of that their compilation with the optimization turned

on can produce very simplified machine code, particularly the B type. We have discussed this matter in detail in §3.1 and justified running and investigating the non-optimized machine code, especially if analyzing its disassembly in assembly language. These benchmarks can be regarded as the starting ones upon which the future methodology will be built. In the rest of section 3, we have explained the machine instructions of the compiled code and discussed a few pitfalls that we had encountered in this research and had to deal with. In doing that, we could notice the consequences of the non-orthogonality of Intel's x86 instruction set, in the sense that not all addressing modes are allowed with all operations, which can result in changes in how the source code can and will be compiled (§3.2).

However, despite all the analysis, explaining and understanding many varieties and peculiarities of the benchmark execution times (BETs) in section 4 was everything but an easy task. That is, not only that Intel's x86 instructions are of variable length—which complicates the analysis of their execution times—but we have shown that these times significantly depend on the broader context within which these instructions appear. Such as the data type of the array elements and the use of special operations and the corresponding (special) registers.

By looking at the results of this research, there is no simple conclusion and no unique best choice. The best method largely depends on the type of the assignment/operation in the loop body and on the array (data) type. We have managed to repeat the results of the early measurements on our old benchmarks and showed an even greater advantage of the use of the incrementing pointers over indices than before. However, the analysis showed that these benchmarks are rather non-general because they have constants in the crucial parts of their loops. Besides that, the advantage occurs for only the first three integer types (Table 3).

For our standard benchmarks, with the assignment-only statements in the loop bodies, we have given the concrete advice based on the execution times in Table 1, summarizing them in the (bulleted) list in §4.3.1. For the modified benchmarks, with addition and assignment, the conclusion based on the results in Table 2 is summarized in the list in §4.3.2. Interpreting these results the other way around, we could say that using the indices can be as good as anything else if we primarily do iterative summations and if a few percentages better performance by incrementing the pointer is irrelevant. Alternatively, we could say that the access via indices is considerably outperformed by the use of the pointer arithmetic in the case of the assignment-only operations, but not forgetting that the `llint` type is an exception. Etc.

Though, one conclusion is rather simple: the x64 code does not improve the benchmark performance significantly, not even for the 64-bit types. Quite the contrary, from Table 4 we see that the execution times for most cases are worse than for the x86 code.

From all this we can derive a conclusion of the presented subject on the standard Wintel platform: the best approach—especially in critical applications—is to measure the execution times of the intensively used piece of code as we did, and to choose the access method with the shortest execution times.

This conclusion, and especially the last piece of advice, also hints us at a few suggestions for future work. First, the additional benchmarks should be written in a way that could not be trivialized by the compiler optimizations, despite the fact that this would prevent easy disassembly within the VS-CV 2019 IDE. Some of them could perform the standard array and matrix operations. Second, it would be interesting to make these measurements on some other platforms.

6 References

C++ reference (2022). `std::chrono` library, `high_resolution_clock` class. Web contents: https://en.cppreference.com/w/cpp/chrono/high_resolution_clock.

Evans, D (2022). *x86 Assembly Guide*. Web contents: https://www.cs.virginia.edu/~evans/cs216/guides/x86.html.

Intel (2022-1). *View disassembly code in the Visual Studio debugger.* Web contents: https://docs.microsoft.com/en-us/visualstudio/debugger/how-to-use-the-disassembly-window?view=vs-2022

Intel (2022-2). *Intel® 64 and IA-32 Architectures Software Developer Manuals*. Web contents: https://www.intel.com/content/www/us/en/developer/articles/technical/intel-sdm.html.

Kernighan, B. W., & Ritchie D. M. (1978, 1988). *The C Programming Language (1st and 2nd Editions, respectively), Englewood Cliffs, NJ: Prentice Hall.*

Logozar, R. (2012-1). Recursive and Nonrecursive Traversal Algorithms for Dynamically Created Binary Trees. *Computer Technology and Application (David Publishing)*, 3**(**5**)**, 374-382.

Logozar, R. (2012-2). Algorithms and Data Structures for the Modeling of Dynamical Systems by Means of Stochastic Finite Automata. *Technical Gazette* (*Tehnički vjesnik*), 19(2), 227-242.

Microsoft, MS (2022-1). Documentation: SYSTEMTIME structure (miniwinbase.h). Web contents: https://docs.microsoft.com/en-us/windows/win32/api/minwinbase/ns-minwinbase-systemtime

Microsoft, MS (2022-2). Documentation: Acquiring high-resolution time stamps. Web contents: https://docs. microsoft.com/en-us/windows/win32/sysinfo/acquiring-high-resolution-time-stamps?redirectedfrom=MSDN#resolution-precision-accuracy-and-stability

Stroustrup, B. (1997), *The C++ Programming Language (3rd Edition).* Murray Hill, New Jersey: AT&T Labs.

Author Guidelines

Microsoft Word Template for the Central European Conference on Information and Intelligent Systems (CECIIS)

Author 1, Author 2
Author's Common Affiliation
Department/Institute
Full Address
{author1, author2}@domain.com

Author(s) Name(s)
Author Affiliation(s)
Department/Institute
Full Address(es)
E-mail(s)

Abstract. *The abstract is to be in fully-justified italicized text, at the top of the left-hand column as presented here, below the author information. Use the word "Abstract." in 10-point Times, boldface type, left positioned, initially capitalized, followed by the abstract in 10-point, single-spaced type, up to 100 words long.*
Leave one blank line after the abstract, and then begin the keywords. Use the word "Keywords." in 10-point Times, boldface type, left positioned, initially capitalized, followed by up to ten keywords in 10-point, separated by comma, as below.
Leave two blank lines after the keywords, and then begin the main text.

Keywords. CECIIS, conference paper, template

1 Introduction

This is an example of a Central European Conference on Information and Intelligent Systems (CECIIS) submission. These guidelines include complete description of the paper style including formatting, fonts, spacing, and related information for producing your proceedings manuscripts.

This template can be used to create your own paper but if you have any troubles or if in doubt do not hesitate to contact us via e-mail: ceciisoo@foi.hr or via the conference web site at http://www.ceciis.foi.hr.

2 Structure and Requirements for Research Paper (Examples)

Central European Conference on Information and Intelligent Systems (CECIIS) welcomes research papers on all topics of interest (Data and Knowledge Bases, Economics and Information Systems, Education for Information Society, Information and Communication Technologies, Information Systems Security, Intelligent Information Systems, Intelligent Transport Systems and Autonomous Vehicles, Software Engineering, Quality of Software and Services) as well as on the Conference theme/special emphasis announced on yearly base, addressing both, research theory and practice.

The following paper structure examples regarding the research type could be used for preparing a CECIIS paper:

1. Quantitative questionnaire-based: background research (state of the art); research problem; hypothesis or research question; description of instrument used; sampling; statistical methods used for analysis; discussion of results with limitations of research.
2. Qualitative research: background research (state of the art); research problem; research question; qualitative method used (case study research, interviews, content analysis etc.); discussion of results with limitations of research.
3. Algorithms and theoretical research: theory description with references; original theoretical contribution (theorem, algorithm etc.) with proofs or confirmations; future research and conclusion
4. Systematic literature review with meta-analysis: scope of research, research questions or research purpose; sources of literature with period covered (data bases, journals etc.), methods of analysis, conclusions with future research

Research paper should be theoretically and evidence based, clearly and consistently as well as ethically prepared and written. Central European Conference on Information and Intelligent Systems (CECIIS) has zero-tolerance policy on any kind of plagiarism. Paper containing text copied and used from another author(s) work without permission or from authors previously published own work, without referencing, will be retracted.

3 Paper Organization and Formatting

All papers should be written in English, up to 8 pages long (approximately 20 000 characters), and arranged in the following order:

- Main title
- Author(s), affiliation(s), full and e-mail addresses
- Abstract
- Keywords
- Body text (Main text)
- Footnotes
- Acknowledgements
- References

All printed material, including text and figures, must be kept within a print area of 16 cm (6.3") wide by 24.7 (9.7") high. Do not write or print anything outside the print area. Paper size is A4: 21x29.7 cm (8.3x11.7"). Margins (top, bottom, left, right) are 2.5 cm (1"). All *text* must be in a two-column format. Columns are to be 7.7 cm (3") wide, with a 0.6 cm (0.24") space between them. Text must be fully justified. Indent each paragraph by 0.5 cm (0.2").

The final submission has to be submitted in a single PDF file with all fonts that are used embedded, and also the source file (Microsoft Word, LaTeX - zip original files and images).

4 Main Title

The main title (on the first page) should begin 3,5 cm from the top edge of the page, centred, and in Times 18-point, bold face. Capitalize the first letter of nouns, pronouns, verbs, adjectives, and adverbs; do not capitalize articles, coordinate conjunctions, or prepositions (unless the title begins with such a word). Leave a blank line after the title.

5 Author Name(s) and Affiliation(s)

Author names and affiliations are to be centred beneath the title and printed in Times 10-point type. Author names should be in bold face. Multiple authors shall be grouped by affiliation as shown in the title above. Include also e-mail addresses in 9-point `Courier/Typewriter` face. Follow the author information by two blank lines before main text.

6 Type-style and Fonts

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times that you have access to. Please avoid using bit-mapped fonts if possible. True-Type 1 fonts are preferred.

Also make sure that all fonts in the final PDF file are embedded. Thus we strongly encourage the use of Latex especially if you use non-standard fonts for formulas for your examples. You can make sure that your fonts are embedded by opening the PDF file in some PDF viewing program and taking a look at the properties of the document.

7 First-order Headings (May Extend to the Next Line)

For example, “1 Introduction”, should be Times 14-point boldface, initially capitalized, flush left, with two blank lines before, and one blank line after.

Don't use periods (“.”) after the heading number, use them only in lower order headings to separate them from higher order headings. For *long* headings use *a hanging indent* aligning the text to the right of the heading number as shown above.

All paragraphs in the text should be indented except for the first paragraph in a section.

7.1 Second-order Headings (May Extend to the Next Line)

As in this heading, they should be Times 12-point boldface, initially capitalized, flush left, with one blank line before, and one after. Use a hanging indent for long headings.

7.1.1 Third-order Headings (May Extend to the Next Line)

Third-order headings, as in this paragraph, are discouraged. However, if you must use them, use 10-point Times, boldface, initially capitalized, flush left, with one blank line before, and one after. Use a hanging indent for long headings.

8 Main Text

Type your main text in 10-point Times, single-spaced. Do not use double-spacing. All paragraphs should be indented 0.5 cm (0.2") except for the first paragraph in a section.

Be sure your text is fully justified—that is, flush left and flush right. Please do not place any additional blank lines between paragraphs.

8.1 Figures

All figures are to be included within the text. **Figure captions** are to be *below* the figures, in 10-point Times

(or a similar serif font), normal face. Initially capitalize only the first word of each figure caption.

Figures are to be numbered consecutively with Arabic numerals throughout the paper, for example: "Figure 1. Database contexts", and are referred to in the text as Fig. 1, Fig. 2, etc.

8.1.1 Illustrations, Photographs and Graphs

Illustrations, photographs and graphs are considered as figures. All graphics should be centred. Your artwork must be in place in the article (preferably printed as part of the text rather than pasted up). If you are using photographs and are able to have halftones made at a print shop, use a 100- or 110-line screen. Supply the best quality photographs and illustrations possible.

Pencilled lines and very fine lines do not reproduce well. Remember, the quality of the book cannot be better than the originals provided. The conference proceedings will be printed in grey scale, but you can provide images in colour since the papers will also be published on the web site of the conference.

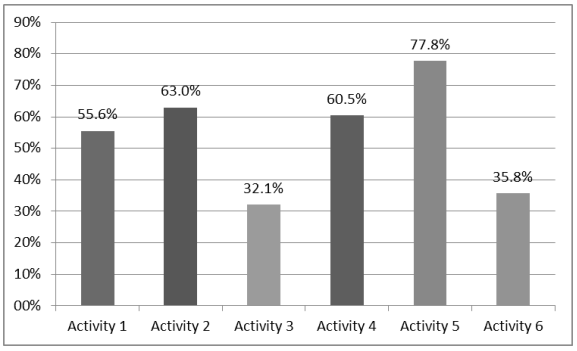


Figure 1. Times, 10 points, normal face

8.2 Tables

All tables are to be included within the text. **Table titles** are to be *above* the tables, in 10-point Times (or a similar serif font), normal face. Initially capitalize only the first word of each table title. Tables are to be numbered consecutively with Arabic numerals throughout the paper, for example: "Table 1. Input data", and are referred to in the text as Table 1, Table 2, etc. Column headings should be as brief as possible.

Table 1. Times, 10 points, normal face

	Column 1	Column 2	Column 3
Row 1	1	2	3
Row 2	4	5	6
Row 3	3	4	7
Row 4	2	6	4

[1] Use Times 8-point type, single-spaced. Avoid using footnotes altogether and include necessary peripheral observations in the text (within parentheses, if you prefer, as in this sentence).

8.3 Equations

Equations should be numbered serially on the right-hand side by Arabic numerals in parentheses, and referred to in the text by eq. 1, eq. 2 etc.

$$3x^2 + 3x - 23 = 0 \quad (1)$$

8.4 Footnotes

Use footnotes sparingly (or not at all!) and place them at the bottom of the column on the page on which they are referenced.[1]

9 In-text Citations and Referencing

In-text citation must be used to denote all text used from another author(s) work or from authors previously published own work. When referenced in the text, enclose the author surname(s) and the publication year in brackets, for example (Surname, year) or (Surname, year, p. x) where x is the page number of the cited reference. When referenced paper has more than three authors, cite the paper as (Surname et al., year).

List and number all bibliographical references, alphabetically sorted, in 10-point Times, single-spaced, with a hanging indent, at the end of your paper. Use 6-point paragraph spacing after each reference. References should be distributed evenly in both columns as much as possible, so use a column break where appropriate.

In the References section below you will find examples of common types of bibliographical items: book (Mayer, 2009, p. 52), journal paper (Bule & Peer, 2014), journal paper with more than three authors (Shukor et al., 2014), conference proceedings paper (Steingartner & Novitzká, 2015), report or standard (*ISO/IEC 25010: Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models*, 2011) and web page ("Web Content Accessibility Guidelines (WCAG) 2.0," 2008). For other types of bibliographical items refer the **APA style guidelines**. It should be ensured that every reference cited in the text is also listed in the References section (and vice versa).

Acknowledgments

Acknowledgments, if necessary, should appear in a separate paragraph preceding the references.

References

Bule, J., & Peer, P. (2014). Technical, Legal, Economic and Social Aspects of Biometrics for Cloud Computing. *Journal of Information and Organizational Sciences*, 38(2), 83–95.

ISO/IEC 25010: Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models. (2011).

Mayer, R. E. (2009). *Multimedia Learning* (Second Edi.). New York: Cambridge University Press.

Shukor, N. A., Tasir, Z., Van der Meijden, H., & Harun, J. (2014). A Predictive Model to Evaluate Students' Cognitive Engagement in Online Learning. *Procedia - Social and Behavioral Sciences*, 116, 4844–4853. doi:10.1016/j.sbspro.2014.01.1036

Steingartner, W., & Novitzká, V. (2015). A new approach to operational semantics by categories. In T. Hunjak, V. Kirinić, & M. Konecki (Eds.), *Proceedings of the 26th Central European Conference on Information and Intelligent System (CECIIS 2015)* (pp. 247–254). University of Zagreb, Faculty of Organization and Informatics Varaždin.

Web Content Accessibility Guidelines (WCAG) 2.0. (2008). Retrieved from https://www.w3.org/TR/WCAG20/

CECIIS 2022 Sponsors

CECIIS 2022 GOLD SPONSORS

INFODOM


MOBILISIS®


ORACLE

CECIIS 2022 SILVER SPONSOR



CECIIS 2022 BRONZE SPONSORS





CECIIS 2022 Financial Support



REPUBLIC OF CROATIA
Ministry of Science and Education

CECIIS 2022 Program Committee

Neven Vrček, Program Committee Chair, University of Zagreb, Croatia

Victoria Appatova, University of Cincinnati, USA
Igor Balaban, University of Zagreb, Croatia
Nina Begičević Ređep, University of Zagreb, Croatia
Marko Bohanec, Jožef Stefan Institute, Slovenia
Leo Budin, University of Zagreb, Croatia
Tonči Carić, University of Zagreb, Croatia
Carlos Carrascosa, Universitat Politècnica de València, Spain
Anna Czarczyńska, Kozminski University, Poland
Luis de Marcos Ortega, University of Alcalá, Spain
Blaženka Divjak, University of Zagreb, Croatia
Jasminka Dobša, University of Zagreb, Croatia
Matjaž Gams, Jožef Stefan Institute, Slovenia
Gordan Gledec, University of Zagreb, Croatia
Andrina Granić, University of Split, Croatia
Lourdes Guàrdia, Open University of Catalonia, Spain
Marjan Heričko, University of Maribor, Slovenia
Dieter Hertweck, Herman Hollerith Research Centre, Germany
Edouard Ivanjko, University of Zagreb, Croatia
Dimitris Karagiannis, University of Vienna, Austria
Valentina Kirinić, University of Zagreb, Croatia
Marina Klačmer Čalopa, University of Zagreb, Croatia
Michal Kohani, University of Zilina, Slovakia
Mario Konecki, University of Zagreb, Croatia
Melita Kozina, University of Zagreb, Croatia
Marjan Krašna, University of Maribor, Slovenia
Alen Lovrenčić, University of Zagreb, Croatia
Sandra Lovrenčić, University of Zagreb, Croatia
Ivan Luković, University of Belgrade, Serbia
Ivan Magdalenić, University of Zagreb, Croatia
Aleksandar Marković, University of Belgrade, Serbia
Marta Mas Machuca, UIC Barcelona, Spain
Dunja Mladenić, Jožef Stefan Institute, Slovenia
Oliver Moravčik, Slovak University of Technology, Slovakia
Enrique Mu, Carlow University, USA
Karlo Nyarko, University of Osijek, Croatia
Jan Paralič, Technical University of Košice, Slovakia
Elisabeth Pergler, University of Applied sciences, Austria
Wolf Rauch, University of Graz, Austria
Sonja Ristić, University of Novi Sad, Serbia
Markus Schatten, University of Zagreb, Croatia
Zlatko Stapić, University of Zagreb, Croatia
William Steingartner, Technical University of Košice, Slovakia
Nebojša Stojčić, University of Dubrovnik, Croatia
Vjeran Strahonja, University of Zagreb, Croatia
Diana Šimić, University of Zagreb, Croatia
Katarina Tomičić-Pupek, University of Zagreb, Croatia
Jan Tretmans, Radboud University Nijmegen, The Netherlands
Wim Van Petegem, KU Lueven, Belgium
Violeta Vidaček Hainš, University of Zagreb, Croatia
Sirje Virkus, Tallinn University, Estonia
Mladen Vouk, North Carolina State University, USA
Ksenija Vuković, University of Zagreb, Croatia
Ivica Crnković, Chalmers University of Technology, Sweden

Honorary Members - Former Program Committee Chairs

Boris Aurer, University of Zagreb, Croatia
Tihomir Hunjak, University of Zagreb, Croatia
Mirko Maleković, University of Zagreb, Croatia
Miroslav Žugaj, University of Zagreb, Croatia

CECIIS 2022 Organizing Committee

Petra Grd, Organizing Committee Chair, University of Zagreb, Croatia

Josipa Bađari, University of Zagreb, Croatia
Ena Barčić, University of Zagreb, Croatia
Ivana Fojs, University of Zagreb, Croatia
Darko Grabar, University of Zagreb, Croatia
Vika Kerekeš, University of Zagreb, Croatia
Valentina Kirinić, University of Zagreb, Croatia
Marina Klačmer Čalopa, University of Zagreb, Croatia
Amalija Koren Cavaliero, University of Zagreb, Croatia
Bernarda Kos, University of Zagreb, Croatia
Izabela Oletić Tušek, University of Zagreb, Croatia
Katarina Pažur Aničić, University of Zagreb, Croatia
Laura Strupar, University of Zagreb, Croatia
Matija Šajn, University of Zagreb, Croatia
Lana Škvorc, University of Zagreb, Croatia
Petra Vondra, University of Zagreb, Croatia

CECIIS 2022 Research Tracks' Chairs

Computer Games: Markus Schatten, University of Zagreb, Croatia

Data and Knowledge Bases: Sandra Lovrenčić, University of Zagreb, Croatia and Patrizia Poščić, University of Rijeka, Croatia

Digital Response to Pandemic: Neven Vrček, University of Zagreb, Croatia

Digital Transformation and Corporate Culture: Anna Czarczyńska, Kozminski University, Poland and Katarina Tomičić Pupek, University of Zagreb, Croatia

Education and Learning Analytics: Blaženka Divjak, University of Zagreb, Croatia, Igor Balaban, University of Zagreb, Croatia

ICT Entrepreneurship and Innovation: Ksenija Vuković, University of Zagreb, Croatia, Marina Klačmer Čalopa University of Zagreb, Croatia and Nebojša Stojčić, University of Dubrovnik, Croatia

Intelligent Energy: Danijel Topić and Marinko Barukčić, University of Osijek, Croatia

Intelligent Systems: Jasminka Dobša, University of Zagreb, Croatia and Dunja Mladenić, Jožef Stefan Institute, Slovenia

Quality in ICT: Valentina Kirinić and Melita Kozina, University of Zagreb, Croatia

Strategic Planning and Decision Making: Nina Begičević Ređep and Tihomir Hunjak, University of Zagreb, Croatia

Systems for Smart Mobility: Tonči Carić, Edouard Ivanjko, University of Zagreb, Croatia

Trends in Software Engineering: Zlatko Stapić, University of Zagreb, Croatia, William Steingartner, Technical University of Košice, Slovakia and Luis de Marcos Ortega, University of Alcalá, Spain

CECIIS 2022 Poster Section Chairs

Neven Vrček, University of Zagreb, Croatia

Petra Grd, University of Zagreb, Croatia

CECIIS 2022 Student Online Presentation Section Chairs

Violeta Vidaček Hainš, University of Zagreb, Croatia

Victoria Appatova, University of Cincinnati, USA



ISSN 1847-2001 (Print)
ISSN 1848-2295 (Online)