UNIVERSITY "ST. KLIMENT OHRIDSKI" - BITOLA
FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGIES - BITOLA
REPUBLIC OF NORTH MACEDONIA

Proceedings of the 13th International Conference on Applied Internet and Information Technologies

AIIT 2023

13 October, 2023, Bitola, Republic of North Macedonia

AIIT
International Conference

UNIVERSITY "ST. KLIMENT OHRIDSKI" - BITOLA
FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGIES - BITOLA
REPUBLIC OF NORTH MACEDONIA

13TH INTERNATIONAL CONFERENCE ON APPLIED INTERNET AND INFORMATION TECHNOLOGIES

AIIT 2023
PROCEEDINGS

Bitola, 2023

Proceedings publisher and organizer of the conference:
University “St. Kliment Ohridski” - Bitola, Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia

For publisher:
Blagoj Ristevski, PhD
Dean of Faculty of Information and Communication Technologies - Bitola University “St. Kliment Ohridski” - Bitola, North Macedonia

Proceedings editors:
Kostandina Veljanovska, PhD
Eleonora Brtka, PhD

Conference Chairmans:
Kostandina Veljanovska, University “St. Kliment Ohridski”, Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia - chair
Eleonora Brtka, University of Novi Sad, Technical Faculty “Mihajlo Pupin”, Zrenjanin, Serbia – co-chair

Technical preparation of the proceedings:
Andrijana Bocevska, PhD
Kostandina Veljanovska, PhD

Cover design:
Andrijana Bocevska, PhD

e-Proceedings
ISBN 978-608-5003-03-7

Disclaimer:

CIP - Каталогизација во публикација
Национална и универзитетска библиотека "Св. Климент Охридски", Скопје

004-049.8(062)

INTERNATIONAL conference on applied internet and information technologies AIIT 2023 (13 ; 2023 ; Bitola, Republic of North Macedonia)
Proceedings / 13th International conference on applied internet and information technologies AIIT 2023, 13 October, 2023, Bitola, Republic of North Macedonia ; [editors Kostandina Veljanovska, Eleonora Brtka]. - Bitola : University "St. Kliment Ohridski", Bitola Faculty of information and communication technologies, 2023. - 411 стр. : илустр. ; 30 см
Библиографија кон трудовите
ISBN 978-608-5003-03-7

а) Информатичка технологија -- Примена -- Собири
COBISS.MK-ID 62559493

Introduction

The International conference on Applied Internet and Information Technologies is a traditional meeting held every year, that sprouts out of collaboration between the University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia and the University "St. Kliment Ohridski", Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia. The XIII AIIT2023 was held in Bitola, Macedonia on which besides the participants from Serbia and Macedonia there were researchers from Croatia, Bosnia and Herzegovina, Hungary, Finland, Russia, Turkey, Egypt, India and Australia whose contribution was either as authors or as reviewers of the papers.

At the Conference were presented innovative findings in the field of information systems, communications and computer networks, software engineering and applications, data science and big data technologies, artificial intelligence, intelligent systems, business intelligence and IT support to decision-making, data and system security, distributed systems, Internet of Things and smart systems, embedded systems, computer graphics, IT management, e-commerce, e-government, e-education, Internet marketing, and IT practice and experience.

The Conference chairs would like to express gratitude to the authors for their contributions and to express special gratitude to the reviewers for their tremendous work done for selecting the papers with their valuable comments and suggestions that contributed to improve the quality of the papers. Out of more than 60 submitted papers, 51 were selected, presented at the Conference and are published in this proceedings.

The work during the conference was organized in nine sessions: plenary session, five in-person oral sessions, one video session and two poster sessions. During the conference, a round table with participants from academic organizations and IT industry was successfully organized. The theme of the discussions at the round table was "Strengthening the capacities of Faculty of ICT for the realization of strategic cooperation with companies from the IT industry".

AIIT 2023 was very successful conference with fruitful exchange of experiences among the participants reviving the hope of further strengthening a friendly environment after the pandemic crisis. We hope that we will continue with the contribution to the further deepening the development of Internet and information technologies research.

Conference chairs:

Kostandina Veljanovska, University "St. Kliment Ohridski", Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia (chair)
Eleonora Brtka, University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia (co-chair)

MAIN ORGANIZERS

Faculty of Information and Communication Technologies - Bitola
University "St. Kliment Ohridski" - Bitola
NORTH MACEDONIA
http://www.fikt.edu.mk/

Technical Faculty "Mihajlo Pupin" Zrenjanin
University of Novi Sad
SERBIA
http://www.tfzr.uns.ac.rs/

ORGANIZATION PARTNERS

Faculty of Computer Science
Irkutsk National Research Technical University
Institute of Informational Technologies and Data Analysis
Irkutsk, RUSSIA
http://www.istu.edu/

Matrosov Institute for System Dynamics and Control Theory of Siberian Branch of Russian Academy of Sciences
Irkutsk, RUSSIA
http://www.idstu.irk.ru/

Irkutsk State Transport University (IrGUPS)
Irkutsk, RUSSIA
https://www.irgups.ru/

Faculty of Engineering South-west university "Neophyte Rilsky"
Blagoevgrad, BULGARIA
https://www.irgups.ru/

Conference Chairs

Kostandina Veljanovska, University "St. Kliment Ohridski", Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia

Kostandina Veljanovska, Ph.D. finished BSc in Computer Science at the University "Sts. Kiril i Metodi", Skopje. Her first MASc in Applied Engineering she received at the University of Toronto, Toronto, Canada. Her second MSc and also her PhD in Technical Sciences she received at the University "St. Kliment Ohridski" - Bitola, R. Macedonia. Her postdoctoral studies in Artificial Intelligence she attended at the Laboratory of Informatics, Robotics and Microelectronics at the University of Montpellier, Montpellier, France. She worked as a Research assistant at the Faculty of Applied Science, University of Toronto, Canada. She also, worked as a researcher in research team for Constraints, Learning and Agents at LIRMM, University of Montpellier. Since 2008, she works as a Full Professor in Information Systems and Networks, Artificial Intelligence and Systems and Data Processing at the Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, Rebublic of North Macedonia. Her research work is focused on artificial intelligence, machine learning techniques and intelligent systems. She has published numerous scientific papers in the area of interest, as well as several monographic items. She is a reviewing referee for well-known publishing house, journals with significant impact factor in science and also, member of editorial board of several international conferences.

Eleonora Brtka, University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia

Eleonora Brtka, Ph.D. is an associate professor at the Information Technology department at the University of Novi Sad, Technical Faculty „Mihajlo Pupin", Zrenjanin. She has received her PhD in Information technology in 2015. Current research interests include: data science techniques and methods, programming languages, education, artificial intelligence and intelligent agent technologies. She is the author or co-author of several articles published in international journals and in the proceedings of international conferences. She participated in several national funded projects.

Organizing Committee

Chairs

Kostandina Veljanovska (chair), Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia
Eleonora Brtka, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia (co-chair)
Zoran Kotevski, Faculty of Information and Communication Technologies - Bitola, Republic of North Macedonia (co-chair)

Members

Željko Stojanov, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia
Blagoj Ristevski, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Višnja Ognjenović, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia
Dalibor Dobrilovic, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia
Monika Markovska, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Dragica Radosav, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Biljana Radulovic, Technical Faculty "Mihajlo Pupin" Zrenjanin, Serbia
Božidar Milenkovski, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Ivana Berković, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Ljubica Kazi, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Snežana Savoska, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Vladimir Brtka, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Zoltan Kazi, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Siniša Mihajlović, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Velibor Premčevski, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Nikola Rendevski , Faculty of Information and Communication Technologies - Bitola, North Macedonia
Aleksandra Stojkov, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Maja Gaborov, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Andrijana Bocevska, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Milica Mazalica, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Igor Vecštejn, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Tome Dimovski, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Marko Blažić, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Vuk Amižić, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Natasa Blazeska-Tabakovska, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Jovana Borovina, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Dalibor Šeljmeši, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Mimoza Bogdanoska-Jovanovska, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Vladimir Šinik, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Nadežda Ljubojev, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Marina Blaženković Toshevski, Faculty of Information and Communication Technologies - Bitola, North Macedonia
Vladimir Karuović, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Evgeny Cherkashin, Institute of System Dynamic and Control Theory SB RAS, Russia
Anastasia Popova, Matrosov Institute for System Dynamics and Control Theory of the Siberian Branch of the Russian Academy of Sciences, Russia
Filip Tsvetanov, South-west University "Neophyte Rilsky", Faculty of Engineering, Blagoevgrad, Bulgaria

Program Committee

Blagoj Ristevski (president), Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Eleonora Brtka, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia (vice-president)
Željko Stojanov, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Višnja Ognjenović, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Dalibor Dobrilović, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Ljubica Kazi, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Dragica Radosav, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Dragana Glušac, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Borislav Odadžić, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Miodrag Ivković, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Biljana Radulović, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Ivana Berković, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Vladimir Brtka, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Zoltan Kazi, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Jelena Stojanov, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Vesna Makitan, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Nadežda Ljubojev, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Vladimir Šinik, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia
Igor Nedelkovski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Aleksandar Markoski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Violeta Manevska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Pece Mitrevski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Ilija Jolevski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Dragan Gruevski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Kostandina Veljanovska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Monika Markovska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Snežana Savoska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Ramona Markoska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Sonja Mančevska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Mimoza Bogdanoska-Jovanovska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Nataša Blaženska Tabakovska, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Božidar Milenkovski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Zoran Kotevski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Nikola Rendevski, Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" - Bitola, North Macedonia
Andrijana Bocevska, Faculty of Information and Communication Technologies, University "St. Kliment

C O N T E N T

The Latest Trends in IT Project Management

Vesna Makitan [1]

[1]*University of Novi Sad, Technical Faculty "Mihajlo Pupin" Djure Djakovica bb, Zrenjanin, Republic of Serbia*

vesna.makitan@tfzr.rs

Abstract:
Considering the fact that when we talk about business, we talk about the realisation of projects, projects have become an integral part of every enterprise. Also, regardless of the area of application, every project today shares common characteristics with IT projects. Therefore, IT projects also represent one of the mainstream topics among academic circles.
The landscape of IT project management is evolving in response to new technologies such as the Internet of Things, Artificial Intelligence, Cloud Computing, etc. This paper addresses the ways these technologies have influenced and improved existing methodologies and tools in the field of IT project management, enabling them to adapt to these new possibilities. It also provides a review of research studies which highlight the latest trends in the field of IT project management.
Moreover, this paper offers insights into the real situation in project realisation, based on the author's own experiences.
The aforementioned may contribute to the improvement of the implementation of future projects.
Keywords:
IT project management, Industry 4.0 technologies

1. Introduction

Although the beginning of the development of a scientific approach to project management is linked to the 1940s and the Manhattan Project [1] (implemented by the US Army for the purpose of building an atomic bomb in World War II), people have been implementing and managing projects throughout history.

Project management includes continually evolving methodologies and tools whose application plays a pivotal role in ensuring project success. Appropriate project management methodologies are applied depending on factors such as technology, requirements and people, and project dimensions. In addition, the development of technologies has brought about changes to project management, enriching its capabilities to adapt to the demands of modern business. [2] The development of technology has also meant that modern projects exhibit the characteristics of IT projects, irrespective of the domain in which they are implemented. Therefore, in a broader sense, project management can be considered as a form of IT project management.

Given that, as members of the academic community, we participate in project realisation, it is important to follow current trends in this area. Therefore, this paper will present research which highlights the latest trends in the field of IT project management. The selected studies explore trends in project management from various aspects, such as: the technologies and components of Industry 4.0 and their impact on project management and project success, the possibility of choosing a singular environment for integrated project management, a comparison of the latest versions of the PMBOK standard and the characteristics of project managers in the era of industry 4.0.

Furthermore, it is important to have a perspective on our position in relation to global trends. To this end, based on the author's personal experience, a description of the current situation in the realization of IT projects in the academic and business environment is provided.

2. Research on the latest trends in IT project management

There are numerous studies dealing with the trends in IT project management. The research [3] was conducted with the aim of examining the extent to which Industry 4.0 technologies, Cloud Computing, the Internet of Things and Artificial Intelligence in particular, are applied in the IT sector in Serbia, as well as the degree to which project managers and technical team members believe that these technologies contribute to a project's success. It was found that those who use the aforementioned technologies assess their projects as generally more successful. The potential exists for technologies to be applied in the field of project management, but it has not been fully utilized in the Serbian IT sector. Given that the importance of technologies is acknowledged by project managers and technical team members, increased usage is expected to lead to greater success of IT projects. Table 1 and Figure 1 illustrate the use of these technologies.

Table 1:
Overview of the use of modern technologies [4]

Technologies	Average grade of technologies implementation	The most common answer	Chronbach's alpha
Cloud technologies	3.9		
Cloud software for any project management activities	3.6	5	
Cloud document management systems	2.5	1	
Cloud storage systems	4.3	5	0.707
Cloud tools for project communication and collaboration	4.7	5	
Online shared calendars	4.4	5	
AI technologies	2.3		
Tools for project selection for realisation or portfolio	2.4	1	
Tools for project monitoring and risk assessment	3.1	4	
Tools for assessing project feasibility based on historical data	1.9	1	
Intelligent software agents	1.4	1	
Tools for selecting team members	1.5	1	
Project realisation support tools	3.5	4	0.762
Tools for the automatic generation of reports or documentation	2.9	3	
Software which independently performs repetitive tasks	2.5	1	
Project planning tools	1.8	1	
Decision support tools based on historical data	1.6	1	
IoT technologies	2.4		
Smart phones	4.6	5	
Smart watches	2.1	1	
Smart tables	1.9	1	
Smart cameras	1.6	1	0.665
Tools for managing and monitoring connected smart devices	1.6	1	

Previous research findings were confirmed by the research [5], which found that "The crucial success factor is the extensive application of Industry 4.0 components to support the actions carried out as part of a project." The researchers [5] adopted a set of 4.0 Industry components: real-time data management, interoperability, virtualisation, decentralisation, agility, service orientation and integrated business process [6 in 5] for their study. The study showed that two components: data management and virtualisation, play a key role and their presence increases the likelihood of project success. "The greater the synergy between individual components of Industry 4.0, the more likely the project is to succeed." [5]

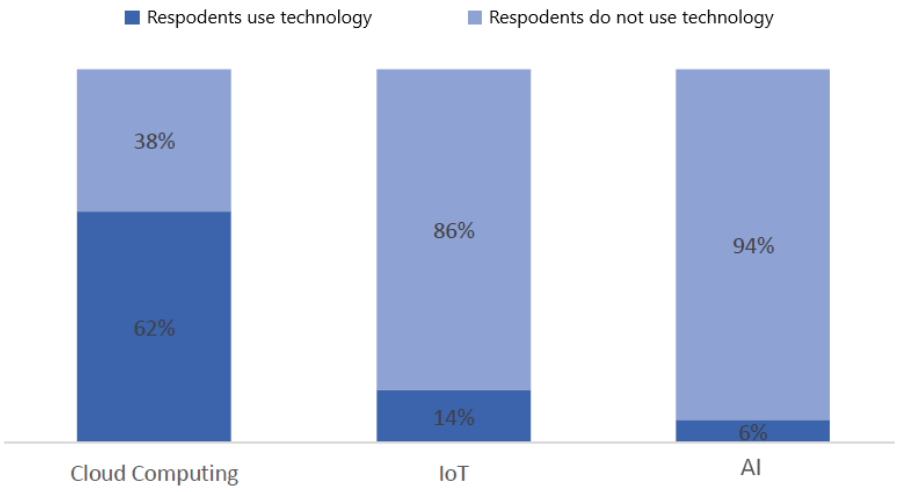


Figure 1: Percentage of respondents using modern technologies. [4]

Researchers in [7] suggested the singular environment for integrated project management. They underlined that "According to KPMG et al. [8] in a study about the global perspective of the future of project management, only 19% of organizations achieve successful projects, just 30% can deliver projects on time, 36% of organizations meet the budget, and 46% of projects are delivered according to stakeholders' expectations." This is reminiscent of the situation in 1995 and the results of the "CHAOS Report", the research carried out by the Standish Group [9]. According to that report, in 1995, the success rate of IT projects was only 16.2%, which means that only these projects were implemented within the projected time frame and budget. Also, 31% of IT projects were cancelled prior to completion, which, together with projects which exceeded the expected time, scope, budget, etc. represented a cost of 140 billion dollars. [9] "These findings show that the project management area still faces some challenges in terms of achieving time, cost, and stakeholder satisfaction, among others. Thereby, to remain relevant, it is critical to discover methods and tools suitably appropriate and up to date, allowing these professionals to keep current with today's trends and new realities [10]." The researchers in [7] proposed the OpenProject software solution as an environment which supports the most features in the eight performance domains of the PMBOK standard [7]. They also suggested future trends in project management: sustainability, agility, requirements management, risk management and benefits management.

The research [11] compares the PMBOK 6th and 7th editions because "the coexistence of these two perspectives was initially an unclear subject". [11] They analysed the PMBOK 6th and PMBOK 7th editions and managed to integrate the knowledge of both versions through correlations between the Techniques and Tools used in the PM Processes of the PMBOK 6th and the Models, Methods and Artifacts of the PMBOK 7th editions. [11]

As its name suggests, the research [12] explores project manager competencies within the context of Industry 4.0. The study highlights the changes in project managers' soft and hard skills. In Industry 4.0 the soft skills vital for project managers are mainly related to the new ways of interacting with project stakeholders, encompassing communication skills, authority, team management, management of unforeseen events and negotiation skills. [13 in 12] "The most important hard skill for project managers is experience with innovative technologies and projects, big data analysis and predictive

algorithms that will help them to manage projects correctly and focused on the objectives to be achieved." [13 in 12] They also underscored that in terms of project manager competencies, the future of project management will be influenced by Industry 4.0 and technological improvements which will change the course of how project management tasks are executed and monitored in the future [14 in 12]. "Information in Industry 4.0 is expected to flow more quickly and dynamically. Work teams will be diverse and will comprise very different areas of study. CPS and IoT will allow the development and use of faster and more predictive management tools. The robotisation of the industry will increase its efficiency, but it can also remove the human factor in certain circumstances. The use of Big Data tools will allow the flow of a massive amount of information, quickly, and with a wide spectrum (5G) [15 in 12]." [12]

3. IT projects in higher education

In terms of the implementation of projects in higher education in the Republic of Serbia, the Ministry of Science, Education and Technological Development, as well as the Provincial Secretariat for Higher Education, have been issuing tenders for projects for several decades. This provides an opportunity for higher education institutions to submit their research proposals and compete for funds to complete their projects.

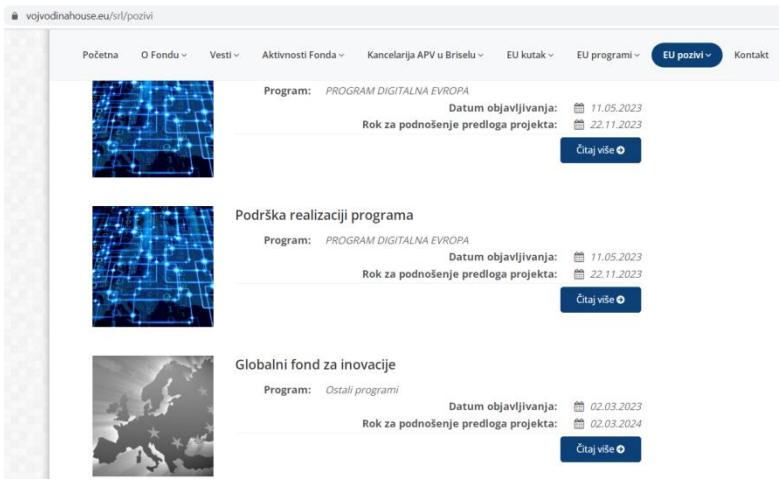


Figure 2: European Affairs Fund of the Autonomous Province of Vojvodina. [16]

In addition, there are also calls for participation in European Union projects of various types including cross-border cooperation (CBC), and educational and research programmes such as Tempus, Erasmus and Horizon, to which higher education institutions can apply. They are promoted by various institutions/sites within the Republic of Serbia. Figures 2-4 present some examples , with the last one best illustrating the extent of the opportunities to win funding. At the time of visiting the site, there were 630 active calls.

Each of the aforementioned calls has specific requirements which must be met. They include, among other things: project objectives, implementers (a consortium – usually made up of various higher education institutions from within the country and abroad, as well as governmental and other organisations which can contribute to the realisation of the project or even companies), project activities with estimated timeframes, human and material resources, as well as project costs. Once the funds have been won for the project implementation, the main indicators of its success are the results achieved within the predicted timeframe and budget. There are no special limitations regarding the project management methodology or the technologies used for the project realisation. The sole guideline for project deliverables is a project report (submitted quarterly, in most cases) with accompanying documentation (receipts for paid invoices, project procurement documents, documentation on suppliers/collaborators engaged in the project, project promotion materials, project budget expenditure records, etc.).

Figure 3: The European Commission's framework programmes in Serbia. [17]

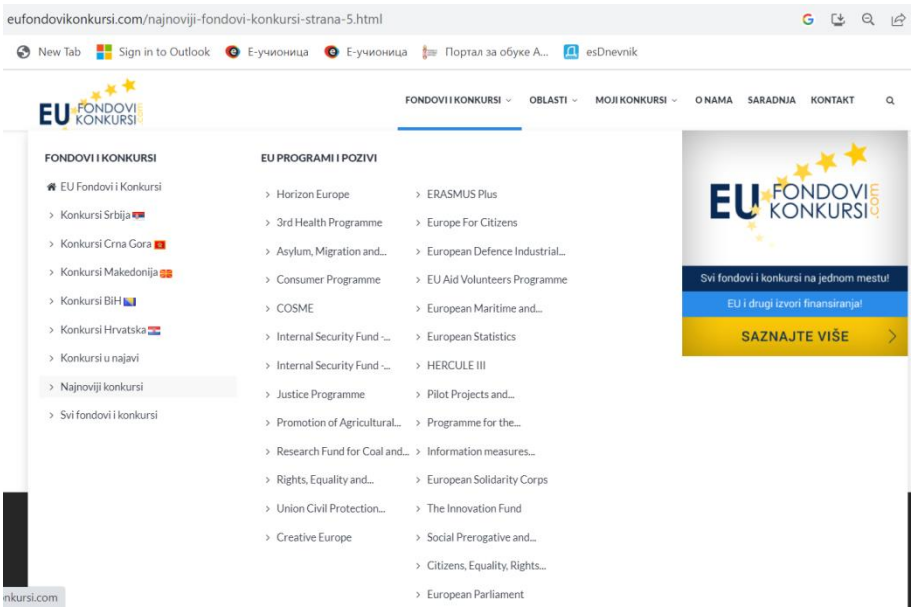


Figure 4: EU funds and calls. [18]

Based on my personal experience, I can say that a few decades ago, the situation was such that it was necessary to introduce training projects for applying for various types of projects, because it was established that the funds allocated by the EU to the countries of the Western Balkans were insufficiently used. [19] At that time, the University of Novi Sad opened a project management office to support its members/faculties in applying for European projects. One of the areas of engagement of this office was the organisation of training/seminars focused on project implementation and

responding to calls for various project types. Subsequently, the situation shifted in favour of institutions in the Republic of Serbia, enabling them to apply and secure funds for the implementation of various types of projects. Our institution used the opportunity to participate in such projects, thus contributing to personal improvement and development.

When considering the technology, requirements and people as key dimensions, Figure 5 best illustrates the situations in which the appropriate methodology is applied (the third dimension is people – the project implementers, whose characteristics may influence the level of uncertainty of implementation). [20] In cases where the requirements and technology are known, traditional project management methodologies are the best choice. With an increase in the uncertainty of technologies and requirements, agile methodologies are available, while for scenarios with the highest values of these dimensions, extreme methodologies are recommended [21].

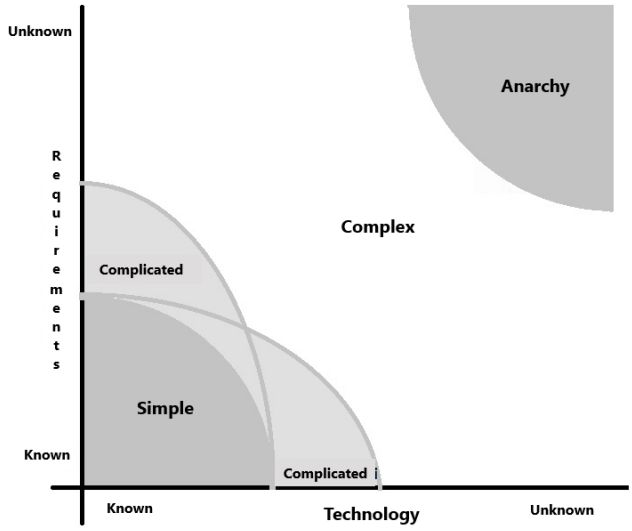


Figure 5: The complexity of projects depending on technology and requirements. [20]

It is interesting to note that the projects in higher education institutions are realised according to the traditional methodology supported by the PMBOK standard, because the most important issue is to adhere to the deadlines, the budget and the planned deliverables. This applies to all projects, regardless of type. Naturally, the choice of technologies used for project delivery is left to the discretion of the project managers and the members of the project teams. It is also interesting that the teams which participate in the implementation of projects exhibit the characteristics of teams in agile methodologies: they are self-organising and the project manager loses the role she/he has in traditional project management.

What characterises the current situation in applying for project calls mirrors that we had a few decades ago: there is a shortage of personnel willing to participate in this process, and training sessions for project applications are being organised again. For example, the City Administration of Zrenjanin, where it has been determined that training programmes are essential, is currently undertaking an initiative for the development of a Science strategy for young people. This action has been supported by the Regional Center for Socio-Economic Development – Banat. [22]

If we recall the findings from the research results [3], the introduction of new technologies in IT project management is taking place more slowly in our country. Moreover, it is necessary to update existing knowledge and involve new people in the process of applying for project calls and IT project management.

The technical faculty “Mihajlo Pupin” is actively involved in all these processes, thus contributing to the development of IT project management.

4. Conclusions

This paper presented the latest trends in IT project management based on current research dealing with the topic. The selected studies facilitated insights into:

- The extent to which Industry 4.0 technologies, especially Cloud Computing, the Internet of Things and Artificial Intelligence, are applied in the IT sector in Serbia;
- The impact of selected components of industry 4.0, especially data management and virtualisation, on project management;
- The singular environment for integrated project management;
- A comparison of the PMBOK 6th and 7th editions;
- Project Manager Competencies in the context of Industry 4.0.

The paper also showed some of the experiences in IT project realisation in higher education institutions. It may be concluded that:

- IT project management is greatly influenced by technological development;
- IT project management exhibits iterative development;
- The methodology used in IT project management is comprised of numerous different methodologies.

Further research may incorporate these findings in order to establish the situation "in the field" so as to identify those factors which contribute most to project success, as well as recommendations for good IT project management practices.

Acknowledgment:
This research is financially supported by the Ministry of Education and Science of the Republic of Serbia under project number TR32044 "The development of software tools for business process analysis and improvement", 2011-2023.

References:

[1] K. Schwalbe, Information Technology Project Management, Revised. 8th ed. Boston, MA: Cengage Learning, 2016.
[2] V. Makitan, M. Ivković, E. Brtka, V. Brtka, New Technologies Impact on the Modern Project Management, Konferencija YUInfo 2019, Kopaonik, 10.03.2019. – 13.03.2019., Društvo za informacione sisteme i računarske mreže: Zbornik radova, Miodrag Ivković (editor), Beograd, 2019.
[3] T. Zorić, V. Makitan, E. Brtka, S. Mrđen, Modern technologies influence on the project success factors in IT sector of Serbia, Technical Faculty "Mihajlo Pupin", Zrenjanin: Journal of Engineering Management and Competitiveness (JEMC), 11(2), pp.124-137. 2021.
[4] T. Zorić, Examination of the impact of modern technologies on the project success factors in the Serbian IT sector, master thesis, Author's reprint, Technical Faculty "Mihajlo Pupin", Zrenjanin, 2020.
[5] L. Kanskia, J. Pizon, The impact of selected components of industry 4.0 on project management, Elsevier: Journal of Innovation & Knowledge, 8, 2023.
[6] C. Salkin, M. Oner, A. Ustundag, E. Cevikcan, A conceptual framework for industry 4.0. (pp. 3−23). Cham: Springer. https://doi.org/10.1007/978-3-319-57870-5_1. 2018.
[7] I. Reis, P. Ribeiro. A Singular Environment: Featuring a Framework for Integrated Project Management. Procedia Computer Science 219, pp. 2019–2026, 2023.
[8] KPMG, AIPM, IPMA. The future of project management: Global Outlook 2019. 2019.
[9] The Standish Group International, The Chaos Report, 1995. URL: https://www.csus.edu/indiv/v/velianitis/161/chaosreport.pdf
[10] S. Wawak, K. Woźniak, Evolution of project management studies in the XXI century. IJMPB 13(4), pp. 867–88, 2020.
[11] F. Amaro, L. Domingues, PMBOK 6th meets 7th: How to link both guides in order to support project tailoring? Procedia Computer Science 219, pp. 1877–1884, 2023.
[12] A. Ribeiro, A. Amaral, T. Barros, Project Manager Competencies in the context of the Industry 4.0, Procedia Computer Science 181, pp. 803–810, 2021.

[13] T. Win, S. Kham, Transformation of Project Management in Industry 4.0. Proceedings of 12th International Conference on Project Management, Bucharest, Romania, November 1st-2nd, 2018.
[14] M. Cakmakci, Interaction in project management approach within industry 4.0. In Lecture Notes in Mechanical Engineering, pp. 176-189. https://doi.org/10.1007/978-3-030-18715-6_15, 2019.
[15] Z. Whysall, M. Owtram, S. Brittain, The new talent management challenges of Industry 4.0. Journal of Management Development, 38(3). https://doi.org/10.1108/JMD-06-2018-0181, 2019.
[16] Фонд Европски послови Аутономне покрајине Војводине, Отворени ЕУ позиви, 2023. URL: https://vojvodinahouse.eu/src/pozivi
[17] EU u Srbiji, EU programi, 2023. URL: https://europa.rs/programi/
[18] EU fondovi i konkursi, Fondovi i konkursi Srbija, 2023. URL: https://www.eufondovikonkursi.com/
[19] Project "European Expert Training for Accession EXTRA 2", TEMPUS, no.: CM_SCM-C026B06-2006, V. Makitan team member, 2007-2008.
[20] K. S. Rubin, Essential Scrum: A Practical Guide to the Most Popular Agile Process, Addison-Wesley, 2012.
[21] R. K. Wysocki, Effective project management – Traditional, Agile, Extreme. John Wiley and Sons, inc. USA, 2014.
[22] Banat Regionalni centar za društveno-ekonomski razvoj, Projekti, 2023. URL: https://banat.rs/

Virtual Reality: The Gateway to Next-Generation Skill Development and Talent Attraction

Nikola Rendevski[1,2], *Blagoj Ristevski*[1]

[1,] *University "St Kliment Ohridski"- Bitola, 1 Maj bb, 7000 Bitola, RN Macedonia*
[2,] *Cythero, Todor Skalovski 10, 7000 Bitola, R. North Macedonia*

nikola.rendevski@uklo.edu.mk, blagoj.ristevski@uklo.edu.mk

Abstract:

This paper explores the transformative role of virtual reality (VR) as a gateway to the future of skill development and talent attraction. As the technological landscape continues to evolve rapidly, traditional approaches to learning and recruitment are being redefined. VR emerges as a disruptive force, offering immersive and interactive experiences that transcend the limitations of conventional training and recruitment methods. We are examining the current challenges in skill development and talent acquisition, emphasizing the need for innovative solutions to bridge the gap between education and industry demands. It then goes into the capabilities of virtual reality technologies, showcasing their potential to revolutionize skill acquisition by providing realistic and context-rich learning environments. Through simulations and interactive scenarios, VR not only enhances traditional training methods but also enables the acquisition of complex skills in a risk-free and adaptive manner. The discussion presents facts from various industries to illustrate the successful implementation of VR in skill development and talent attraction.

Keywords:
VR training, gamification, skills shortage, talent attraction

1. Introduction

VR's roots date back to the mid-20th century (Figure 1). However, it took until the end of the 20th century, with the development of rudimentary head-mounted displays and interactive simulations, for VR to take concrete form. Over the past few decades, VR technology has grown exponentially, driven by advancements in hardware capabilities, improvements in software, and a deeper understanding of the psychological aspects of immersion. The journey in VR development has been an amazing one, but it hasn't been without its challenges. High cost, limited access, and the need for powerful computing hardware were the initial barriers to widespread adoption. Additionally, the quality of early VR experiences was often not good enough to create truly engaging and immersive virtual worlds. Overcoming these hurdles required a collaborative effort by researchers, developers, and industry pioneers.

An important change in the VR environment is the transition from PC-based VR (PCVR) systems to standalone VR devices [1]. Traditionally, VR experiences have been associated with high-powered computers, limiting accessibility and practicality. The advent of standalone VR, featuring standalone devices with integrated computing power, frees users from relying on fixed setups. This move democratized VR, making it more accessible to a wider audience and opening the door to new applications beyond gaming and entertainment.

As we consider the transformative potential of VR [3] in skill development and talent acquisition, it is important to recognize the complex evolutionary history that has brought us here. This article explores the challenges being overcome and the opportunities presented as VR evolves, and explores the new ways in which immersive experiences can change and

redefine how individuals acquire skills and organizations interact with top talent. We will focus on the transition to standalone VR as a catalyst for the times.

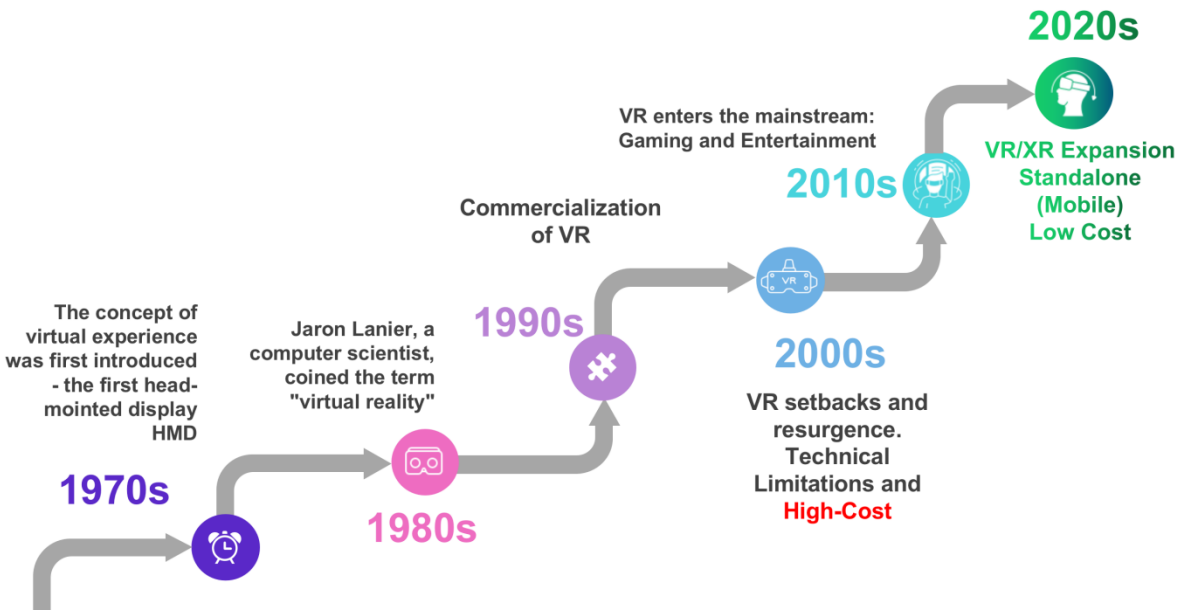


Figure 1. History of VR Technology Development

2. Transformation from PCVR to Standalone VR

VR technology has undergone a remarkable evolution, marked by a technological transformation from connected PC-based VR systems to the revolutionary era of standalone VR. As we undertake this research, it is important to understand the technical complexities underlying this transformative change [2][3].

In the early stages of VR development, users were tied to powerful PCs and required cables and external devices to achieve a seamless VR experience. The demand for high-end computing power and specialized equipment has limited access to VR to a select few and limited its use to specific fields. However, dynamic advances in hardware and software technology have ushered in a new era for VR, and standalone VR devices have proven to be a game-changer. Moving from PCVR to standalone VR is more than just a change in form factor. This represents a fundamental shift away from dependence on external computing resources, with standalone VR devices integrating powerful processors, graphics capabilities, and sensors in a compact format, freeing users from the limitations of physical connectivity. I'll release it. This change will not only improve mobility, but also make VR more accessible to a wider audience, democratizing it.

Technological advances, such as the development of efficient mobile processors, have played a key role in making standalone VR a reality. These processors are designed to deliver robust computing performance in compact devices, enabling VR experiences without sacrificing quality. Additionally, advances in sensor technology such as inside-out tracking and gesture recognition are increasing the immersive and interactive nature of standalone VR, further blurring the lines between the virtual and physical worlds.

3. Gamification

Gamification, a concept based on the integration of game elements and mechanics into non-game environments, is a transformative force in the field of human interaction and motivation. Originating from the ancient use of games and challenges in early civilizations to teach skills, gamification has evolved significantly, especially with the advent of digital technology.
Since ancient times games have been an integral part of human culture serving not only as a source of entertainment but also as a powerful tool for education and skill development. Early civilizations recognized the educational value of games and challenges and used them as effective means of teaching basic skills and knowledge.

The digital age ushered in a new dimension of gamification, marked by the rise of serious gaming in the late 20th century. These serious games were notable for serving purposes beyond entertainment, such as training, simulation, and behavior modification. In pursuit of more effective learning methods, serious games seamlessly integrate game mechanics to increase engagement and make learning new skills and knowledge a more dynamic and interactive process.

Serious games represented a paradigm shift in using game elements for purposes beyond traditional entertainment. By combining game design principles with educational and behavioral goals, serious games have provided a powerful combination that engages users while promoting learning and skill development. Incorporating game mechanics such as points, levels, and rewards adds a level of immersion to these experiences, creating an environment that fosters cognitive engagement and improved memory retention. The field of gamification is constantly evolving with continuous research and innovation aimed at finding new ways to promote engagement, motivation, and behavior change in a variety of situations. This development goes beyond traditional gaming platforms and permeates many aspects of daily life, from education and training to health and professional development.

In today's context, gamification is a dynamic field that provides a versatile toolkit for designers and educators looking to create experiences that attract, motivate, and encourage desired behaviors. As technology advances, the potential applications of gamification are expanding, opening new frontiers for leveraging the inherent appeal of games to achieve meaningful outcomes in non-gaming contexts.
Gamification in VR training [4][5] is a powerful and innovative approach to improving the learning experience. By integrating gaming elements into the training process, companies/organizations can make learning more engaging, immersive, and effective.

Here we consider important aspects of gamification for training in VR:

1. Immersive learning environment:

VR offers unique advantages by creating immersive environments that recreate real-world scenarios. Gamification goes a step further and adds interactive elements such as challenges, quests, and simulations to make the learning experience more engaging and memorable.

2. Motivation and commitment:

Games are designed to be entertaining by their nature. By implementing game mechanics like points, rewards, and levels into the VR training, participants are more likely to stay motivated and engaged. The sense of accomplishment from overcoming a challenge or reaching a higher level contributes to a positive learning experience.

3. Scenario-based learning:

VR gamification allows creation of realistic scenarios that for example employees might experience in their roles. This approach allows learners to apply knowledge to practical

situations and improve problem-solving and decision-making skills. VR simulation can recreate complex situations that are difficult to simulate using traditional training methods like classroom or e-learning.

4. Progress tracking and feedback:

Games provide instant feedback on performance and help learners and trainees recognize their strengths and opportunities for improvement. VR training with gamification features can track progress, provide real-time feedback, and provide personalized learning paths. This not only contributes to continuous improvement, but also promotes a sense of accomplishment.

5. Collaborative learning:

Many games encourage collaboration and teamwork. Applying these principles to VR training encourages participants to work together to achieve a common goal. This collaborative learning environment can improve communication skills, teamwork, and the ability to solve problems as a group.

6. Adaptive learning path:

VR gamification enables adaptive learning experiences. The system can dynamically adjust task difficulty based on learner performance, ensuring training is both challenging and achievable. This personalized approach maximizes the efficiency of the learning process.

7. Behavior change and skill transfer:

The immersive nature of VR combined with gamification increases the potential for behavioral change and skill transfer. Learners can practice and reinforce desired behaviors in a risk-free environment, improving retention and application of skills in real-world scenarios.

8. Competitive factors:

Incorporating an element of competition through leaderboards and timed challenges can encourage healthy competition among learners. This not only brings more fun, but also motivates individuals to perform better than others, leading to a sense of accomplishment.

9. Cost-effective training:

Although the initial investment in VR technology can be sometimes significant for many organizations, gamified VR training can be a cost-effective solution in the long run. It reduces the need for physical resources, travel, and on-site training, making it a scalable and efficient training method.

4. VR training efficiency

To the best of our knowledge, the most recent study on the efficiency of VR training in comparison with traditional learning methods is the PwC report from 2020 [6]. The summary of this extensive study and the most important learning efficiency parameters are presented in Figure 2 and the following:

Training speed: Employees completed VR training up to 4x faster than in-person training and up to 1.5x faster than e-learning.

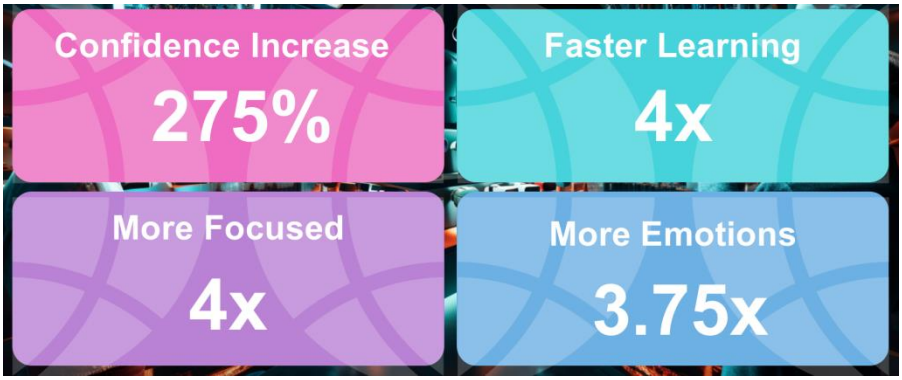


Figure 2. Main VR training efficiency findings in comparison with the traditional learning methods. Source: PWC Study on Effectiveness of Virtual Reality Training in Enterprises, 2020

Improved confidence: Employees who received VR training were up to 275% more confident in applying the skills they learned, an increase of 40% compared to classroom learning and 35% compared to e-learning.

Emotional connection: VR learners felt 3.75 times more emotionally connected to content than classroom learners and 2.3 times more than e-learners.

Focus and commitment: Learners trained in VR were up to 4 times more focused than e-learning learners and 1.5 times more focused than their classmates.

As a conclusion from the study, it can be derived that VR skills training is more effective than traditional methods and can lead to increased confidence, emotional connection, and focus. Also, using VR for soft skills training is a more cost-effective and worthwhile investment than in-person instruction or e-learning. This report suggests that VR is poised for enterprise-scale use, as falling costs and advances in technology make it an increasingly viable option for effective training programs.

A review of the existing literature shows that the effectiveness of VR in learning and training has been previously studied. Table 1 provides an overview of various studies demonstrating mainly positive outcomes from integrating VR into learning and training contexts.

Table 1. Selected Previous VR Efficiency in Training Studies, Research Context and Findings

Title	Authors	Findings	Published In
The Effectiveness of Virtual Reality-Based Training in Emergency Obstetric Care: A Systematic Review [7]	Forneris, S. G., et al.	Positive outcomes in training healthcare professionals in emergency obstetric care using VR.	Simulation in Healthcare, 2019
The Effect of Virtual Reality on Learning Outcomes in Health and Safety Training [8]	Gao, Z., et al.	VR positively impacted learning outcomes in health and safety training compared to traditional methods.	Journal of Science Education and Technology, 2019
A Meta-analysis of the Cognitive and Motivational Effects of Serious Games [9]	Wouters, P., et al.	Investigates the cognitive and motivational effects of serious games, including those implemented in virtual reality.	Journal of Educational Psychology, 2013

5. Impact of Labor and Skills Shortage and Talent Attraction

Uncover talent by immersing yourself into their social ecosystems: Prospective employees thrive where community converges

According to a 2021 report from the TechForce Foundation, the impending retirement of baby boomers, workforce mobility, turnover, and new job creation will increase the demand for more than 19,000 trauma technicians annually in the U.S. from 2021 to 2025. is expected to occur.

At the same time, the American Welding Society has highlighted worrying trends in the welding industry, predicting a shortage of approximately 400,000 welders by 2024. This compelling data strongly suggests that the shortage of skilled technicians is not just a short-term challenge, but a long-term problem that threatens to impact industry capabilities in the coming years.

To address this critical challenge in the world of work, virtual reality (VR) technology is proving to be a promising solution [10]. VR has the potential to increase engagement, accelerate upskilling, and facilitate remote and self-guided training. This technology represents an innovative approach to addressing the ongoing skills shortage and provides a scalable and innovative way to prepare the workforce for the demands of these specialities.

Given that individuals often succeed within their communities, the use of VR technology offers an innovative solution.

An effective approach is to create virtual gamification simulations that are carefully designed to replicate the challenges and dynamics of real-world work environments [10]. These simulations provide a comprehensive and immersive experience, giving potential candidates a glimpse into the complexities of real-world scenarios. Using this method, companies can effectively assess candidate skills and ensure a match between skills and job requirements. Additionally, collaborative VR experiences have great potential, especially when it comes to allowing young talent to connect remotely with peers, mentors, and industry experts. This makes it easier to expand the professional network, even in a virtual space, and fosters a sense of community. Through collaborative VR initiatives, companies are not only bridging geographic gaps, but creating environments that help share knowledge, mentor, and seamlessly integrate new talent into the professional world [11].

6. Conclusions

Proactive strategies such as immersion in social ecosystems and the integration of game-based simulations can proactively address pressing talent shortages and effectively bridge the growing gap between education and industry needs. It has proven to be an important method. Immersing in the social ecosystem represents a paradigm shift beyond traditional recruitment and leveraging the diverse and dynamic networks in which potential talent grows. This approach not only allows for a more holistic understanding of candidates, but also a community-focused approach to talent acquisition. At the same time, the incorporation of gamified simulation represents a breakthrough in training methods within virtual reality (VR), creating realistic and adaptable scenarios that reflect the complexity of real-world work environments. The gamified VR simulations improve the practical application of learned skills. This not only enriches the learning experience but also better prepares individuals for the challenges they will face in their careers and attract the future workforce.

Moreover, VR's role as a dynamic force goes beyond the immediate problem of talent shortages. This represents a fundamental shift in the entire skills development and talent acquisition landscape, and VR technology, with its immersive capabilities and versatility is

positioned as a catalyst for change in organizations. This transformation is marked by a shift to a more engaging, personalized, and experiential learning environment. By seamlessly combining immersive experiences with innovative learning methods, VR is a pioneering solution that not only bridges the immediate skills gap but also anticipates and adapts to the changing needs of the modern world.

References:

[1] N. Rendevski et al., "PC VR vs Standalone VR Fully-Immersive Applications: History, Technical Aspects and Performance," 2022 57th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), Ohrid, North Macedonia, 2022, pp. 1-4, doi: 10.1109/ICEST55168.2022.9828656.

[2] N. Rendevski, K. Veljanovski and N. Emini, "FPS Performance Factors in Standalone Virtual Reality Applications," 2023 58th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST), Nis, Serbia, 2023, pp. 349-352, doi: 10.1109/ICEST58410.2023.10187389.

[3] N. Rendevski et al., "Immersive technologies in 5G-enabled applications: Some technical challenges of the novel usage models", AIRCC's International Journal of Computer Science and Information Technology, Vol. 13, Issue 6, pp. 31-37

[4] M. Sünсksen, H. Bendig, M. Teistler, M. Wagner, O. J. Bott and K. Dresing, "Gamification and virtual reality for teaching mobile x-ray imaging," 2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH), Vienna, Austria, 2018, pp. 1-7, doi: 10.1109/SeGAH.2018.8401364.

[5] V. C. Camarena Quispe, J. C. Cubas Aguinaga, L. G. Diaz Huaco and J. T. Cruz, "DENTORACULUS - A gamified Virtual Reality experience for dental morphology learning," 2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Shanghai, China, 2023, pp. 1007-1008, doi: 10.1109/VRW58643.2023.00349.

[6] PricewaterhouseCoopers. "PwC's study into the effectiveness of VR for soft skills training." PwC, 01 Oct. 2023, www.pwc.co.uk/issues/technology/immersive-technologies/study-into-vr-training-effectiveness.html.

[7] Ameh CA, Mdegela M, White S, van den Broek N. The effectiveness of training in emergency obstetric care: a systematic literature review. Health Policy Plan. 2019 May 1;34(4):257-270. doi: 10.1093/heapol/czz028. PMID: 31056670; PMCID: PMC6661541.\

[8] Gao, Yifan, et al. "The effectiveness of traditional tools and computer-aided technologies for health and safety training in the construction sector: A systematic review." Computers & Education, vol. 138, 1 Sept. 2019, pp. 101-15, doi:10.1016/j.compedu.2019.05.003.

[9] Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. Journal of Educational Psychology, 105(2), 249–265. https://doi.org/10.1037/a0031311

[10] SprayVerse VR Training Simulator, https://sprayverse.com, Accessed October 1, 2023

[11] WeldVR VR Welding Training Simulator, https://weldvr.com, Accesed October 1, 2023

Cyber Risk Management Tool for Improving the Cybersecurity Maturity in the Companies

Kire Jakimoski[1], Oliver Iliev[2], Gorgi Kakashevski[3], Biljana Stojchevska[4], Jelena Gjorgjev[5], Elena Petrova[6]

[1] *Faculty of Informatics, AUE-FON, Str. Kiro Gligorov, No. 5, Skopje, Republic of N. Macedonia*
[2] *Faculty of Informatics, AUE-FON, Str. Kiro Gligorov, No. 5, Skopje, Republic of N. Macedonia*
[3] *Faculty of Informatics, AUE-FON, Str. Kiro Gligorov, No. 5, Skopje, Republic of N. Macedonia*
[4] *Faculty of Informatics, AUE-FON, Str. Kiro Gligorov, No. 5, Skopje, Republic of N. Macedonia*
[5] *Faculty of Informatics, AUE-FON, Str. Kiro Gligorov, No. 5, Skopje, Republic of N. Macedonia*
[6] *Faculty of Informatics, AUE-FON, Str. Kiro Gligorov, No. 5, Skopje, Republic of N. Macedonia*

kire.jakimoski@fon.edu.mk; oliver.iliev@fon.edu.mk, gorgik@fon.edu.mk; biljana.stojcevska@fon.edu.mk; jelena.gjorgjev@fon.edu.mk; elena.petrova@aue.edu.mk

Abstract:
Successful cyber risk management is more than needed nowadays in extended detection and response cyber security environments. Integration and automation of modern risk management includes implementation of well-known platforms and standards like ISO/IEC 27001, ISO/IEC 27005, CIS v8 that should be synchronized and well mapped. Many companies, especially small and medium-sized ones, usually don't have cyber risk teams and for that reason their cyber risk management processes should be simplified and automated. Companies can fully understand their risks only if they successfully implement a Cyber Risk Management process through successful implementation of the Cyber Risk Management Policy. So, it is very important to have automated and simple tool to assess the cyber risks and understand which security controls to perform to mitigate the cyber risks. In this paper authors developed cyber risk assessment tool for fast and efficient cyber risk assessment for the companies. This tool is asset based and it evaluates the risks with high precision according to the status of the security controls.

Keywords:
CIS (Center of Internet Security) Controls, Compliance, Cyber Risks, Risk Assessment, Security Controls, Vulnerabilities.

1. Introduction

Management of cyber risks is very important part of the Information Security Management System (ISMS). Risk is the happening of an unwanted event or the non-happening of a wanted event which adversely affects business. Cybersecurity Risk Management is an ongoing process for identification, analysis, and description of potential events and circumstances that can produce impacts to Information Security objectives. According to this information, management can take decisions about what risks are at acceptable level, and which risks require treatment to ensure potential impacts do not materialize.

A risk assessment and treatment process are key areas for implementation and maintenance of a successful ISMS, and they are crucial part of the ISO/IEC 27001 and ISO/IEC 27005 standards. Protection against information security threats is on appropriate level only when risks are completely understandable in order adequate security controls to be ensured. It is essential that organizations have an adequate risk assessment and treatment process in place to ensure that potential impacts do not become real, or if they do, contingencies are in place to deal with them. It is also essential that processes are sufficiently clear so that subsequent assessments produce consistent, valid, and comparable results, even when carried out by different people. Implementation of security controls is very important for the risk treatment process, so Annex A of the ISO/IEC 27001 standard and CIS v8 security controls are very important documents for the identification of appropriate controls. Tools for cyber risk assessment

are of great importance for implementation of successful risk assessment and treatment process. That is the focus of this paper.

Authors in [1] present case studies and real-life examples including risk assessment software like RAVEN and SAPPHIRE, as well as questions for students. In [2] authors propose a framework for cyber risk assessment and mitigation to estimate the attack probability and predict the required security technology to reduce the attack probability to some given level in the next year. Furthermore, expected loss is also calculated due to cyber attacks using risk modeling. Authors in [3] describe developed methodology using NIST cybersecurity framework and propose cybersecurity tool for evaluation with 35 questions to assess the maturity of the small to medium-sized enterprises according to the five NIST categories. In [4] a method for qualitative risk assessment algorithms is presented. This method ensures that the algorithm and the risk model are easy to understand. Method of this author includes creation of assessment algorithms for ten common cyber-attacks. Authors in [5] propose methodology to assess the cyber risks in smart homes to present the risks on home inhabitants and to suggest mitigation of the identified risks. Authors in [6] propose cyber and privacy risk management tool for assessing cyber and privacy risks in automated manner with decision-supportive capabilities. Integrated cybersecurity risk management is presented in [7] including prediction of risk types through machine learning techniques and systematic identification of critical assets.

Considering the above references, in this paper we focus on a cyber risk assessment tool aimed for small to medium-sized companies. It uses twenty generic cyber risks mapped with CIS v8 security controls. In this way, cyber risk tool achieves relatively simple way to assess the cyber risks and map them to well-known security controls related to the identified risks. Furthermore, this tool also gives instructions which security controls from CIS v8 to be implemented after identifying the level of risks. So, this tool achieves also view of the compliance to the standards like CIS v8 and ISO/IEC 27001 Annex A Controls.

The remainder of this paper is organized as follows. Section 2 presents the methodology used for developing the cyber risk assessment tool. In section 3 cyber risk assessment tool is described in detail. Section 4 concludes this paper.

2. Methodology for Developing the Cyber Risk Assessment Tool

Many companies usually do not pay too much attention to cyber security and do not have cyber security experts other than the system administrator that covers cyber security parts for them as well. For that reason, it will be highly beneficial to make their cyber risk management processes automated to the degree possible. Implementation and maintenance of a solid Information Security Management System (ISMS) cannot be done without risk assessment and treatment routines, which are crucial parts of the ISO/IEC 27001 [8] and ISO/IEC 27005 standards. So, it is very important to have automated and simple tool quickly to assess the risks and understand which security controls to perform to mitigate the cyber risks.

This research will improve the cyber risk management process in the companies by using well-known standards, methodologies, and tools. It will be done by implementing a simple model for the cyber risk management process of companies. Cyber risks and cyber security controls that are applicable especially to small companies and to mid-level companies are presented in detail with this tool. So, it can automate and improve the efficiency of the cyber risk management in these organizations. Cyber risk assessment tool is developed integrating the simplerisk tool from https://www.simplerisk.com/ and CIS controls from the Center of Internet Security https://www.cisecurity.org/. The tool consists set of questions with simple answers (Yes or No) that automatically generate 20 cyber security risks that are mapped to the IG1 Security Controls from CIS controls.

Hence, well-known standards, methodologies, and tools are used for building the tool for automation of the cyber risk management process in companies. Simplerisk.com tool is used for generating general cyber risks. These 20 cyber risks are used as a basis for creating the tool for assessing cyber risks.

So, in the first phase of this research generic risks from simplerisk platform are mapped with the security controls in CIS v8 standard to build a more sophisticated tool for cyber risk management

processes in the companies. Furthermore, cyber risks and controls are classified according to the asset types that are used in CIS security controls.

Then, for each risk according to the mapped security controls and subcontrols from CIS standard, set of new questions are created. Therefore, after answering each set of questions for each risk, the tool automatically generates the level of the cyber risk.

Furthermore, it also gives instructions for obtaining security controls for the questions that are answered with No. Benefits of this tools are that it generates different risk levels according to the results from the questionnaire and offers instructions for solving the vulnerabilities. Another benefit is that automatically solutions are mapped with CIS v8 standard [9] because it uses the security controls from CIS. That also enables to relate to ISO/IEC 27001, considering the available tools for mapping CIS v8 to ISO 27001:20222 security controls [10]. This kind of integrated and detailed cyber risk assessment tools is more than needed for the companies, especially for small to mid-sized companies.

3. Description of the Cyber Risk Assessment Tool

As it was already explained in the previous section, for developing the cyber risk assessment tool in the research, reference for the cyber risks was the platform from simplerisk (https://www.simplerisk.com/). Cyber security controls to mitigate the cyber risks generated with this tool are used from the CIS v8 standard (https://www.cisecurity.org/controls/v8) to build a simple mechanism for decreasing the cyber risks related to these controls**.** IG1 group of subcontrols in CIS v8 that are under each of the 18 main controls in CIS v8 is exactly what is used for this part of the proposed cyber risk assessment tool.

The level of implementation of subcontrols under each main control is used to define the level of the appropriate risk, whether it is insignificant, low, medium, high, or very high. I used the risk matrix presented in Fig. 1 for defining the level of cyber risk generated after answering the set of questions regarding the specific risk, so levels are: Insignificant, Low, Medium, High, and Very High risks.

Fig. 1 Cyber Risk Matrix

Formula in excel presented in (1) is created for defining the risk level according to the answers on the questions related to each risk. If the total score when answering the questions is 0, Insignificant risk level will be generated. If the total score when answering the questions is higher than 0 and lower than 4, Low risk level will be generated. If the total score when answering the questions for the appropriate risk is higher than 4 and lower than 7, Medium risk level will be generated. If the total score when

answering the questions is higher than 7 and lower than 10, High risk level will be generated. If the total score equals to 10 and above, Very High Risk will be generated.

=IF(F2="","BLANK",IF(F2=0,"INSIGNIFICANT",IF(F2<4,"LOW",IF(F2<7,"MEDIUM",IF(F2<10,"HIGH",IF(F2>=10,"VERY HIGH")))))) (1)

Cyber Risk Assessment Tool is firstly developed in excel format with formulas and functions. In Fig. 2 part of the excel format view of the tool is presented for the second main control "Inventory and Control of Software Assets".

In CIS Control v8 there is "Asset type" for each of the 18 security controls subjects. So, in excel each of the asset types is put in a separate sheet. There are five asset types: Devices, Applications, Users, Network and Data. Under each of the controls, all subcontrols that are for IG1 implementation group are taken into consideration. Then, questions for each of the controls and appropriate subcontrols are created with given points according to the number of questions for each risk. When answered with Yes, there are no points, and when answered with No, appropriate points are inserted. Per example, if there are 5 control and its subcontrols, each sub control weights 2 points. If all questions regarding these subcontrols are answered with Yes, there are no points, so in total we have 0 points, and that gives Insignificant Risk. If all questions are answered with No, we have in total 10 points, and that gives Very High Risk.

Column 1 in Fig. 2 presents the main controls and subcontrols from CIS v8 that belong to each asset type. In this case, main control "Inventory and Control of Software Assets" and its subcontrols are shown for the asset type "Applications". In column 2 created questions are presented for each category of CIS v8 controls. Third column in Fig. 2 presents the answers that should be answered with simple Yes/No answers from the clients. Next two columns present the Risk details for each category and the Risk Level. Then, column "CIS v8 Instructions" presents the instructions for implementing security controls to decrease the cyber risks. After their implementation, client will answer related questions with "Yes" and risk level will be decreased.

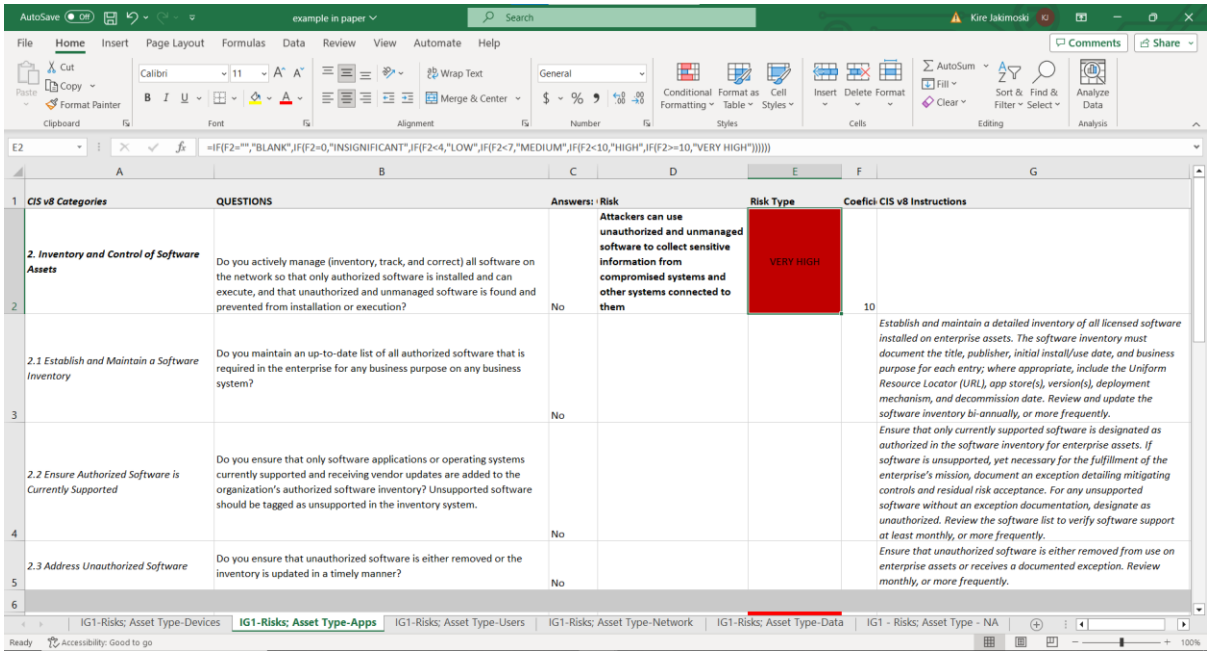


Fig. 2 Cyber Risk Assessment Tool in excel format

Cyber Risk Assessment Tool is also developed in web form using the already created functions and formulas in the excel format. Fig 3 presents the web form view of the tool.

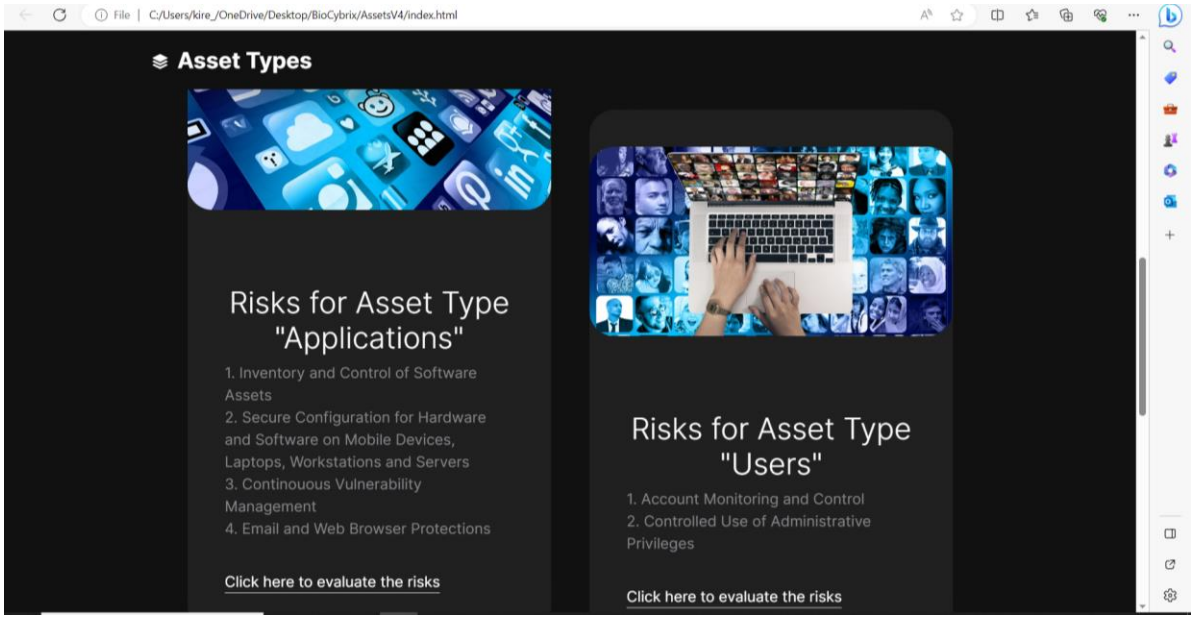


Fig. 3 Web form view of the cyber risk assessment tool

Hence, generated cyber risks and instructions for implementing the security controls to decrease them in the Risk Assessment tool gives simple control and comfort of the clients that use this tool. This gives great benefit in the process of cybersecurity risk management automation.

Table 1 on the other side, presents mapping of the controls in CIS v8 with the Annex A Controls in the ISO/IEC 27001 standard. So, presented cyber risk assessment tool is well mapped with CIS v8 and ISO 27001 Annex A controls. This makes the tool compliant to well-known compliance standards for cyber security.

Table 1:
Mapping of the pending risks with CIS v.8 controls and ISO 27001 Annex A Controls

General Risks	Related Controls in CIS Controls Version 8	Related ISO 27001 Annex A Controls
1. Attackers can use unauthorized and unmanaged devices to gain access to network.	1. Inventory and Control of Enterprise Assets	A.8.1.1; A.11.2.5; A.13.1.1; A.9.1.2
2. Attackers can use unauthorized and unmanaged software to collect sensitive information from compromised systems and other systems connected to them.	2. Inventory and Control of Software Assets	A.8.1.1; A.12.5.1; A.12.6.2
3. Attackers can exploit vulnerable services and settings to compromise operating systems and applications.	4. Secure Configuration of Enterprise Assets and Software	A.8.1.3; A.14.2.5; A.14.2.2; A.12.1.2
4. Attackers can take advantage of gaps between the appearance of new knowledge and remediation to compromise computer systems.	7. Continuous Vulnerability Management	A.9.2.3; A.12.6.1
5. Attackers can misuse administrative privileges to spread inside the enterprise.	6. Access Control Management	A.9.2.3; A.9.4.2; A.9.4.3; A.9.4.4; A.12.4.3
6. Attackers can hide their location, malicious software, and activities on victim machines due to deficiencies in security logging and analysis.	8. Audit Log Management	A.12.4.1; A.12.4.3: A.12.4.4
7. Attackers can craft content to entice or spoof users into taking actions that greatly increase risk and allow introduction of malicious code, loss of valuable data, and other attacks.	9. Email and Web Browser Protections	A.8.1.3; A.12.2.1: A.12.6.2; A.13.1.1; A.13.2.3
8. Attackers can use malicious software to attack our systems, devices, and data.	10. Malware Defences	A.12.2.1: A.12.4.1
9. Attackers can scan for remotely accessible network services that are vulnerable to exploitation.	4. Secure Configuration of Enterprise Assets and Software 13. Network Monitoring and Defence	A.13.1.1; A.13.1.2; A.13.1.3

10. Attackers can make significant changes to configurations and software on compromised machines, and it may be extremely difficult to remove all aspects of their presence.	11. Data Recovery	A.12.3.1
11. Attackers can gain access to sensitive data, alter important information, or use compromised machines to pose as trusted systems on our network by exploiting vulnerable services and settings	12. Network Infrastructure Management 4. Secure Configuration of Enterprise Assets and Software	A.12.1.2; A.13.1.1; A.13.1.3
12. Attackers can exploit vulnerable systems on extranet perimeters to gain access inside our network.	13. Network Monitoring and Defence 15. Service Provider Management	A.9.4.2; A.13.1.1
13. Attackers can exfiltrate data from our networks compromising the privacy and integrity of sensitive information.	3. Data Protection	A.8.2.1; A.13.2.3; A.6.2.1; A.8.3.1
14. Attackers can find and exfiltrate important information, cause physical damage, or disrupt operations due to improper separation of sensitive and critical assets from less sensitive information.	6. Access Control Management	A.8.1.1; A.9.1.1; A.10.1.1; A.12.4.3; A.13.1.1; A.13.1.3
15. Attackers can gain wireless access and bypass our security perimeters in order to steal data.	4. Secure Configuration of Enterprise Assets and Software 6. Access Control Management	A.8.1.1; A.8.1.3; A.10.1.1; A.13.1.1; A.13.1.3
16. Attackers can impersonate legitimate users by exploiting legitimate but inactive user accounts.	5. Account Management	A.8.1.1; A.8.1.3; A.9.2.1; A.9.2.6; A.10.1.1; A.12.4.1; A.13.1.1
17. Attackers can exploit employee knowledge gaps to compromise systems and networks.	14. Security Awareness and Skills Training	A.7.2.2
18. Attackers can take advantage of vulnerabilities in software to gain control over vulnerable machines.	16. Application Software Security	A.10.1.1; A.12.1.4; A.12.6.1; A.14.2.1; A.14.2.5
19. An attacker may have a greater impact, cause more damage, infect more systems, and exfiltrate more sensitive data due to a poor incident response plan	17. Incident Response Management	A.16.1.1; A.16.1.3
20. Attackers can take advantage of unknown vulnerabilities due to a lack of testing of organization defences.	18. Penetration Testing	A.12.6.1: A.16.1

4. Conclusions

This paper presents integrated approach for cybersecurity risk management considering platforms and standards for cyber risks and security controls. Dealing the cyber risk management in the companies considering the approach presented in this paper will help cybersecurity experts in improving the cyber risk management in the companies, which is one of the most important and complex issues in cyber security.

Proposed cyber risk assessment tool in this paper will help companies to assess the cyber risks relatively fast with a lot of details for the risk levels and the appropriate security controls that should be implemented to decrease the identified cyber risks. That is of great benefit, especially for the small to mid-sized companies which usually don't have cyber security team in their staff.

References:

[1] Ostrom LT, Wilhelmsen CA. Risk assessment: tools, techniques, and their applications. John Wiley & Sons; 2019 Jul 30.

[2] Mukhopadhyay A, Chatterjee S, Bagchi KK, Kirs PJ, Shukla GK. Cyber risk assessment and mitigation (CRAM) framework using logit and probit models for cyber insurance. Information Systems Frontiers. 2019 Oct; 21:997-1018.

[3] Benz M, Chatterjee D. Calculated risk? A cybersecurity evaluation tool for SMEs. Business Horizons. 2020 Jul 1;63(4):531-40.

[4] Erdogan G, Refsdal A. A method for developing qualitative security risk assessment algorithms. InRisks and Security of Internet and Systems: 12th International Conference, CRiSIS 2017, Dinard, France, September 19-21, 2017, Revised Selected Papers 12 2018 (pp. 244-259). Springer International Publishing.

[5] Ali B, Awad AI. Cyber and physical security vulnerability assessment for IoT-based smart homes. sensors. 2018 Mar 8;18(3):817.

[6] Gonzalez-Granadillo, G.; Menesidou, S.A.; Papamartzivanos, D.; Romeu, R.; Navarro-Llobet, D.; Okoh, C.; Nifakos, S.; Xenakis, C.; Panaousis, E. Automated Cyber and Privacy Risk Management Toolkit. *Sensors* 2021, *21*, 5493. https://doi.org/10.3390/s21165493

[7] Kure HI, Islam S, Mouratidis H. An integrated cyber security risk management framework and risk predication for the critical infrastructure protection. Neural Computing and Applications. 2022 Feb 2:1-31.

[8] ISO 27001 Annex A Controls. International Organization for Standardization. Available at: < https://www.iso.org/standard/54534.html> [Accessed 28 April 2023].

[9] CIS Critical Security Controls Version 8. Center for Internet Security. Available at: < https://www.cisecurity.org/controls/v8> [Accessed 28 April 2023.]

[10] Mapping of CIS v8 to ISO/IEC 27001:2022 controls Available at:<https://www.cisecurity.org/insights/white-papers/cis-controls-v8-mapping-to-iso-iec-27001-2022>[Accessed 05 October 2023].

Cyberbullying: Characteristics, Consequences, and Impact on Youth in North Macedonia

Marina Dodevska[1] and Nikola Janev[1]

[1]*Ss.Cyril and Methodius University, Skopje, North Macedonia*

marina.dodevska@students.finki.ukim.mk; nikola.janev@students.finki.ukim.mk;

Abstract:
The paper aims to provide an overview of cyberbullying as a concept, its characteristics and consequences, as well as its impact on the young population. For this purpose, we conducted a survey among people from 10 to 25 years of age to find out if they were a victim of cyberbullying and how they solved the problem. At the end, we present the answers from the survey. We will explain how parents can check on their kids to see what they do when they go online.
Keywords:
cyberbullying, young people, social media

1. Introduction

[1] The world is more connected than it has ever been. Skype, email, Facebook, Twitter. We can snap a picture on our phones, post a thought, or share a joke with thousands of other people - all within a matter of seconds. Today's teens, having never known a world without the Internet, are especially adept at picking up new technology and using electronic communication to create vital social networks. Problems arise when the technology that is supposed to bring people together is used instead to abuse others, pushing peers out of their social network into a world that is filled with loneliness, embarrassment, fear, or shame.

[2] The parents often buy different appliances like laptops, gaming consoles, and mobile phones for their kids who are not even 13 years old (which is the minimum age to join most social networks), just not to "disturb" them. This has a bad impact on them because parents don't know what their kids watch or do on their mobile phones. Most websites like YouTube made a special application just for kids called "YouTube Kids" which you can watch videos that are made for kids. When the kids are on their phones, they can do whatever they want, from messaging strangers to posting comments on social networks.

2. Definition and characteristics

[3] The definition of cyberbullying varies among scholars. But generally, cyberbullying (or online bullying) can be defined as a deliberate, unwanted, and repeated act of aggression that occurs over digital devices such as computers, smartphones, gaming consoles, or tablets. Cyberbullying takes on several different forms – both direct and indirect – which can make it difficult to identify as online harassment.
What we mean by the direct form of cyberbullying is when the one who commits the cyberbullying sends direct text, images, recording voices, and videos to the bullied person via messaging apps.
On the other side, the indirect type of bullying is when the one who commits the cyberbullying posts on social media (publicly) something about the bullied person, for example, that might be video, photo, etc.
Generally, online bullying will have the following characteristics:
*intent,
*aggression,
*Contact through electronic devices,

*repetition,
*causes harm to target
Online bullying involves the use of technology to harass, intimidate, or embarrass cyberbullies who often hide behind the anatomy of the Internet, which can make it more challenging to identify and hold them accountable for their actions.

There are several ways a person can be bullied online whether it is direct or indirect:
*sending harmful messages or threats,
*spreading rumors online,
*sharing embarrassing photos or videos,
*creating fake profiles to harass others,
*posting comments on the social networks

Of course, when cyberbullying occurs, we have consequences.
These are just some of the consequences that can happen if cyberbullying is not addressed in time:
- Falling into depression (feeling of unacceptance)
- Occurrence of anxiety
- Feeling of "suffocation"
- Sleep problems (Insomnia)
- Isolation from society and lack of the need for socializing
- Refusing to go to school, college, or social events

3. Related works

[4] This paper emphasizes the need for parents and educators to be aware of the dangers of cyberbullying and to be the first line of defense for an adolescent in recognizing, addressing, and resolving problems. Furthermore, they call on pediatricians, doctors, and psychiatric consultants to create a comfortable atmosphere for adolescents to discover and report their problems early in order to raise awareness in their communities.
[5] In this research, poor relationship skills in boys and low self-management in girls account for a part of the explanatory percentage for being a victim of bullying in both physical and online settings. In this context, they consider the provision of socio-emotional tools to be important to its early prevention.
This research [6] is another reminder that the standardization of the field and the application of theoretical models are crucial for a deeper understanding of the various aspects of online violence in order to move away from descriptive presentations of prevalence data.
This study [7] surveyed students in Malaysia and found that students were not spared from the harmful effects of cyberbullying. Victims of cyberbullies suffer emotional and psychological stress, and inevitably their grades will suffer. Institutions, teachers, and parents must realize that cyberspace is a risky place and must take effective steps to protect victims.
An answer from clinical psychologist [8] tells us “What is the right age for kids to be on social media?”. He considers that unfettered access, without any screen controls or (parental) monitoring, should be delayed for as long as possible, at least not under 16 years old. He also says that a parent`s decision should not be made based on what other parents allow their kids to do. The parents need to know if their kids are ready to be on social media or not.
[9] A survey from the Pew Research Center finds that 59% of U.S. teens have personally experienced at least one of the six types of abusive online behaviors which are:
-Offensive name-calling
-Spreading of false rumors
-Receiving explicit images they didn`t ask for
-Constant asking of where they are, and what they are doing, by someone other than a parent
-Physical threats
-Having explicit images of them shared without their consent

And the most common type of cyberbullying was offensive name-calling.
In our research, we talk about the impact of cyberbullying on young people aged 10-25 in North Macedonia.

4. Methodology

The total sample included 73 young adults from the Republic of North Macedonia. 42.5% were female and 57.5% were male with an average age of 20-25 years old, ranging from 10-25 years.
In this research was used Google Forms as a survey tool. In our survey, we had different questions like which sex our respondents are, how old are they, what social networks they use, and whether they share posts privately or publicly.
Our questions were answered by multiple choice, submitting a text and only one answer. In our work, we surveyed 73 people in North Macedonia.
We distributed the survey through communication channels such as Discord, Facebook, Reddit, LinkedIn, and among students from 10 to 25 years of age.
We collected the answers in about three months.
The results of the research are presented and discussed in the next section.

5. Results

According to the picture, we can notice that 57, 5 % of our survey was answered by boys, and 42, 5 % by girls.

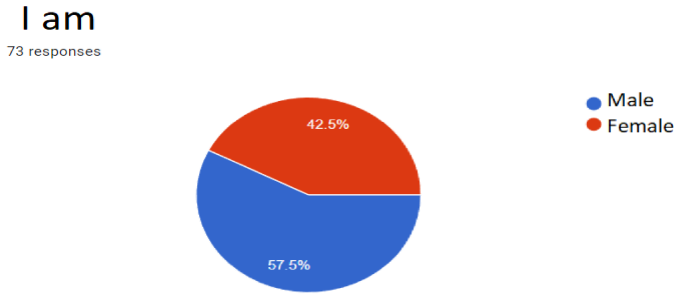


Fig1: More than half of our respondents are male

We asked our respondents what age they are and 80.8% answered that they are 20-25 years old, 17.8% are 15-20 years old and very few are 10-15 years old.

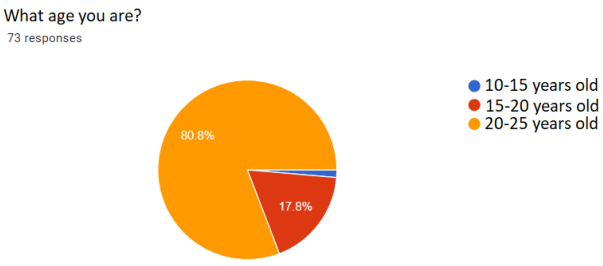


Fig.2: The average age is 20-25 years old

We asked our respondents which social networks they use, and as we can see, the largest percentage of them use Instagram (94.5%), and the least use Snapchat (24.7%).

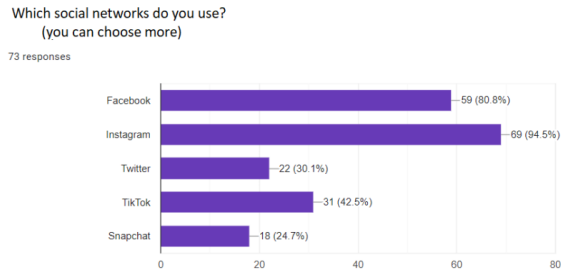


Fig.3: Instagram and Facebook are the most popular social networks in our survey

We asked our respondents with whom they share their posts on social networks and 67.1% share them with their Facebook friends, 26% publicly, and a small number of them share privately, that is, only they can see the posts.

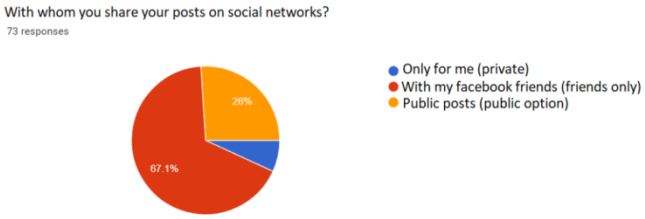


Fig.4: Friends only option on Facebook is used more than the public option

What pleases us is that 100% of our respondents answered that they had heard of the term cyberbullying.

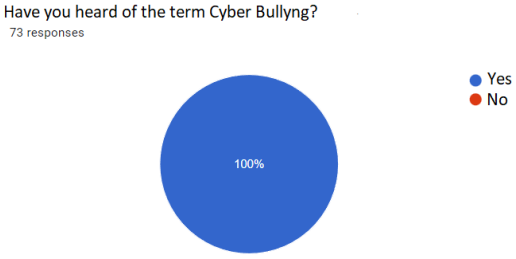


Fig.5: All of our respondents heard of the term cyber bullying

We asked our respondents if they had experienced cyberbullying or if they knew someone with such a case, so 35.6% knew someone with such a case, 23.3% had experienced bullying and 41.1% had not experienced bullying.

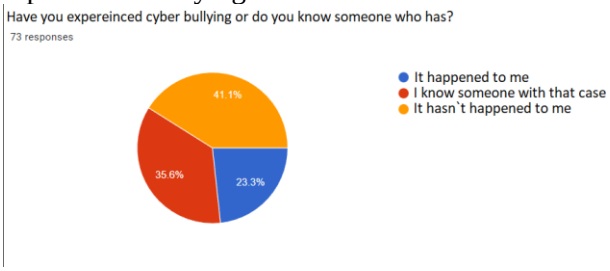


Fig.6: Most of our respondents never experienced cyberbullying, after that they know someone

We asked our respondents whether in their case ordinary violence (physical or psychological) turned into cyberbullying or vice versa and 57.5% answered that it did not turn into cyberbullying, 21.9% started as usual and then turned into bullying and 20.5% it started as bullying and turned into ordinary

violence.

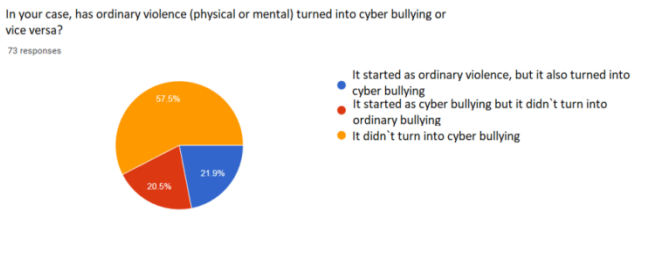


Fig.7: Peer bullying among students is here, whether is online or physical

We asked our respondents if they had received a threatening message online and as we can see 53.4% never received it, 41.1% rarely received it and a small percentage often received such messages.

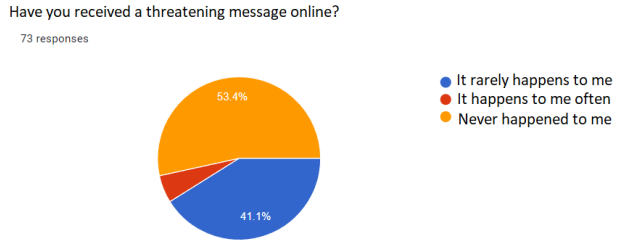


Fig.8: Most of our respondents never encountered a threatening message online

We asked our respondents how they felt when they received a threatening message online and as we see most of them (55.8%) did not pay attention to the message, 32.7% got annoyed and a small part of them (21.2%) were sad.

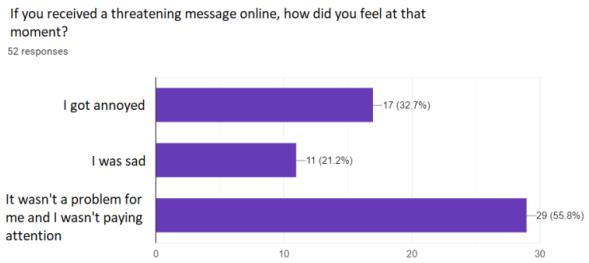


Fig.9: Most of our youth had no problem if they ever received a threatening message online

We asked our respondents if they shared their problem with someone older than them and most of them (60.3%) did not share, 34.2% shared the problem, while a small part of them did not share because of fear.

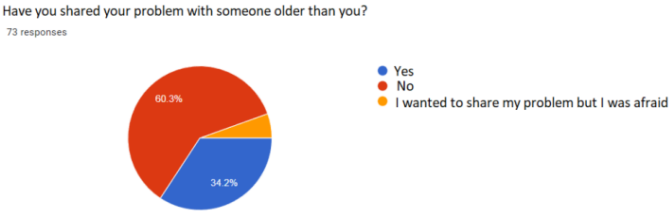


Fig.10: Most of our respondents didn’t share the problem

We asked our respondents how often they encounter bad words on social networks and 78.1% often encounter them, 19.2% sometimes, and a small part of them never encounter them.

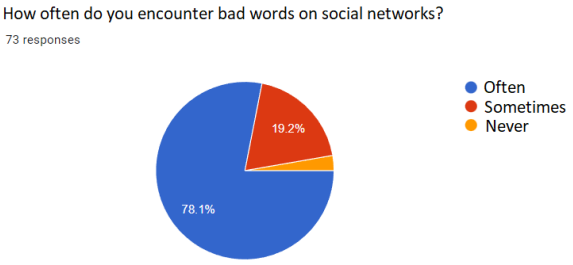


Fig.11: Bad words and threatening messages are often present

We asked our respondents if their pictures/personal information had been misused by friends/society people and 64.4% said no, while 35.6% said yes.

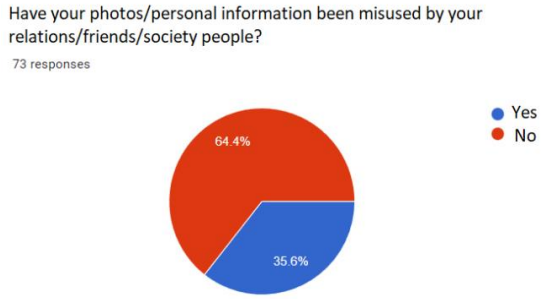


Fig.12: A small percent of our respondents said Yes according to our question

We asked our respondents if they would seek help from a psychologist if they experienced bullying in order to resolve the situation more easily and 58.9% answered no, while 41.1% answered yes.

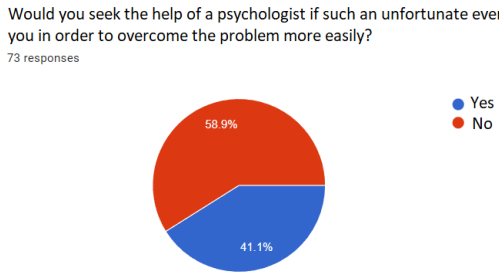


Fig.13: The scary part of the answers to this question is that most of our respondents won`t seek help if bullying occurs to them.

We asked our respondents how long it took them or a peer to recover from experiencing cyberbullying and 74% answered less than a year, 16.4% more than a year, and 9.6% a year.

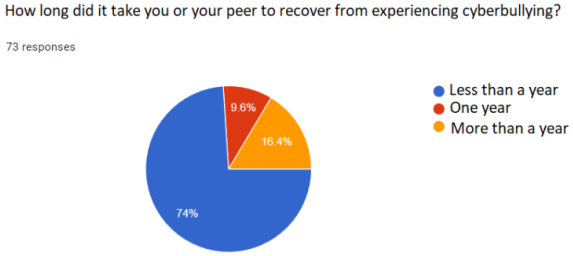


Fig.14: The good part is that our respondents recover quickly from cyberbullying

We asked our respondents if they think that social networks successfully deal with cyberbullying and most of them (91.8%) answered no, while 8.2% answered yes.

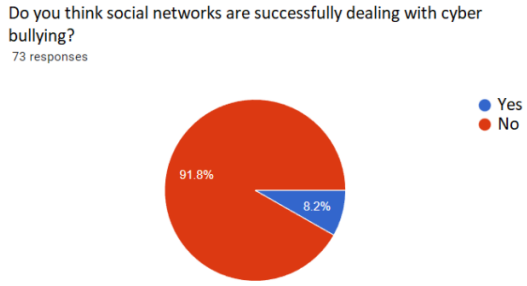


Fig.15: Our respondents think that social networks successfully deal with cyberbullying

We asked our respondents if they had ever reported a comment/picture/profile that was associated with cyberbullying and most of them (75.3%) answered more than once, 16.4% reported once, and a small part of them not once.

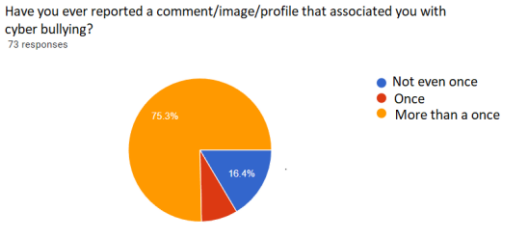


Fig.16: The reporting occurs more often according to our respondents

Our respondents gave several answers regarding what they would do to prevent cyberbullying:

-Education
-Providing support, in order to reduce the consequences
-Making friends if the person knows each other, helping them overcome it and talk to someone
-Reporting of profiles that commit violence in cyberspace
-Telling the person doing the cyberbullying to stop it because
hurts someone else doing it knowingly/unknowingly
-Advice for seeking help from an adult (educators, friends, school psychologists, and parents)

6. Conclusion

As we can see from the answers of our respondents, to help people who are bullied online, it is necessary to seek help from an adult (educator, parent, psychologist). We agree with the answers that suggest asking for help from an adult and not hiding the problem because things don't get solved without sharing.

We are all human and everyone needs help from the "outside" in life. If the problem is kept inside, the person may shut down and ignore the environment. Man is not created to live alone, every problem, if it cannot be solved alone, is solved with the help of others. Not all people have bad intentions towards us.

The other step is to always report such profiles no matter where we are on the Internet and not pay attention to them. There are thousands of comments on social media, but not all of them should be considered because they are all on a different topic.

Parents need to talk to the parents of the cyber bully and find a non-confrontational solution to the problem, without adding fuel to the fire and supporting the cyber bully.

Advising classmates to report the problem and not remain silent, to tell the cyberbully that it is wrong, that we are all human, and that sometimes we can feel the same regardless of whether the person being bullied did something wrong (insulting, humiliating, etc.).
Parents should show more interest in their children as and when they use their electronic devices. They should at least see if their children are messaging strangers on social media or posting inappropriate comments on social media or online sites. For example, they should check browser history, view history, comment history, etc.
We sincerely hope that cyberbullying will stop because everyone deserves to have a good life.

References:

[1] PREVNet: Canada's Healthy Relationship Hub, section: Cyberbullying, Link: https://www.prevnet.ca/bullying/cyber-bullying
[2] Melinda Wenner Moyer: Kids as Young as 8 Are Using Social Media More Than Ever, Study Finds, Link: https://www.pacesconnection.com/blog/kids-as-young-as-8-are-using-social-media-more-than-ever-study-finds-nytimes-com
[3] DISPARTI Law Group: What is Cyberbullying and how does it differ from traditional bullying, Link: https://www.dispartilaw.com/what-is-cyberbullying-definition/
[4] Gassem Gohal, Ahmad Alqassim, Ebtihal Eltyeb, Ahmed Rayyani, Bassam Hakami, Abdullah Al-Faqih, Abdullah Hakami, Almuhannad Qadri and Mohamed Mahfouz: Prevalence and related risks of cyberbullying and its effects on adolescent (Research), BMC Psychiatry, 2023, Link: https://bmcpsychiatry.biomedcentral.com/articles/10.1186/s12888-023-04542-0
[5] Juan Manuel Rodríguez-Álvarez, Santiago Yubero, Raúl Navarro, and Elisa Larrañaga: Relationship between Socio-Emotional Competencies and the Overlap of Bullying and Cyberbullying Behaviors in Primary School Students (Article), European Journal of Psychology and Education, 2021, Link: https://www.mdpi.com/2254-9625/11/3/49
[6] Lucija Vejmelka, Roberta Matkovic, and Miroslav Rajte: Cyberbullying in COVID-19 Pandemic Decreases? Research of Internet Habits of Croatian Adolescents (Article), MDPI, 2022, Link: https://www.researchgate.net/publication/366421939_Cyberbullying_in_COVD-19_Pandemic_Decreases_Research_of_Internet_Habits_of_Croatian_Adolescents
[7] Dr. Qais Faryadi: CYBER BULLYING AND ACADEMIC PERFORMANCE, International Journal Of Computational Engineering Research / ISSN: 2250-3005 / Vol 1, Issue 1 / Dec. 2011, Link: https://files.eric.ed.gov/fulltext/ED574784.pdf
[8] Tom Huddleston Jr: Is there a 'right' age for kids to be on social media, CNBC.COM, 2023, Link: https://www.cnbc.com/2023/05/26/the-right-age-for-kids-to-be-on-social-media-experts-say.html
[9] Monica Anderson, A Majority of Teens Have Experienced Some Form of Cyberbullying, Pew Research Center, 2018, Link: https://www.pewresearch.org/internet/2018/09/27/a-majority-of-teens-have-experienced-some-form-of-cyberbullying/

Internet Of Things, Big Data and Privacy Issues

Nebojsa Denić[1], Sonja D. Radenković[2], Ana Kovačević[3], Kostadinka Stojanović[4]

[1] *Faculty of Sciences, University of Pristina in Kosovska Mitrovica,*
[2] *Belgrade Banking Academy – Faculty of Banking, Insurance and Finance, Union University Belgrade,*
[3] *Faculty of Security Studies, University of Belgrade,*
[4]*Alfa BK University Belgrade, Faculty of Mathematics and Computer Sciences,*

nebojsa.dnic@pr.ac.rs, sonja.radenkovic@bba.edu.rs, kana@rcub.bg.ac.rs, kostadinka.stojanovic@alfa.edu.rs

Abstract. Although there are certainly benefits to using the Internet of Things, it is vital to provide improved security and privacy of the used devices so that their usability, speed, and scalability are not compromised. The question of whether an individual's privacy is compromised by this storage of the vast amount of data that the sensors have collected emerges. We lose control and the freedom to display ourselves whichever we choose if our privacy is violated. The issue of added dangers for these devices emerges given the enormous number of Internet-connected gadgets and the growing trend in this number. Are they secure enough? How trustworthy are they? Are people's privacy rights compromised by ongoing data collection?

Keywords: Internet of things, privacy, Big Data, smart cities

1. Introduction

Today, the phrase "Internet of things" (Internet of things) or "IoT" is frequently used to describe physical objects or "things" that have embedded electronics, software, sensors, and are connected or networked to enable data exchange with servers, centralized systems, and/or other connected devices based on various communication infrastructures. IoT applications enable network objects to affect their physical surroundings by acting physically as well (Bertino, 2016). The Alliance for IoT Innovation supports the European Commission's principal work programs on the Internet of Things (EC, 2016).

The phrase "Internet of Everything" has lately surfaced in addition to the term "Internet of Things," and it is described as the intelligent linking of people, processes, data, and devices (Erl et al., 2016).

The Internet of Things, like all new technology, has both positive aspects (such savings and a decreased greenhouse effect) and unsettling aspects (like the potential for privacy violations). By installing light sensors, it is possible to assess how much light is present in each area and switch on the lights only when necessary. This saves energy and lowers lighting expenses, but it also allows for the detection of pedestrian traffic..

Anytime, anywhere, and anything is a common association with the Internet of Things concept. It is important to keep in mind that with the idea of being "always on," 24 hours a day, 7 days a week, recording of the environment and our bodies' parameters may occur in some circumstances against our will. Can IoT harm our privacy in this way? How safe is it? IoT is a growing part of both people's and businesses' everyday lives, so it's important to strike the right balance between keeping corporate operations running smoothly on the one hand, and privacy and security on the other.

2. Smart Devices

Households already employ "smart" appliances like the smart thermostat, smart air conditioner, smart refrigerator, and smart smoke detector. The smart thermostat adjusts its energy use in accordance with the owner's behavioural patterns. However, the thermostat tracks and logs more than just energy use; it also keeps track of indoor temperature, humidity, ambient light, movement, and other factors (Schneier, 2015). In order to enhance energy efficiency, a smart air conditioner may learn about owner preferences. The intelligent refrigerator keeps track of when the food is correct/fresh. A smart house can use a wider variety of sensors. If we want to build a smart electrical system that affects the efficient use of energy and lessens the greenhouse effect, all of these gadgets will be crucial.

The first thermostat with Internet connectivity was developed by the Nest[1] firm, which is now a part of the Google Store. The current price of the third-generation smart thermostat is $249. For $3.2 billion, Google acquired Nest in 2014. On its website, the firm claims that since 2011, smart thermostats have prevented the use of about 14 million MWh (Nest, 2017). Along with selling a smart thermostat, carbon monoxide and smoke detectors, Nest also has a full line of future sensors planned.

A smart thermostat, for instance, analyses the owner's behavior patterns and adjusts as necessary to use energy effectively. The smart thermostat not only captures the indoor temperature but also humidity, ambient light, and movement.

Health-improving fitness equipment gathers and analyzes information about our bodies. For instance, a fitness tracker like a Fitbit (www.fitbit.com) or Jawbone (www.jawbone.com) gathers data on our motions while awake or asleep and examines our exercise and sleeping patterns. Additionally, this information can be found online.

Figure 1 displays the data obtained using the Jawbone bracelet, which tracks and analyzes information on our activity level, sleep patterns, and the number of calories we've burned.

Figure 1. View analysis results using the Jawbone[2] app

Applications for personalized medicine and health care are constantly collecting data from numerous devices, with the human body emerging as a significant source of data. All information gathered is stored in the cloud and/or sent to other devices such mobile phones (and may be sent to third parties)..

3. Big data and Internet of Things

Big data is a relatively recent phrase that is being used more frequently. Big data is the field that deals with the analysis, processing, and storage of substantial amounts of data, which frequently originate from several sources. Data sets are growing larger and more varied.

The idea of big data is not new and is influenced by technology, therefore the amount of data that qualifies as large data has evolved through time. For instance, 20 years ago, 1 GB of data was seen as a Big Data problem and required specialized computing resources to be used, however today, 1 GB can be moved, processed, and stored without any issues.

Based on the first letters of the phrases volume, velocity, and veracity in the English language, big data is frequently said to have three V-shaped properties. The five V qualities (Volume, Velocity, Variety, Veracity, Value) will be employed in the current definition because other characteristics were later added to these three fundamental ones.

When conventional data analysis, processing, and procedures are insufficient, big data solutions are used. In these cases, many unrelated data sets are matched, a sizable volume of unstructured data is processed, and hidden information is discovered in time-dependent data.

[1] https://store.google.com/us/category/google_nest?hl=en-US
[2] jawbone.com

The following characteristics of data produced by Internet of Things applications are listed by Chen et al. (2014):

- A significant volume of artificially manufactured data, ranging from straightforward numerical data to sophisticated multimedia data (video). Data that is "historical" as well as currently collected are evaluated and processed. With more gadgets connected to the Internet, there will be an increase in the volume of data generated.
- Due to the wide range of devices used to collect the data, the data is heterogeneous. Most of the data that were gathered were unstructured or semi-structured.
- Jaka korelacija vremena i prostora u podacima generisanim u uređajima *Internet of Things*. Svaki uređaj za skupljanje podataka je postavljen na određenu geografsku lokaciju, a u podacima je sadržano i p'vreme. Vremensko-prostorna komponenta podataka je izuzetno važna i prilikom analize podataka.
- Postojanje velike količina šuma tokom skupljanja i prenosa podataka u IoT. Među skupljenim podacima sa uređaja, nekada će samo mala količina neuobičajenih podataka biti značajna. Na primer, tokom snimanja (nadziranja) saobraćaja videom, nekoliko video frejmova koji su snimili saobraćajne prekršaje ili nesreće su mnogo značajniji od onih koji su snimili normalan tok saobraćaja.
- The data produced by Internet of Things devices show a strong link between time and space. The data contains p'time and each data collection device is positioned at a particular geographic location. When assessing the data, the data's time-space component is equally crucial.
- The presence of significant noise during IoT data collecting and transmission. Only a small portion of the unexpected data that the device collects will occasionally be important. For instance, many camera frames that captured traffic offenses or accidents during video recording (surveillance) of the traffic are far more important than those that captured the usual flow of traffic.

It is obvious that gathering a lot of data makes sense if it is examined. Based on real-time data from traffic, social media, and weather information, smart cities can optimize traffic flow.

4. Privacy of Data

The security of these devices, their networks, and the data itself will be of the utmost significance as the IoT expands and more gadgets become connected. The greenhouse impact can be minimized, living can be made easier, and big savings can be gained by gathering a lot of data, but one should also be mindful of the issues they can lead to. Management based on sensor data has the drawback of being more susceptible to attacks from various hackers.

With a balance between security and flexibility, encryption techniques must be effective and scalable over a large number of IoT devices with constrained computational resources. The software that runs on these devices needs to be safeguarded, and patch updates, network attacks.

In 2015, 90% of major corporations experienced a cyber assault, up from 81% in 2014, and 81% of smaller enterprises registered attacks, up from 74% in 2014, according to research from the year prior (PriceWaterhouse, 2015).

Although the Internet of Things has many benefits, including lower prices and a reduction in some global consequences (such as the greenhouse effect), it also comes with a significant issue. The issue is that the release of sensitive information, which violates an individual's right to privacy, occurs concurrently with the use of IoT. Sensitive medical information may so mistakenly be divulged, for instance, when keeping track of a person's health. Even non-personal information, like the apartment's interior temperature, can be used to determine whether or not there are occupants and how they spend their time if it intersects with knowledge of the air conditioning system's operation.

Additionally, the issue of whether decisions made without using human expertise and instead using analytical algorithms arises. Do we need to worry about potential flaws in the software that uses sensors to manage the outdated infrastructure?

Privacy is one of the characteristics that must be achieved in order to attain information security by the very meaning of the term. The protection of information's availability, privacy, and integrity on the Internet is referred to as internet security. (2015) SRPS ISO/IEC 27032. According to ISO/IEC 27033-1:2009, the Internet is a networked worldwide system that is open to the public. According to SRPS

ISO/IEC 27000:2016, the definition of privacy is the property that information is not made accessible to or disclosed to processes, entities, or people that are not allowed. However, there are situations when security may come at the expense of privacy.

Due to the widespread surveillance of people and the massive amounts of data we receive from various sources, privacy is one of the key issues in IoT applications. When adopting the IoT, this requirement poses a significant difficulty for the user because several sensors are gathering personal data (Vasilomanolakis et al. 2015). This will be an even bigger difficulty in the future if we consider the fact that the number of sensors will keep increasing. Models that "anonymize data" do exist, however Vasilomanolakis et al. (2015) found that they were either insufficient or difficult to utilize.

It's a prevalent misconception that since people haven't done anything wrong, they shouldn't worry about data protection. But is it truly the case? Do you want the public to hear the private chat you had? Do you require drapes for your bedroom? When we are alone or with close friends or family at home, we can unwind and act in a way that we most definitely wouldn't act at work without breaking the law.

In order to uphold a person's dignity and respect, they have the fundamental right to privacy. We can choose how we want to display ourselves to the world by choosing our level of privacy.

Biologist Peter Watts stated that mammals perceive surveillance as a threat, and this makes them paranoid, aggressive and vengeful. (Watts, 2014). Surveillance makes us feel like prey, and controllers feel like predators.

Studies have shown that constant surveillance is not healthy physically or emotionally, even the perception of constant surveillance, and creates feelings of depression and/or anxiety (Schneier, 2015). Violation of privacy is not the same for everyone, marginalized groups are more exposed, and also those in important positions, who are under the constant approval of people are even more exposed.

5. Conclusion

The issue of data privacy is growing more important as a result of the vast number of devices that are online. It is possible to consume less electricity because of Internet of Things applications, but data will also be supplied about how people move around their homes and how they spend their time; intelligent street lighting will gather information about how people move outside. We will all constantly be under surveillance, different types of environmental data will be continuously collected, and what is particularly unsettling is that that data will be preserved forever. Cameras and other utilized sensors will only get better, smaller, and more mobile.

With the development of the Internet of Things, new problems have emerged, such as how to stop devices from gathering and/or broadcasting information about the user's location and other contextual data, as well as how to help users comprehend the advantages and disadvantages of sharing personal data..

Because of this, the question of how much of an advantage the Internet of Things has in comparison to its risks and drawbacks arises. Is it possible (and how) to optimize and manage the kind and quantity of data that is gathered, as well as how it is processed, in order to lessen the danger of misuse and keep the advantages of improved resource management? The lives of citizens, the economy, and the environment would all be significantly impacted by a potential cyber attack on any essential infrastructure or social function in a large city.

All parties engaged in its storage, management, and use should take careful precautions to protect such information. It is crucial that users/subjects understand the breadth, content, and potential uses of the information acquired in this fashion. This is especially true for the population living in large urban areas. The following stage of development should offer convenient, straightforward solutions that will enable users to preserve their privacy and promote anonymity depending on the unique circumstances in order to capitalize on the advantages of IoT and prevent potential abuses.

REFERENCES:

1. Bertino, E. (2016). Data Security and Privacy in the IoT. In *EDBT* (pp. 1-3).
2. Belgrade, 2017: http://www.smartsantander.eu/index.php/testbeds/item/135-belgrade-summary

3. Carlo Maria Medaglia and Alexandru Serbanati. An overview of privacy and security issues in the internet of things. In The Internet of Things, pages 389–395. Springer, 2010.
4. Chen, M., Mao, S., & Liu, Y. (2014). Big data: A survey. *Mobile Networks and Applications*, *19*(2), 171-209.
5. Erl, T., Khattak, W., & Buhler, P. (2016). *Big Data Fundamentals: Concepts, Drivers & Techniques*. Prentice Hall Press
6. EC, 2016: European Commission. The alliance for internet of things innovation (AIOTI). Available online: https://ec.europa.eu/digital-single-market/alliance-internet-things-innovation-aioti
7. HP, 2015: Internet of things research study, http://h20195.www2.hpe.com/V4/getpdf.aspx/4aa5-4759enw
8. Gartner, 2017: http://www.gartner.com/newsroom/id/3598917
9. GEO,2015: https://geobrava.wordpress.com/2015/03/28/how-smart-cities-will-lead-the-internet-of-things-growth).
10. GovSantander, 2014: http://www.governing.com/topics/urban/gov-santander-spain-smart-city.html
11. IBM, 2016: http://ecc.ibm.com/case-study/us-en/ECCF-WGC12503USEN
12. Idtechex (2016): http://www.idtechex.com/internet-of-things-usa-16/show/en/
13. ISO/IEC 27033-1:2009 Information technology -- Security techniques -- Network security -- Part 1: Overview and concepts
14. Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal*, *79*(1), 1-14.
15. Lund, D., Turner, V., MacGillivray, C., & Morales, M. (2014, May). Worldwide and Regional Internet of Things (IoT) 2014 – 2020 Forecast: A Virtuous Circle of Proven Value and Demand. Retrieved January 25, 2016, http://www.business.att.com/content/article/IoT-worldwide_regional_2014-2020-forecast.pdf
16. PricewaterhouseCoopers (PwC). Information Security Breaches Survey 2015; HM Government: London,UK, 2015.
17. Santander, 2017: http://www.smartsantander.eu/
18. Schneier, B. (2015). *Data and Goliath: The hidden battles to collect your data and control your world*. WW Norton & Company.
19. SRPS ISO/IEC 27032:2015 Informacione tehnologije — Tehnike bezbednosti — Smernice za sajber bezbednost
20. SRPS ISO/IEC 27000:2016 Informacione tehnologije — Tehnike bezbednosti — Sistemi menadžmenta bezbednošću informacija — Pregled i rečnik
21. Vormetric, 2016: http://enterprise-encryption.vormetric.com/rs/480-LWA970/images/Vormetric_2016_Data_Threat_Report_Global_WEB.pdf
22. Vasilomanolakis, E., Daubert, J., Luthra, M., Gazis, V., Wiesmaier, A., & Kikiras, P. (2015, September). On the Security and Privacy of Internet of Things Architectures and Systems. In *Secure Internet of Things (SIoT), 2015 International Workshop on* (pp. 49-57). IEEE.
23. Watts, P. (2014) IAPP Canada Privacy SymposiumIAPP, 2014: https://iapp.org/news/a/if-you-cant-protect-the-data-burn-it-to-the-ground

Leveraging Social Media within Integrated Service Delivery of Personalized Public Services towards Proactive E-government

Mimoza Bogdanoska Jovanovska[1], Jadranka Denkova[2], Dragan Grueski[1,] Renata Petrevska Nechkoska[3], Natasha Blazheska-Tabakovska[1]

1 University "St. Kliment Ohridski", Faculty of Information and Communication Technologies, Partizanska bb, 7000 Bitola, RSM
2 University "Goce Delcev", Faculty of Law, Krste Misirkov No.10-A, 2000 Stip, RNM
3 University "St. Kliment Ohridski", Faculty of Economics, Prilepski braniteli, No. 143, 7500 Prilep, RNM

mimoza.jovanovska@uklo.edu.mk, jadranka.denkova@ugd.edu.mk, dragan.gruevski@uklo.edu.mk, renata.petrevska@uklo.edu.mk, natasa.tabakovska@uklo.edu.mk

Abstract:
Governments all over the world have had to continually come up with new strategies to advance and build e-government as a result of the incorporation of the newly developed digital technologies. As the front office is directly tied to communication, there is an ongoing push for new ways to communicate during the process of public services delivery. Social media platforms, as the most common phenomenon that technology has to offer to humanity, have shown to be effective tools for communication exchange. Social media are recognized as "window of opportunity" and crucial component of integrated service delivery for e-government development. This article focuses on personalized public services as the highest level of sophistication that some public services can acquire in the context of e-government projects aimed at becoming smart government. The paper's research technique is a case study of a current situation in the Republic of North Macedonia. The motivation for the proposed solution comes from the most recent published research on nations with comparable levels of e-government development. The purpose of this study is to point researchers and IT professionals in the direction of incorporating social media into the service delivery process for public services that can achieve a level of personalisation.

Keywords:
social media, integrated service delivery (ISD), proactive public services,

1. Introduction

From past to the present, e-government has been associated with monitoring and applying the latest technologies, with the aim of leveraging their benefits in enhancing the level of operation of institutions, on the one hand, and improving and facilitating the delivery of public services to users, on the other hand. Although the back office and the functioning of the institutions, along with their connection and external functioning, is the core of e-government [6], the segment of the front office is the part of the public sector that receives the most attention from those who invest (i.e. the state), and those for whom those investments are intended (i.e. citizens).

The incorporation of digital technologies has revolutionized the way governments provide services to their citizens. The complexity of digital public services varies, and such services can be delivered in a variety of manners [18]. Hence, the improvement of the citizen-government relationship is the most important task placed on the way to achieving the government's challenges for the citizens and within their objectives. Although at the beginning the goal of online services was to make the delivery of services more user-friendly, today this is no longer enough: the focus is placed on the delivery being adapted to the habits of citizens regarding the use of technologies. Even more, the focus is slowly shifting towards reducing the

administrative burden of citizens, by achieving proactivity on the part of the state in the realization of a certain number of services that allow it. Scholta and Lindgren [26] suggest that public organizations need to be proactive in digital public services in the approach to their clients / for what concerns the approach to their clients / so as to approach their clients. This means that the state has to sustain a process of continuous improvement of the interactions between users and government. In this context, IT specialists, fostered by the intense development of numerous digital technologies and their widespread application, were pressured to provide various creative ideas for integrating those technologies. The purpose of such integrations is to meet the citizens' requirements, as much as feasible. So, IT experts use already existing information on citizens' technology use and application habits , and create service delivery customized to the users' habits. This has led governments to consider integrating existing and widely used technologies and accomplishing integrated service delivery (ISD).

ISD, can be defined as the act of bringing together and merging digital technologies for the delivery of public services, in order for citizens to have seamless access to these services, depending on their preferences and needs [14].Such system opens the possibility for further development: namely, it gives the government an opportunity to change its approach from reactive to proactive. Although a number of solutions, frequently technical, have been elaborated to support ISD, the authors agree that there is still opportunity for future research in terms of the digital technologies that may be included in ISD.

As social media become "window opportunity" because of their widespread use, attempts are carried out to identify the ways to leverage these platforms in e-government initiatives, in order to facilitate communication, encourage citizen engagement, and optimize public service delivery [5]. UN [28] as an international survey, indicate that four out of five countries now have a national information portal containing links to government social media accounts, on platforms such as Facebook and Twitter. This situation results in a high uptake of online services, which is especially important on the path towards a fully digital government. Unfortunately, only a limited number of countries can brag of such accomplishments, which also serve as an excellent example of a benchmark. According to the EU's e-Government Benchmark for 2020, North Macedonia is ranked at the bottom of the list comprised of 44 countries covered by the study, in respect to the possibility for complete online service [7].

The paper shed light on the specific challenges faced by Macedonian citizens in the process of renewing personal documents, showing the shortcomings and the wrong, extremely outdated and dysfunctional concept behind the system created in e-government direction. As a result, the Macedonian scenario presents an intriguing framework for exploring the potential benefits and problems of integrating social media into e-government services, as part of an ISD which can lead to the creation of a stable base for a proactive approach of the state towards the citizens, especially for what concerns the delivery of certain public services.

This paper is structured in five sections, as follows: section 2 discusses the state of the art of social media and e-government research. In the third section we briefly elaborate on social media and integrated service delivery. Next, in section 4 we present a specific case study that can be used as contribution to theory and to streams of administrative burden and proactive digital public service produce – this represents the scenario where social media are used as part of the ISD. Finally, section 5 of this paper draws several conclusions and recommendations for further research.

2. Social media and e-government: the state-of-the-art

Almost a quarter of a century ago, an increase was noticed in the need to get more value out of public administration information technology investments, and the discussion over how to do so is becoming progressively more heated. As a result of this, many e-government programs and research efforts are centered on improving online service delivery [15]. Modernization of services through processes of reconstruction, and the adoption of new information technologies and systems, are often fundamental components of such improvement efforts [12].

In the second phase, following the rapid increase of ICT (Web 2.0), the possibility of social media becoming an important aspect of e-government development appears obvious, and the hope is that such development will have a substantial impact on how services are delivered in the public sector [19]. Creating new service delivery channels to serve citizens, and especially including social media as a channel, in conjunction with other technologies and applications, is recognized and recommended as a way of providing citizens with accessible tailored e-government services [3].

During the past decade, a body of research articles on the potential importance of social media in e-government emerged. Some papers are related to literature reviews [10, 12], while others are focused on describing and analyzing specific case studies where social media are involved in e-government functionality [9, 13]; Tursunbayeva and her colleagues [27] "identify, classify, critically appraise and synthesize the corpus of published research evidence relevant to the adoption, use and impacts of social media for e-government in the public health sector" with focus to social media as a channel for organization-citizen interaction (dissemination and feedback). Bharosa and his colleagues [4] analyze integrated service delivery from a government institutions aspect, having for focus a set of principles needed for ISD functioning, including social media as one of the channels. However, there is still very little empirical research on governments' use of social media platforms [16]. Some of the studies explore how to leverage social media to achieve e-government goals such as e-participation [2]. Going even further in this direction, some of the authors investigate Facebook posts as something that requires human attention, and therefore makes it a challenge for government organizations to engage through this medium to communicate with their citizens [8]. Only a small portion of this pool of research publications explores how social media could potentially be integrated into the service delivery process [19].

Recently, the researchers gathered citizens' opinions and attitudes on the use of social media in communication with public institutions. The goal of this research [12] is to discover what factors can persuade citizens to use social media for public service delivery; the outcomes of this study are intended to aid government organizations and policymakers in making decisions to provide a more citizen-centered service delivery.

Adae et al. [1] give examples of cases for integrating e-government portals with social media accounts in the African public sector. They determined that the usage of government portals to provide e-government services can be supplemented by social media accounts that provide information about the services and "serve as an avenue to engage the public about e-government services".

Some researchers point out the proactivity of public organizations to be desirable in the digital public services delivery. Proactivity advocates advise public institutions to change their approach to the citizens; citizens to receive public service even if they do not do anything or merely only confirm that they want that public service. Scholta and Lindgren [26] contribute to this theory by defining the conceptual changes that proactivity necessitates in digital public services. The importance of their work comes from their call for collaborative research by scientists from public administration, information systems, and service management, to connect the flows (administrative procedures) related to the administrative burden of citizens in order to achieve state proactivity in terms of delivery, through the integrated delivery of services employing various technologies.

3. Social Media and Integrated Public Service Delivery

3.1. Integrated service delivery (ISD)

The main problem in public service delivery is the fragmentation of each department and the scarce attention paid to information system development and user needs. The fragmentation and „silo" government structure complicates easy communication, and this results in customer dissatisfaction.

Public agencies continuously supplement their service delivery channels with online services, in order to improve service delivery and to make it more user-friendly. The policies and rules for improving service delivery are mostly based on customer oriented management paradigms, and they do not take into

consideration complex customer interaction processes [4]. As a remedy for "the fragmented landscape of public administration" [22] Integrated Service Delivery (ISD) is developed.

ISD is the process of bringing and fitting together government services, so that citizens can access these services in a seamless fashion, based on their wants and needs [14]. All actors included in the process of creation of ISD (architects, designers, system developers, programmers, administrative staff) need customer expectations as input, as well as information concerning citizens' use of public services through technological means.

ISD covers a wide range of initiatives: in the past it concerned traditional channels (such as the telephone and service desk) which remained from the online channel that include services over the Internet or via unstaffed kiosks. As a result, citizens and businesses now use a variety of channels to contact a public agency; the separate development of different channels of communication available for a single service (multi-channel delivery) has resulted in inconsistencies of interfaces. To overcome the drawbacks of multiple-channel service delivery, the different channels should be integrated and coordinated. Creating ISD has advantages for both the government and its clients.

3.2. Social media

Social media is regarded as a component of the Web 2.0 movement, which is distinguished by user-generated content, online identity building, and relational networking. Social media programs such as Facebook™ and Google+™, microblogging services such as Twitter™, blogs, wikis, and media sharing sites such as YouTube™ and Flickr™ are all examples of social media [23]. According to Landsbergen [16] the social media is a set of tools that serve several social communication needs; a communication tool rather than an IT application; a tool that supports communication within social networks.

Social media is rapidly growing, because it supports some important social needs. Governments will need to use social media in the process of service delivery because of the following two reasons:

- firstly, social media as a tool: 1) allows individuals to use human and other networks more easily; 2) expects interactive rather than broadcast communications; 3) is powerful because it uses "multimedia" content (video and audio), not only text; and 4) relies on measures and objectives to facilitate communication. Its effectiveness comes from allowing two-way, interactive communication. People desire quick access to the information they require and to the relationships they wish to build.
- Secondly, metrics are a critical point, since they are an integral component of this culture. Messages, postings, and communications, are all counted by people in various groups. These network-influence indicators emphasize the importance of the message and the communicator. The open-source nature of the technology allows for the collection of this data.

3.3. Proactive public services

Public services are defined as services provided to the public to address social or economic ills or issues, as well as any instance in which citizens, businesses, or others, interact with the administration and some form of information or financial exchange occurs [21].

Proactive public services are services that a public organization "pushes" toward citizens or businesses based on their needs, circumstance, personal preferences, life events and location [17]. A public sector organization operates proactively when it approaches the recipient before the recipient contacts the public sector organization [25], thus acting ahead of time.

As proactive public services can be considered a next step in digital government development [17], they are usually based on digital public services, services that are delivered or mediated via digital technology.

4. Macedonian Case Study

Macedonia's adoption of e-government, and thus of the means of realizing service delivery, is quite disintegrated and unclear, as in many other countries that are in the middle of the process of e-government development. In such countries there are a variety of avenues for information sharing, through both traditional and online channels, although the traditional one remains dominating. Through different phases of e-government development there are attempts for online public service delivery: each institution has its own website with online services; there is a One-Stop-Shop - www.uslugi.gov.mk; but also e-government kiosks have been constructed in the larger towns. Some of the online services just provide information, such as contact information (e.g. information for phone calls, as a method of communicating with government institutions, and e-mail address), but others provide different online applications or tools. It is very rare to have a complete online service delivery. Within the context of e-health, Macedonians receive individualized service: for instance, when they have an appointment with a specialist they receive a SMS reminder. In general, the majority of service delivery still happens through traditional means (e.g. institutional front desk and paper documents). The TV and the radio are the most commonly utilized channels by the government, to deliver information to the citizens (one-to-many), and sometimes social media are employed, too. The One-Stop-Shop option is weak, and mostly information-based. Moreover, if some examples of complete online service delivery exist, they are not user friendly. Some websites should allow citizens to simply apply for scheduled terms to approach institutions, but this functionality is frequently out of order and poorly organized (e.g. the citizen must manually check which term is free - the system does not automatically present free terms, which leads to an unnecessarily time-consuming research for free terms).

This study presents an uncommon circumstance that occurred in the Republic of North Macedonia. As a result of the country's renaming, it is required by the law that all citizens' personal documents, regardless of their expiration date, be replaced with new ones, from May 2023 to February 2024. As a result, the Ministry of Internal Affairs (MIA) is required to issue personal paperwork such as ID cards, passports, and driving licenses to all Macedonians. This places a significant strain on the institution's capabilities, especially since many Macedonian residents living abroad visit Macedonia over the summer and, as a result, must apply for, and receive, these public services. It is critical to remember that these services are sequential: first, ID cards, then passports. Each document issuing procedure takes an average of 10 days.

All service delivery begins with informing, especially when it is not the typical issuing of documents, as in this case. The procedure of issuing documentation necessitates applying for a term, filling out an application in person in MIA premises, being present at the institution for taking photos, waiting for documents to be issued, and physically being present at the institution to collect the personal document. The overall process of issuing these services was as follows: citizens were informed by the government mainly utilizing TV and radio as traditional tools, and online information was also employed, via the ministry of Internal Affairs' website. Social media platforms such as Facebook and Twitter were used, too. At the beginning, to take an appointment to visit MIA it was necessary to use a tool on the Ministry of Interior Affairs' (MIA) website, that was, however, not sophisticated. This function has stopped working in the past 3-4 months, shortly after the law went into effect, and now, at its place, the MIA's website provides information regarding phone calls as a channel for service dating. The number found on the website refers to a call center that provides assistance to MIA, but getting a hold of service is often problematic. A large number of citizens confronted with this circumstance visit MIA offices in search of information, or to schedule appointments. Furthermore, another absurd situation takes place - a citizen who needs to change all three personal document, in an ideal time table of 1 month, needs to visit the MIA facilities minimum three times.

5. Conclusion and Contributions

A number of public services, such as personal documents (ID, passport, driving license), have limited duration. This characteristic allows such public services an opportunity to potentially achieve the highest

level of sophistication of online services in e-government. To reach the so-called personalization phase of e-government development, the government must "take care" of the day of expiry of documents and notify individuals about it on time. Attaining this level of sophistication in public sector delivery relates to a higher level of e-government development known as smart government. Smart governments shift their function from reactive to proactive, which means they initiate and execute services without prior engagement or application by residents. This kind of situation requires the governments to find a cheap, but also very close communication channel to the citizens, closest to their habits.

Social media has become a platform that is easily accessible to anyone with Internet connection, and the favorite communication channel for a large number of people. So, it has opened the door for governments to think of utilizing such channel in a strategic way and to benefit from social media in the process of service delivery as part of the ISD. This is especially related to public services delivery that can lead the government to become a smart government by interacting with its citizens in an effective and efficient way, according to citizens' habits of using technology. According to Dadashzadeh [8] a distinct approach is required, which means that governments' use of social media should be organized; governments should focus on simplification and collaboration and,by using the most widespread tools, such as social media, it should provide efficient client-focused services, though electronic delivery. In the strategy of creation of ISD Scholta and Lindgren [26] suggest that public organizations need to be proactive in digital public services delivery, where the organizations have to approach their clients, rather than the other way around.

The present Macedonian case study is a good example for the identification of obstacles as possible opportunities for improvement. The lack of communication, need of two-way communication, user friendly application form, and approach to service delivery through technological tools that citizens habitually use, suggest that Macedonian e-government initiatives need to provide significant insights linked to the integration of social media in the integrated delivery of public services within e-government.

The purpose of this study is to support and guide government officials, policymakers, and researchers in making decisions on the strategic incorporation of social media in e-government service delivery. The SMS reminder example in Macedonian e-health demonstrates that Macedonian e-government has the potential to be proactive if it uses the most generally widespread communication technologies, which are already a part of the population's daily life, and are essentially inexpensive to use. The goal is to encourage IT experts to explore ways to include social media into integrated public service delivery, for instance, to create a tool to reach citizens with reminders; through such innovations the country would finally become smart. This case study can be used as an excellent example that contributes to existing theory, since it points out the problems that exist in digital public services and currently keep them from achieving a high level of sophistication, and call for joint research by scholars of public administration, information systems, and service management, to relate the research streams of administrative burden and proactive digital public service.

This paper will contribute to the body of overall literature and knowledge on the incorporation of social media platforms into the ISD process. The findings will shed light on the potential benefits of increased involvement, improved communication, and streamlined service delivery as a result of social media integration into the service delivery system. Future research is expected to show how integration can streamline application processes, promote inclusion, and improve overall service delivery by demonstrating the capabilities of social media toward achieving the 'human-centered digital transition' that Ursula von der Leyen, President of the European Commission, believes in [7]. In that direction, the challenge for the next paper can be answering to the question "How can social media enable governments to enact a better proactive service delivery in a way that they have never done before?"

References

[1] Adae, E. K., Twum, K. K., Hinson, R. E., Duh, H. I., and Odame, D. A. 2023. Social Media and E-Government Services in Africa, In book: Social Media and Africa's Public Sector (pp.103-118), April 2023, DOI:10.1007/978-3-031-22642-7_6

[2] Alarabiat, A., Soares, D. S. and Estevez, E. 2016. Electronic Participation with a Special Reference to Social Media-A Literature Review. In International Conference on Electronic Participation 41-52. Springer International Publishing.

[3] Ayachi, R., Boukhris, I., Mellouli, S. 2016. Proactive and reactive e-government services recommendation. Univ Access Inf Soc 15, 681–697. https://doi.org/10.1007/s10209-015-0442-z

[4] Bharosa, N., Janssen, M., Klievink, B., van Veenstra A-F., and Overbeek, S. 2010. Guiding Integrated Service Delivery: Synthesizing and Embedding Principles Using Role-Playing Games. Electronic Journal of e-Government Volume 8 Issue 2, 2010, pp. 83-92

[5] Bertot, J.C.; Jaeger, P.T.; Grimes, J.M. 2010. Using ICTs to create a culture of transparency: E-government and social media as openness and anti-corruption tools for societies. Gov. Inf. Q. 2010, 27, 264–271.

[6] Bogdanoska Jovanovska, M., Erman N., and Todorovski Lj. 2013. Indicators of intensity and development of e-government back-office. Modern approaches, metrics and indicators for monitoring and evaluating public policies. Monograph. Faculty of Administration, University of Ljubljana, 203-228.

[7] eGovernment Benchmarking 2022. Synchronizing digital governments. Insight Report written by Capgemini, Sogeti, IDC and Politecnico di Milano for the European Commission Directorate-General for Communications Networks, Content and Technology July – 2022

[8] Dadashzadeh, M. 2010. Social media in government: From eGovernment to eGovernance. J. Bus. Econ. Res. 2010, 8, 81–86. 36.

[9] Díaz-Díaz, R. and Pérez-González, D. 2016. Implementation of Social Media Concepts for e-Government: Case Study of a Social Media Tool for Value Co-Creation and Citizen Participation. Journal of Organizational and End User Computing Volume 28, Issue 3, July-September 2016.

[10] Dwivedi, K. Y., Rana, P. N., Tajvidi, M., Lal, B., Sahu, G. P., and Gupta, A. 2017. Exploring the Role of Social Media in e-Government: an Analysis of Emerging Literature. ICEGOV '17, March 07-09, 2017, New Delhi , AA, India © 2017 ACM. ISBN 978-1-4503-4825-6/17/03 DOI: http://dx.doi.org/10.1145/3047273.3047374

[11] Goncalves, J., Liu, Y., Xiao, B., Chaudhry, S., Hosio, S., & Kostakos, V. 2015. Increasing the reach of government social media: A case study in modeling government-citizen interaction on Facebook. Policy and Internet, 7,1 80-102. doi:10.1002/poi3.81.

[12] Khan, S., Umer, R., Umer, S., Naqvi, S. 2021. Antecedents of trust in using social media for E-government services: An empirical study in Pakistan. Technology in Society Volume 64, February 2021, 101400 https://doi.org/10.1016/j.techsoc.2020.101400

[13] Khan, S., Rahim, N. Z. Ab., Maarop, N. 2021. A systematic literature review and a proposed model on antecedents of trust to use social media for e-government services. International Journal of Advanced and Applied Sciences, 7(2) 2020, pp. 44-56

[14] Kernaghan, K. 2005. Moving towards the virtual state: integrating services and service channels for citizen-centred delivery. International Review of Administrative Sciences, 71(1), 119–131. doi:10.1177/0020852305051688

[15] Klievink, B. and Janssen, M. 2009. Realizing joined-up government — Dynamic capabilities and stage models for transformation. Government Information Quarterly, Volume 26, Issue 2, April 2009, Pages 275-284

[16] Landsbergen, D. 2010. Government as Part of the Revolution: Using Social Media to Achieve Public Goals. Electronic Journal of e-Government Volume 8 Issue 2, 2010 (pp. 134-146)

[17] Linders, D., Liao, C.Z.-P., and Wang, C.-M. 2018. Proactive e-governance: Flipping the service delivery model from pull to push in Taiwan. Government Information Quarterly, 35 (4) (2018), pp. S68-S76, 10.1016/j.giq.2015.08.004

[18] Lindgren, I., Madsen, C.Ø., Hoglund ¨ Ryd´en, H., & Heggertveit, I. 2022. Exploring citizens' channel behavior in benefit application: Empirical examples from Norwegian welfare services. In L. Amaral, D. Soares, L. Zheng, M. Peixoto, & C. Braga (Eds.), Proceedings of the 15th international conference on theory and practice of electronic governance (ICEGOV 2022) (pp. 416–423).

[19] Liu, S. M., & Yuan, Q. 2015. The evolution of information and communication technology in public administration. Public Administration and Development, 35,2, 140-151. doi:10.1002/pad.1717.

[20] Margetts, H., & Dunleavy, P. 2013. The second wave of digital-era governance: a quasi-paradigm for government on the Web. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 371(1987), 20120382–20120382. doi:10.1098/rsta.2012.0382

[21] Osborne, S. P. 2020. Public service logic: Creating value for public service users, citizens, and society through public service delivery. Routledge Critical Studies in Public Management (1st ed.), Routledge, London (2020), 10.4324/9781003009153

[22] Phythian, M. J., & Taylor, W. G. K. 2001. Progress in electronic service delivery by English District Councils. International Journal of Public Sector Management, 14(7), 569–591. doi:10.1108/eum0000000006190

[23] Porumbescu, A. G. 2016. Linking public sector social media and e-government website use to trust in government. Government Information Quarterly**,** Volume 33, Issue 2, April 2016, Pages 291-304, https://doi.org/10.1016/j.giq.2016.04.006

[24] Roya, H. and Taghi, T. M. A Model for Public Service Delivery in the Context of Social Media (E-Government 2.0) for Developing Countries. Available at SSRN: https://ssrn.com/abstract=4386112 or http://dx.doi.org/10.2139/ssrn.4386112

[25] Scholta, H., and Lindgren, I. 2019. The long and winding road of digital public services - one next step: Proactivity H. Krcmar, W.F.B. Fedorowicz, J.M. Leitmeister, S. Wattal (Eds.), Proceedings of the 40th International Conference on Information Systems (ICIS 2019), Association for Information Systems, Munich (2019)

[26] Scholta, H. and Lindgren, I. 2023.Proactivity in digital public services: A conceptual analysis. Government Information Quarterly, Volume 40, Issue 3, June 2023, 101832 https://doi.org/10.1016/j.giq.2023.101832

[27] Tursunbayeva, A. Franco, M. and Pagliari, C. 2017. Use of social media for e-Government in the public health sector: A systematic review of published studies. Government Information Quarterly, Volume 34, Issue 2, April 2017, Pages 270-282. https://doi.org/10.1016/j.giq.2017.04.001

[28] United Nations 2016. E-government survey. E-government in support of sustainable development.

Social Media Algorithms and their Impact on Sustainable Internet Marketing Strategies

Mihalj Bakator[1*], Dragan Ćoćkalo[2], Mila Kavalić[3] ,Sanja Stanisavljev[4] and Edit Terek-Stojanović[5]

[1] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*
[2] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*
[3] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*
[4] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*
[5] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*

mihalj.bakator@tfzr.rs; dragan.cockalo@tfzr.rs; mila.kavalic@tfzr.rs; sanja.stanisavljev@tfzr.rs; edit.terek@tfzr.rs

Abstract:
In this paper, the authors analyze social media algorithms and how they affect Internet marketing strategies. In the modern business environment, enterprises must consider multiple factors in the market. When developing Internet marketing strategies, the concept of sustainability should be considered, as without it, the volume of new marketing strategies would increase over time, increasing marketing costs. Therefore, effective and long-term strategies are the potential key to success on the Internet. However, social media algorithms pose a challenge when it comes to communicating online, and creating content for customers. In this paper, the goal is to analyze social media algorithms and modern Internet marketing strategies in order to develop a theoretical model. This model aims to concisely present the dynamics between social media algorithms, Internet marketing strategies, and enterprise competitiveness. The methodology used was theoretical synthesis. The data/studies were acquired through databases such as Scopus, WoS, DOAJ, JStor and other. The paper contributes to the existing body of literature and provides an appropriate basis for future studies in this domain. This type of synthesis is not common, but it highlights the complexity of relations between the analyzed subjects.

Keywords:
Social media, algorithms, sustainability, Internet marketing strategies

1. Introduction

In the Digital Age, social media platforms have emerged as dominant communication channels influencing public discourse, consumer behavior, and cultural trends. Their significance has transcended interpersonal communication to become invaluable tools for marketing, brand positioning, and customer engagement. While traditional forms of marketing - such as television ads, print media, and billboards - still play a role in advertising strategies, the focus has increasingly shifted toward leveraging the dynamic environment of social media platforms [1]. A central aspect driving this transition is the algorithms employed through these platforms, which serve as the underlying mechanism for content dissemination and user interaction. Although social media algorithms are complex, adaptive, and often opaque, their impact on Internet marketing strategies can be profound and paradoxical.

On the one hand, algorithms enable unprecedented targeting capabilities that allow brands to engage with particular user demographics [2, 3]. On the other, the ever-changing nature of these algorithms can pose a challenge for sustainable Internet marketing, demanding constant adaptation to ensure long-term visibility and engagement. Additionally, there are ethical and social considerations to account for, such as algorithmic bias and information silos, which have potential ramifications for societal norms and

values [4]. The current body of literature is broad regarding addressing social media, sustainability, and Internet marketing strategies.

Previous studies in this domain noted that social media algorithms play a pivotal role in determining what content users see on their feeds. These algorithms are designed around the primary objective of maximizing user engagement by analyzing user interactions, the relevancy of the content, and various other factors. While they have reshaped the dynamics of content distribution and made it possible for content creators to reach audiences globally, they come with challenges. One major concern is the creation of echo chambers, where users are repeatedly exposed to similar content, leading to a narrow worldview and potentially facilitating the spread of misinformation [5].

Further, the algorithms' lack of transparency raises ethical questions. There's a growing demand for platforms to make their algorithms more transparent and to strike a balance between content personalization and diversity. This could prevent the undue concentration of influence and ensure users are exposed to a broader spectrum of information, helping to combat the dangers of selective exposure [6].

The digital age, while offering myriad benefits, also raises concerns about sustainability. Every digital activity, from sending an email to streaming a video, has a carbon footprint due to the vast digital infrastructures like data centers and networks that power the internet. As our reliance on these digital services grows, so does the environmental impact. It becomes imperative to assess the sustainability of our increasing digital consumption and its impact on the environment [7].

Addressing this challenge requires a dual approach. First, there's a need to transition digital infrastructure to renewable energy sources, reducing the carbon footprint of digital operations. Second, optimization of digital processes and algorithms can lead to reduced computational loads, translating into lower energy consumption. Policy-makers, tech innovators, and stakeholders must collaborate to make the digital ecosystem sustainable [8].

The landscape of Internet marketing has seen a dramatic transformation with the advent of data analytics, artificial intelligence, and evolving digital platforms. Modern marketing strategies are data-centric, allowing businesses to target audiences with unparalleled precision. This personalization, driven by user behavior and preferences, enables companies to enhance their outreach and engagement. However, this precise targeting brings its own set of challenges [9].

The line between personalization and invasion of privacy is thin. With increasing data breaches and growing awareness about data privacy, there's a pressing need for transparent data handling and ethical marketing practices. As the digital age progresses, businesses must constantly adapt and ensure their strategies align with not only technological advancements but also the evolving values and preferences of consumers [10].

However, fewer studies analyze social media algorithms in the context of sustainable Internet marketing strategies and campaigns. This paper aims to fill this gap by qualitatively analyzing the relationships between social media algorithms and sustainable Internet marketing strategies. The goal is to analyze how algorithms shape user behavior, dictate content visibility, and, in turn, influence marketing tactics and outcomes. Further, strategic adaptations essential for sustaining brand engagement in a volatile algorithmic landscape are addressed. Finally, the ethical dimensions and long-term implications of relying on algorithmically curated platforms for Internet marketing are considered, suggesting a framework for sustainable and responsible practices. Understanding this relationship is crucial for scholars, marketers, policymakers, and platform developers.

The paper has three main sections (excluding the Introduction and Conclusion). The first section addresses social media algorithms. Next, sustainable Internet marketing strategies are presented alongside the developed theoretical model. Finally, suggestions and guidelines for improving enterprise competitiveness are given.

2. Social media algorithms

Social media algorithms are essentially sets of rules or a sequence of instructions designed to perform a specific task - in this case, determining which content appears in the user’s feed and in what order. They analyze numerous variables such as the kind of content the user likes, shares, or comments on, who the user interacts with the most, and even how much time the user spends on a particular post

[11]. This information is then used to create a personalized feed aimed to keep the users engaged on the platform for as long as possible. Originally, social media feeds were largely chronological, showing the most recent posts first. However, as these platforms grew and the amount of content increased, it became harder for users to see everything that might be relevant to them. That's where algorithms come in. They filter the content to show what the platform thinks is the most interesting or engaging. This makes it easier to find content that is enjoyable but also means the user is less likely to see posts that don't fit the algorithm's prediction of its interests [12, 13].

The primary goal of these algorithms is to maximize user engagement. The more the user interacts with the platform, the more ads the users see, and the more data can be collected about their behavior. This model is advantageous for businesses advertising on social media, but it also has its drawbacks. For instance, these algorithms can create "echo chambers," where users are only exposed to viewpoints similar to their own, contributing to increased polarization. Algorithms are also constantly evolving. Social media platforms frequently update their algorithms to better capture user behavior or to introduce new features. This means that what worked yesterday in terms of gaining visibility or engagement on a post may not work tomorrow, requiring content creators and marketers to continually adapt their strategies. Another criticism is that algorithms prioritize certain types of content over others, often favoring sensational or controversial material that may not be factual but generates a lot of user interaction. This has raised ethical concerns about the role of social media in disseminating information and shaping public opinion [14].

At their core, social media algorithms are complex computational models that leverage machine learning, data analytics, and user behavior metrics to curate and personalize content [15]. While the specific mechanisms can differ between platforms like Facebook, Instagram, Twitter, TikTok and others, some general principles guide how these algorithms work. These are [16, 17, 18]:

- Data Collection and User Profiling: The first step is collecting data. Every time the user likes a post, shares content, follows a new account, or even spends time watching a video, this data is collected. Some platforms also track behavior outside their app through cookies and other tracking mechanisms.
- Content Ranking: The next step is to use this data to rank content. Not all content is treated equally. Posts are usually assigned a relevance score based on how closely they align with the user’s past behavior and preferences. This score can be influenced by numerous factors such as recency, engagement, relevance, and relationship.
- Personalization and Context: Context also plays a crucial role. Algorithms consider contextual factors like geographical location, the device that is used, and even the time of day to personalize the feed. For example, the user might see more local news stories or posts from local businesses based on the location data.
- Feedback Loops: Feedback loops are another important mechanism. If the users engage with a type of content or a particular user frequently, the algorithm "learns" from this and shows them more of the same, creating a loop that constantly refines their feed based on the most recent behavior.
- A/B Testing: Social media platforms frequently conduct A/B tests to evaluate the effectiveness of new algorithmic features or tweaks. These tests involve showing one set of algorithmic criteria to Group A and a different set to Group B, then measuring which performs better in terms of user engagement or other key metrics.
- Ethical and Societal Implications: While not a "mechanism" per se, it's important to note that algorithms can perpetuate biases present in the data they are trained on or the objectives they are designed to optimize. This can lead to ethical concerns such as the amplification of harmful content, echo chambers, or biased viewpoints.

It's worth noting that most social media platforms are secretive about the specific workings of their algorithms, mainly to prevent a manipulation or "gaming" of the system, but these general principles are widely acknowledged to be in play. Understanding these mechanisms is crucial for anyone looking to maximize their reach on social media, but it's also important for everyday users to understand how their online experience is being shaped.

3. Sustainable Internet marketing strategies

Developing a sustainable Internet marketing strategy that aligns with the dynamic nature of social media algorithms is a challenging but necessary endeavor. One of the foundational principles of a sustainable strategy is content diversification. Given that social media algorithms are designed to respond to user behavior and preferences, a one-dimensional approach to content can leave a brand vulnerable to sudden shifts in algorithmic priorities. Therefore, brands should not only spread their efforts across multiple social platforms but also diversify the types of content they produce. Whether it's long-form articles, short-form videos, or interactive polls, different forms of content will resonate with various segments of the audience and also offer some level of algorithmic unpredictability. In essence, content diversification acts as a hedge against the constantly changing landscape of social media algorithms [19].

Algorithms are increasingly sophisticated, designed to reward content that genuinely engages users rather than merely attracting clicks. The focus should be on producing high-quality, engaging content that adds value to the consumer. This means investing time in understanding the target audience, researching topics that resonate with them, and delivering content that is both compelling and useful. While it may be tempting to flood social media feeds with a high volume of posts, such a strategy could backfire if the content does not meet the quality metrics that algorithms are designed to measure. The dynamics of social media algorithms often reward consistency and frequency. Consistent posting schedules and regular interactions with the audience can improve visibility in their feeds. However, this should not come at the expense of quality. Consistency should be about building a reliable brand image that the audience can expect and look forward to, rather than a predictable volume of output. A calendar of planned, well-thought-out posts can serve as a useful tool to maintain this balance between consistency and quality [20].

Due to the ever-changing nature of social media algorithms, a static strategy is unlikely to remain effective over time. Utilizing analytics tools to monitor metrics like engagement rates, impressions, and conversions in real time can provide invaluable insights. Are certain posts getting more traction at specific times of the day? Are particular topics driving more engagement than others? By answering these questions and dynamically adjusting the strategy, it can remain one step ahead of algorithmic changes [21]. Algorithms also take cues from how users are interacting with content. Simply put, a post that garners more likes, shares, and comments is more likely to be favored by the algorithm. To take advantage of this, marketers should focus on building a strong community around their brand. This involves more than just pushing content; it also requires actively engaging with the audience by encouraging user-generated content, responding to comments, and fostering discussion. As algorithms evolve, they often prioritize new types of content, such as videos or interactive elements like polls and quizzes. Brands should stay abreast of these changes and be ready to integrate new content formats into their strategy. Video content, in particular, has been shown to generate higher engagement rates compared to other forms of content, making it an important component of a sustainable marketing strategy.

While chasing algorithmic favorability, it's crucial not to lose sight of ethical considerations. With the increasing scrutiny of social media's role in public discourse, platforms are taking steps to penalize content that is misleading, inflammatory, or harmful. Brands must ensure that their marketing strategies are not just algorithmically savvy but also socially responsible [22]. Enterprises can better navigate the complexities of social media algorithms by considering these elements in developing an Internet marketing strategy, ensuring not just short-term gains but long-term sustainability.

Furthermore, based on the analyzed literature a theoretical model is developed. The model is presented in Figure 1.

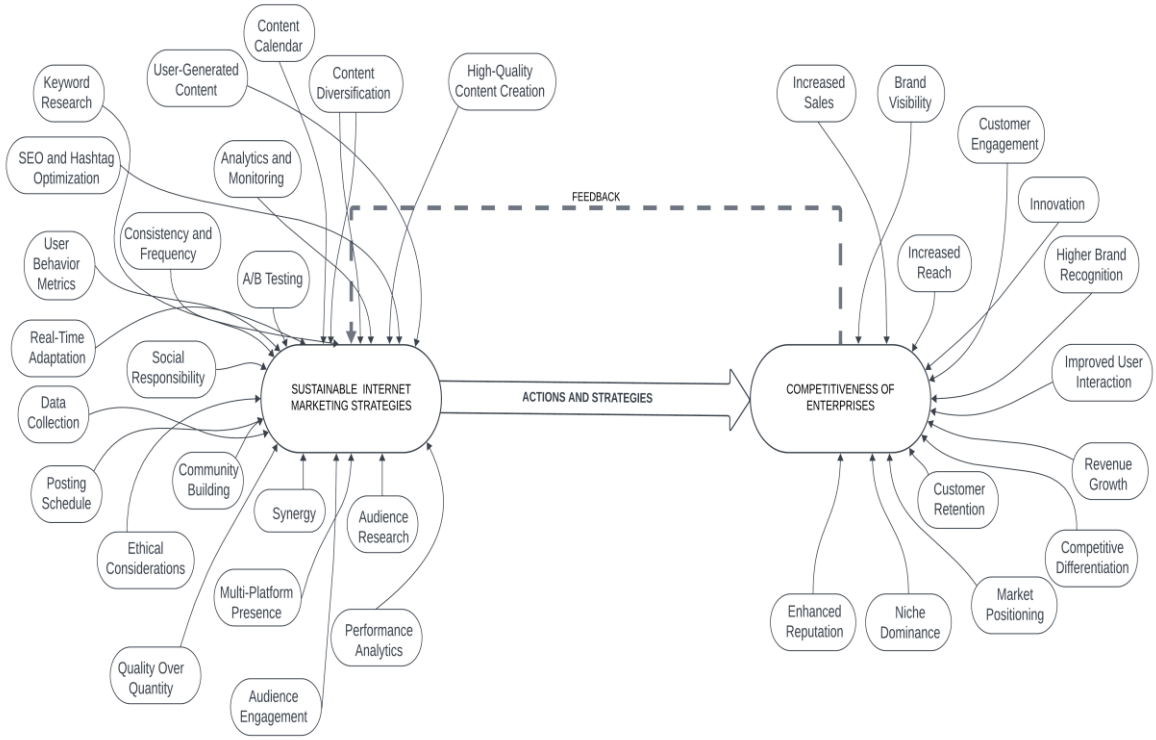


Figure 1: Theoretical model

The model shown in Figure 1 suggests a close connection between sustainability and internet marketing, suggesting that how companies present themselves online directly affects how competitive they are. A significant shift in marketing paradigms has been made with the inclusion of sustainability into the Internet marketing framework. Traditional marketing strategies may become obsolete as businesses and consumers place more value on sustainable business practices. This model suggests that incorporating sustainability into online marketing initiatives has intrinsic value. It implies that businesses that put a priority on sustainable strategies will have a significant competitive advantage as a result, redefining their relationship with customers in terms of loyalty, trust, and perceived value. Additionally, it suggests a potential shift in how resources are allocated within businesses, with an emphasis on more environmentally friendly tools and strategies rather than aggressive, non-sustainable online marketing strategies. In actual operations, businesses can make use of this model. The organization's broader vision and objectives can be set in motion by rethinking strategic planning to incorporate sustainable principles. Budget allocations can favor sustainable online platforms and campaigns as financial considerations are reevaluated. Employees can receive training to effectively understand and incorporate sustainable internet marketing strategies. This model also provides a framework for open stakeholder communication, highlighting the dedication to sustainability of an organization. On a more practical level, market analysis supported by this model can help businesses stay flexible and adapt their strategies to the state of the market.

4. Suggestions and guidelines

Based on the qualitative analysis of literature in the domain of social media algorithms, enterprise competitiveness, and sustainable Internet marketing strategies, the following suggestions and guidelines for improving enterprise competitiveness are noted:

- Expanding a brand's presence across multiple social media platforms can safeguard it against the volatility of algorithmic changes on any individual platform. Different social media sites have unique algorithms and user demographics, so a multi-platform approach maximizes reach and adaptability.

- Content diversification is another effective tactic for increasing a brand's algorithmic resilience. Incorporating a variety of content types—videos, articles, images, and interactive elements like polls—can resonate with different segments of an audience and introduce a level of unpredictability that can be advantageous in navigating algorithmic preferences.
- While the frequency of posting can increase visibility, the quality of content remains paramount. Regular audience surveys and engagement analytics can offer insights into consumer preferences and behaviors. High-value, resonant content is more likely to be shared and engaged with, which in turn, boosts its algorithmic favorability.
- Maintaining a consistent posting schedule can further enhance visibility and engagement. Utilizing a content calendar allows for well-planned posts that coincide with peak audience activity. However, this schedule should remain flexible enough to allow for real time adaptation to trending topics or events.
- In terms of real time strategy adjustments, analytics tools that monitor key performance indicators such as engagement, reach, and ROI are invaluable. These insights enable agile strategy pivots, optimizing the effectiveness of social media campaigns. A/B testing of different post types, timing, and content can offer additional data for refining marketing strategies.
- Community building is a crucial component of a competitive strategy. Direct engagement through live sessions, webinars, and Q&As, as well as encouraging user-generated content, can leverage broader networks and enhance brand reach.
- Ethical considerations should not be compromised in the pursuit of aggressive marketing tactics. Full disclosure of sponsorships, affiliate relationships, and other potential conflicts of interest is essential for maintaining brand integrity and consumer trust.
- For sustained competitiveness, periodic keyword research and the effective use of hashtags are indispensable. While these elements are basic, they ensure that content remains discoverable and reaches targeted demographics.
- Creating easily consumable content can favorably position a brand within social media algorithms. Such content captures attention quickly and encourages user interactions, making it more likely to be promoted through algorithms. Continuous education on the latest algorithmic changes is necessary for ongoing adaptability.

Finally, a well-coordinated, multi-platform marketing strategy that aligns with broader marketing goals can result in a cohesive and more effective brand experience. Creating a synergy between online and offline activities not only maximizes reach but also builds a robust and sustainable competitive strategy.

5. Conclusion

In conclusion, the interplay between sustainable Internet marketing strategies and social media algorithms is a dynamic and evolving landscape that has significant implications for the competitiveness of enterprises. The multi-faceted approach discussed herein emphasizes the need for diversification, quality, ethical considerations, and real time adaptability. The utilization of these strategies in synergy with a deep understanding of social media algorithms can result in increased brand visibility, customer engagement, and ultimately, a more competitive market position.

The strategies suggested offer a comprehensive yet adaptable framework that accounts for the volatile nature of social media algorithms. They reflect the need for ongoing monitoring and adaptability, backed by solid ethical considerations to ensure long-term sustainability and public trust. However, the rapid changes in algorithms and audience behavior necessitate continuous research and adjustment of these strategies for them to remain effective over time.

Future research could focus on several pertinent avenues. Longitudinal studies examining the efficacy of these strategies across different industry sectors would offer valuable insights into their universal applicability or the need for sector-specific modifications. Experimental designs involving A/B tests could provide empirical data on the short-term and long-term effectiveness of particular strategies, thereby contributing to a more nuanced understanding of algorithmic behavior. There is also a need for more qualitative research exploring consumer perceptions and attitudes toward

algorithmically curated content, which can offer a consumer-centric perspective to augment current enterprise-centric approaches.

Acknowledgement

This paper has been supported by the Provincial Secretariat for Higher Education and Scientific Research of the Autonomous Province of Vojvodina, number: 142-451-2963/2022-03.

References:

[1] A. Rasool, F. A. Shah, J. U. Islam, Customer engagement in the digital age: A review and research agenda. Current Opinion in Psychology, 36 (2020), 96-100. https://doi.org/10.1016/j.copsyc.2020.05.003

[2] T. K. Balaji, C. S. R. Annavarapu, A. Bablani, Machine learning algorithms for social media analysis: A survey. Computer Science Review, 40 (2021), 100395.

[3] J. Swart, Experiencing algorithms: How young people understand, feel about, and engage with algorithmic news selection on social media. Social Media+ Society, 7 (2021), 20563051211008828.

[4] H. Taherdoost, Enhancing Social Media Platforms with Machine Learning Algorithms and Neural Networks, Algorithms, 16 (2023), 271. https://doi.org/10.3390/a16060271

[5] M. Petrescu, A. S. Krishen, The dilemma of social media algorithms and analytics, Journal of Marketing Analytics, 8, (2020), 187-188. https://doi.org/10.1057/s41270-020-00094-4

[6] M. A. DeVito, D. Gergle, J. Birnholtz, Algorithms ruin everything" # RIPTwitter, Folk Theories, and Resistance to Algorithmic Change in Social Media, In Proceedings of the 2017 CHI conference on human factors in computing systems, (2017) (pp. 3163-3174). https://doi.org/10.1145/3025453.3025659

[7] S. Schneider, The impacts of digital technologies on innovating for sustainability, Innovation for Sustainability: Business Transformations Towards a Better World, 415-433 (2019). https://doi.org/10.1007/978-3-319-97385-2_22

[8] G. Samuel, F. Lucivero, L. Somavilla, The environmental sustainability of digital technologies: Stakeholder practices and perspectives, Sustainability, 14 (2022), 3791. https://doi.org/10.3390/su14073791

[9] K. O. Kayumovich, F. S. Annamuradovna, The main convenience of internet marketing from traditional marketing, Academy, 1 (2020), 30-32. https://doi.org/10.24411/2412-8236-2020-10101

[10] O. K. Khurramov, & S. A. Fayzieva, The main convenience of internet marketing from traditional marketing, Academy, 1 (2020), 30-32.

[11] A., Al Marouf, M. K., Hasan, H. Mahmud, Comparative analysis of feature selection algorithms for computational personality prediction from social media, IEEE Transactions on Computational Social Systems, 7 (2020), 587-599. https://doi.org/10.1109/TCSS.2020.2966910

[12] A. Pasumponpandian, A hybrid-algorithm for E-project selection on social media, Journal of Information Technology, 2 (2020), 116-122. https://doi.org/10.36548/jitdw.2020.2.00

[13] F., Zimmer, K., Scheibe, M., Stock, W. G. Stock, Echo chambers and filter bubbles of fake news in social media. Man-made or produced by algorithms. In 8th annual arts, humanities, social sciences & education conference, (2019), 1-22.

[14] B. E., Duffy, C. Meisner, Platform governance at the margins: Social media creators' experiences with algorithmic (in) visibility. Media, Culture & Society, 45 (2023), 285-304. https://doi.org/10.1177/01634437221111923

[15] D. Ognibene, R. Wilkens, D. Taibi, D. Hernández-Leo, U. Kruschwitz, G. Donabauer, S. Eimler, Challenging social media threats using collective well-being-aware recommendation algorithms and an educational virtual companion, Frontiers in Artificial Intelligence, 5 (2023), 654930.

[16] B. Zarouali, T. Dobber, G. De Pauw, C. de Vreese, Using a personality-profiling algorithm to investigate political microtargeting: assessing the persuasion effects of personality-tailored ads on social media, Communication Research, 49 (2022), 1066-1091. https://doi.org/10.1177/0093650220961965

[17] D. Shah, A. Singh, S. S. Prasad, Sentimental Analysis Using Supervised Learning Algorithms, In 2022 3rd International Conference on Computation, Automation and Knowledge Management (ICCAKM), (2022), 1-6, IEEE.
[18] D., Yanti, A. D., Subagja, S., Nurhayati, S. R. I., Rezeki, C. H., Limbong, R. S. Hamid, Short Videos & Social Media Algorithms: Effective Communication in Tourism Marketing, International Journal of Artificial Intelligence Research, 6 (2023).
[19] E. M. Denga, N. R. Vajjhala, S. Rakshit, The Role of Digital Marketing in Achieving Sustainable Competitive Advantage, Digital Transformation and Internationalization Strategies in Organizations, 44-60 (2022). https://doi.org/10.4018/978-1-7998-8169-8.ch003
[20] S. E. Najafi, H. Nozari, S. A. Edalatpanah, Investigating the Key Parameters Affecting Sustainable IoT-Based Marketing. In Computational Intelligence Methodologies Applied to Sustainable Development Goals, (2022), 51-61. Cham: Springer International Publishing.
[21] H. Nozari, J. G. Nahr, The Impact of Blockchain Technology and The Internet of Things on the Agile and Sustainable Supply Chain, International Journal of Innovation in Engineering, 2 (2022), 33-41.
[22] F. J. S. Lacárcel, Main uses of artificial intelligence in digital marketing strategies linked to tourism, Journal of Tourism, Sustainability and Well-being, 10 (2022), 215-226.

Modeling in Social Systems Meet Artificial Intelligence

Kalistrat Sandjakoski [1], Ljubinka Gjergjeska Sandjakoska [2]

[1] *University Ss. Cyril and Methodius Skopje, Institute for Sociological, Political and Juridical Research, Partizanski odredi, Skopje, Macedonia*
[2] *University St. Paul the Apostle Ohrid, Faculty of Computer Science and Engineering, Partizanska bb, Ohrid, Macedonia*

kalistrat_s@hotmail.com;ljubinka.gjergjeska@uist.edu.mk

Abstract:

Since artificial intelligence is a new reality, penetrating all spheres of modern existence, the need for a deep understanding of the relationship between it and social systems is more than needed. The objective of this study is to depict the relations between social systems modeling and artificial intelligence, present in both directions. The aim is to shed light on factors that have influenced socio-technological and ethical trends. Moreover, this paper presents an understanding of human-centered artificial intelligence, followed by a framework for an artificial intelligence approach to modeling the dynamics of social systems. The proposed approach tends to answer the following research question: how can artificial intelligence contribute to social systems modeling? The analysis revealed that artificial intelligence can be viewed and used as an effective supporting tool in sociology. This research offers significant and timely insight into the application of technology in the social sciences.

Keywords:
artificial intelligence, social systems modeling, human-centered artificial intelligence

1. Introduction

Since artificial intelligence (AI) is a new reality penetrating all spheres of modern existence, the need for a deep understanding of the relationship between it and social systems is more than needed. The use of the term artificial intelligence has evolved over the years. In the beginning, AI was used to differentiate the new research from the prevailing paradigm of cybernetics [2]. How AI is defined differs across a wide range of contexts [3]. Common knowledge includes the deployment of computing infrastructure and programming code to create systems expected to mimic, augment, or displace human agency. In general, the definitions of AI can be organized into four categories [4]. The first category is related to the phenomenon of acting humanly, which includes: the ability to communicate successfully; to store what it knows or hears (knowledge representation); to use the stored information to answer questions and draw new conclusions (automated reasoning); to adapt to new circumstances and to detect and extrapolate patterns (machine learning); to perceive objects (computer vision); to manipulate objects and move about (robotics). The second category is thinking humanly, which is based on the cognitive modeling approach. The third group of definitions, based on thinking rationally, lay down the "laws of thought" approach. The last is acting rationally, which introduces the rational agent approach. Although there are some key differentiations between these categories of definitions, their synergy covers a wide area of implementation. The justification for the use of AI can be presented in three groups of arguments [5]. The first group is based on the automation of important but repetitive and time-consuming tasks, allowing humans to focus on higher-value work. Next is the ability of AI to reveal insights into massive amounts of unstructured data that once required human management and analysis, such as data generated by videos, photos, written reports, business documents, social media posts, or e-mail messages. The third group of benefits is related to the integration of thousands of computers and other resources to solve the most complex problems. Consequently, AI capabilities should be leveraged to find ways to solve complex problems in social systems. All of the features that create the contemporary content of AI have strong relationships with social phenomena. Combining the

principles of behavioral science and cultural studies to analyze and simulate human behavior forms the core of behavioral-cultural modeling in social systems.

Nowadays, empirical research relies on data collection, and mathematical and computational models are used to make predictive models of individual actions, group dynamics, decision-making processes, etc. Computational modeling of social behavior, including models of social networks, social influence, social learning, and social dynamics, is a research trend. Key questions that scientists in sociology tend to answer are: why to do computer modeling and how to do it? The answers to these questions are placed on several pillars. One of the pillars is the complexity of social systems, focusing on risks and consequences. Since there are risks and consequences, the output cannot be anticipated based on common sense and experience. Next are the feasibility and ethics of doing experiments with social systems where humans are at the center. As a justification for using computers in social modeling, there is a need for the integration of reliable knowledge from different sources into a more complex way to understand the interactions between them. All of these can be indications of the different purposes a model in sociology may have, such as prediction or forecasting, explanation or exploration, understanding theory or designs, and illustration or visualization. [6]

The objective of this study is to depict the relations between social systems modeling and artificial intelligence, present in both directions. The aim is to shed light on factors that influence socio-technological and ethical trends. Moreover, this paper presents an understanding of human-centered AI, followed by a framework for the AI approach in modeling the dynamics of social systems.

The rest of the paper is organized as follows: In the next section, the contribution of this research is presented. That section is organized into three parts. The first part gives a summary of human-centered AI. The second refers to the modeling of social systems, followed by part of the section that discusses the influence of factors on socio-technological and ethical trends. The third section proposes an AI approach for social modeling. The paper ends with concluding remarks.

2. Contribution

We believe that the scientific contribution of this paper will open a wider understanding of interdisciplinary research in the application of AI in sociology. The discussion in this section is organized into the following parts: human-centered AI, social system modeling, and influence factors on socio-technological and ethical trends

2.1. Human-Centered Artificial Intelligence

Human-centered AI is not a new concept in the scientific community. It arises as a response to the ethical and practical challenges posed by AI technologies. This understanding has a crucial role in shaping the responsible development and deployment of AI systems. If the systems are designed, developed, and deployed to augment human capabilities and values, then the center is obvious. In addition, the ability to empower individuals and organizations with AI technologies to solve problems, make informed decisions, and achieve their goals and the goals of society in general supports the human-centricity understanding of AI. The continuation of the improvement of AI systems based on user feedback is another indicator for the justification of this view, which is in consideration of classical sociological theory about the uniqueness of human beings. The hypothesis of human distinctiveness, consequently, is not rejected but expanded and elaborated [8]. Exploiting AI techniques to promote the development of sciences and accelerate their applications to benefit human beings, society, and the world is discussed in [10]. Researchers are addressing concerns of traditional importance within sociology, such as the bases for cooperation, and the role of structure in affecting individual agency, and interaction, using computational models of intelligent adaptive agents. [11]

In the literature, there are five co-existing perspectives on human-centered AI [7]. The first perspective is *deficit-oriented*, viewed in terms of attention, concentration, exercise, and fatigue [12, 13, 14, 15]. The second perspective is a *data reliability-oriented* understanding that copes with the deficits of AI with regard to explaining the ability and trustworthiness of AI in the light of fairness and unbiased data [16, 17, 18]. The next is a *protection-oriented* design that focuses on human-centered design and ergonomics [19, 20]. The fourth perspective, *potential-oriented*, is based on the use of hybrid

intelligence with the human-in-the-loop while leveraging human work potential through AI [21, 22]. The fifth, a *political-oriented understanding* [23, 24], shows how to reach human centricity while using AI in the workplace.

2.2. Social systems modeling

System modeling is one of the most important tasks that should be solved during the simulation of the dynamics of social systems. In general, elements that are included in the modeling process are entities or agents, interactions, rules and norms, environment, feedback mechanisms, goals and objectives, data and information, time dynamics, networks and relationships, outcomes and consequences, sensitivity and uncertainty, validation and evaluation, cultural and contextual factors, and ethical and moral considerations. The complex that all of these elements form is difficult to investigate. That's why, usually, there is some simplification during the simulations. One of the aims of social system modeling is to explore the consequences of different assumptions and scenarios. Also, modeling may involve understanding how individual and collective behaviors change and evolve, including the emergence of patterns and trends.

The modeling of social relations is a trend that silently takes place following two commonly opposed assumptions. The first assumption is that technological development, with special reference to technological objects and their logical application, shapes the individual and interactive behavior of people in the community. The second assumption is based on pragmatic experiments with animals: laboratory conditions are organized, whose working mechanism affects the animal's behavior, tests it, and directs it in a certain direction. The reliability of the latter assumption, which has developed in a special field called social engineering, is intensely questioned by many contemporary scientists. However, we believe that the more prominent penetration of AI into the sphere of society, to a greater or lesser extent, takes place under the sign of the two pre-settings. The progress of a society depends on the positive connection between education and mutual interactions. It is very important to consider not only what kind of literacy and knowledge society will disseminate within itself but also whether quality interactions depend on reliable knowledge. These two intersecting perspectives best show why artificial intelligence should be oriented in the direction of humans and not give them an alibi to suicidally imagine post-humanist dreams.

With the appearance of the most modern and most developed model of computer artificial intelligence, Chat-GPT, many ethical questions are opened, including, above all, the question of the validity of the education that is acquired with the help of this intelligible tool. The quality of mutual interactions, which today are shaken by the increasingly frequent isolation of man in the digital space, depends on the scientific reliability of individual and collective educational processes. One of the ethical challenges facing the scientific community dealing with artificial intelligence in the social field is that this "skillful imitator of the human mind" provides us with data on the subtlest scientific topics without showing us any, not even quasi-valid references. Many suspected that the Google digital search system would lose its modern function simply because it does not have a built-in AI coordinator and AI guide that will walk us through the world of polyvalent data and facilitate our access to it. But Google, unlike Chat-GPT, offers objective intellectual content that is not creatively constructed at the moment within the schemata of elementary associations but gives us reliable and recognized scientific data full of hierarchical references and expedient legal structures. In this sense and at this historical moment, the educational informational content that Google offers us affects society more productively from the perspective of the relationship between education and mutual interactions, which is a fundamental driver and element of social modeling, than the creative and deepened content compiler Chat-GPT. For now, the digital assistant for searching objective, popular, and speculative content has greater advantages than the robot that creates rational schemes based on combining elementary associations because it legitimately promotes the quality division of social systems and communities.

2.3. Influence factors to socio-technological and ethical trends

Understanding how technology and ethics evolve and how they can coexist requires analysis and an interdisciplinary approach. What drives socio-technological and ethical trends is a more general than

specific question related to the interest domain and its complexity. The interplay of factors can be organized into several groups (Table 1).

Table 1:
Influence factors to socio-technological and ethical trends

Factor of Influence	Related features
Advancements, Innovations and Emerging technologies	Biotechnology, Nanotechnology, Quantum computing,
Government regulations and policies	Social movements Advocacy groups
Philosophical discourse	Critical thinking, Argumentation
Cultural norms	Society roles, Expectations, Conventions
Economic and demographic factors,	Market forces, Founding availability, Business models
Crisis events	Wars
Environmental considerations	Sustainability
Globalization	Non-physical networking
Media and communications	Social networks

We should note that these sets of factors are not mutually exclusive. There is a deep connection between them, represented by a subset of features that overlap or are shared. Some examples presented in sociological research show that social media discussions shape public opinion and policy decisions.

3. Artificial Intelligence approach

In this research, we are presenting a framework for an AI approach to modeling social systems. We emphasize that here we present only the theoretical approach that we are trying to realize computationally.

Methodology

- Definition of the problem and objectives; identification of the sociological concepts
- Transformation of sociological concepts into well-defined variables
- Choosing adequate system dynamics techniques
- Construction of a set of differential and algebraic equations that describe the dynamics of those variables
- Simulation for establishing the empirical adequacy of the theory being modeled
- Data collection (simulated data and real-world data)
- Data preprocessing (handling missing values, outliers, inconsistencies, encoding categorical variables, standardization, and normalizing numerical features)
- Feature engineering refers to the transformation and creation of new features, if needed.
- Model selection (such as deep learning architecture)
- Model training followed by optimizing the hyperparameters
- Model evaluation, using defined performance measurements
- Model interpretation (how the prediction is done and what that means for the sociological problem defined)
- Fine-tuning (adjustments)

4. Conclusions

The proposed approach tends to answer the following research question: how can artificial intelligence contribute to social systems modeling? The analysis revealed that artificial intelligence can be viewed and used as an effective supporting tool in sociology. This research offers significant and timely insight into technology application in the social sciences, presenting the complementary potential of AI. The modern meeting of social system modeling and AI is presented.

The potential benefits of the proposed methodology depend on how it is implemented in a given sociological context. One possible benefit is *scalability*, which means that the proposed methodology can be applied to small or more complex social systems. Additionally, the AI approach ensures *consistency* in dynamical social modeling, which can contribute to reducing errors that usually occur because of the stochastic nature of sociological systems. Using the methodology, the *documentation* process is easy to do. The limitations of the proposed methodology can be seen in the *compliance* of sociological concepts and the variables that should be defined and somehow measured. The process of quantifying or measuring can also be limited by some factors in society.

The computational realization of the methodology is an open challenge for future work.

References:

[1] Vinciarelli, Alessandro, Maja Pantic, Dirk Heylen, Catherine Pelachaud, Isabella Poggi, Francesca D'Errico, and Marc Schroeder. "Bridging the gap between social animal and unsocial machine: A survey of social signal processing." IEEE Transactions on Affective Computing 3, no. 1 (2011): 69-87.

[2] McCarthy, John. "Review of the question of artificial intelligence." *Annals of the History of Computing* 10, no. 3 (1988): 224-229.

[3] Hoffman, Steve G. "Thinking science with thinking machines: The multiple realities of basic and applied knowledge in a research border zone." *Social Studies of Science* 45, no. 2 (2015): 242-269.

[4] Russell, Stuart J. *Artificial intelligence a modern approach*. Pearson Education, Inc., 2010.

[5] Nishant, Rohit, Mike Kennedy, and Jacqueline Corbett. "Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda." *International Journal of Information Management* 53 (2020): 102104.

[6] Calder, Muffy, Claire Craig, Dave Culley, Richard De Cani, Christl A. Donnelly, Rowan Douglas, Bruce Edmonds et al. "Computational modelling for decision-making: where, why, what, who and how." *Royal Society open science* 5, no. 6 (2018): 172096.

[7] Wilkens, Uta, Christian COST REYES, Tim Treude, and Annette Kluge. "Understandings and perspectives of human-centered AI–a transdisciplinary literature review." *Frühjahrskongress der Gesellschaft für Arbeitswissenschaf, Bochum* (2021).

[8] Wolfe, Alan. "Mind, self, society, and computer: Artificial intelligence and the sociology of mind." *American Journal of Sociology* 96, no. 5 (1991): 1073-1096.

[9] Glukhikh, V. A., S. M. Eliseev, and N. P. Kirsanova. "Artificial Intelligence as a Problem of Modern Sociology." (2022).

[10] Xu, Yongjun, Xin Liu, Xin Cao, Changping Huang, Enke Liu, Sen Qian, Xingchen Liu et al. "Artificial intelligence: A powerful paradigm for scientific research." *The Innovation* 2, no. 4 (2021).

[11] Carley, Kathleen M. "Artificial intelligence within sociology." *Sociological methods & research* 25, no. 1 (1996): 3-30.

[12] Adnan, Hassan Sami, S. Matthews, M. Hackl, P. P. Das, M. Manaswini, S. Gadamsetti, M. Filali, B. Owoyele, J. Santuber, and J. Edelman. "Human centered AI design for clinical monitoring and data management." *European Journal of Public Health* 30, no. Supplement_5 (2020): ckaa165-225.

[13] Albu-Schäffer, Alin Von drehmomentgeregelten Roboterarmen zum intrinsisch nachgiebigen humanoiden Roboter. In: Christiane Woopen und Marc Jannes (Hg.): Roboter in der Gesellschaft.

Technische Möglichkeiten und menschliche Verantwortung. Berlin, Heidelberg: Springer Berlin Heidelberg (Schriften zu Gesundheit und Gesellschaft - Studies on Health and Society, 2), 1–14 (2019)

[14] Tombeil, Anne-Sophie; Ganz, Walter; Kremer, David (2020) Arbeit und künstliche Intelligenz in wissensintensiven Dienstleistungen. In: Gerhard Ernst, Klaus Zühlke-Robinet, Gerhard Finking und Ursula Bach (Hg.): Digitale Transformation: Arbeit in Dienstleistungssystemen. 1. Aufl. BadenBaden: Nomos Verlagsgesellschaft mbH & Co. KG, 165–180

[15] Bond, R.; Mulvenna, M. D.; Wan, H.; Finlay, D. D.; Wong, A.; Koene, A.; Adel, T. (2019) Human Centered Artificial Intelligence: Weaving UX into Algorithmic Decision Making. In: RoCHI, 2019, 2– 9

[16] Adadi, Amina, and Mohammed Berrada. "Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)." *IEEE access* 6 (2018): 52138-52160.

[17] Arrieta, Alejandro Barredo, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador García et al. "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI." *Information fusion* 58 (2020): 82-115.

[18] Ploug, Thomas, and Søren Holm. "The four dimensions of contestable AI diagnostics-A patient-centric approach to explainable AI." *Artificial Intelligence in Medicine* 107 (2020): 101901.

[19] Buxbaum, Hans-Jürgen; Häusler, Ruth (2020) Ladenburger Thesen zur zukünftigen Gestaltung der Mensch-Roboter-Kollaboration. In: Hans-Jürgen Buxbaum (Hg.): Mensch-Roboter-Kollaboration. 1. Auflage. Wiesbaden: Springer Fachmedien Wiesbaden, 293–317

[20] Steels, Luc. "Personal dynamic memories are necessary to deal with meaning and understanding in human-centric AI." In *NeHuAI@ ECAI*, pp. 11-16. 2020.

[21] Akata, Zeynep, Dan Balliet, Maarten De Rijke, Frank Dignum, Virginia Dignum, Guszti Eiben, Antske Fokkens et al. "A research agenda for hybrid intelligence: augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence." *Computer* 53, no. 8 (2020): 18-28.

[22] Wilson, H. James, and Paul R. Daugherty. "Collaborative intelligence: Humans and AI are joining forces." *Harvard Business Review* 96, no. 4 (2018): 114-123.

[23] Kaiser, Oliver S., and Norbert Malanowski. *Smart Data und Künstliche Intelligenz: Technologie, Arbeit, Akzeptanz*. No. 136. Working Paper Forschungsförderung, 2019.

[24] Schaal, Stefan. "The new robotics—towards human-centered machines." *HFSP journal* 1, no. 2 (2007): 115-126.

Ensuring Web Accessibility of UI Components by following Web Content Accessibility Guidelines

Nikola Mandić[1], Branko Markoski[1], *Velibor Premcevski* [2]

[1] *University of Novi Sad, Faculty of Technical Sciences, Novi Sad, Republic of Serbia*
[2] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Republic of Serbia*

e-mail: nikola.mandic@uns.ac.rs; markoni@uns.ac.rs; velci.aspire@gmail.com

Abstract:
As the number of information and users on the web grows, along with the increase in the number of applications and with the goal of making the web accessible to everyone, there is a need to enable the use of web applications by people with certain disabilities in the normal use of applications. This paper covers specific theoretical concepts of website accessibility for a broader range of users. It also provides an overview of semantic elements within the website. Additionally, the paper showcases the implementation of a UI components that meet accessibility criteria. JavaScript, together with the React library, was used for the implementation of the components. The main goal of the paper is to demonstrate and implement a UI component that meets web accessibility criteria, with a focus on theoretical concepts and practices.

Keywords:
Web applications, Accessibility, User Interface

1. Introduction

Increasing number of users on the internet and the rise in the amount of information available online. There is a growing, worldwide recognition that users with disabilities have the same right as others to access information technologies. This recognition is manifested in the enactment of legislation like that in the United States whose aim is to make the Web and other information technologies accessible to users with disabilities. Such legislation has led to the creation of standards, guidelines, and checklists for accessibility [1, 9]. The paper focuses on important concepts for achieving web accessibility, with an emphasis on implementing a UI component that meets accessibility criteria.

The paper is divided into five chapters. The first chapter describes the structure of the paper, along with a brief summary of the content of each chapter. Within Chapter 2, necessary terms and definitions are described and explained to understand the concept of web applications and client-server architecture. In Chapter 3, there is a review of the concept of web accessibility and the necessary terms for its understanding. In Chapter 4, the focus is on the technologies used for implementing solution. The fifth chapter discusses the implemented solution and the approach taken to resolve it. In the sixth chapter, a review of the implemented solution is provided, and concluding considerations are provided.

2. Web applications

Web applications refer to applications accessed via Web browser over a network and developed using browser-supported languages (e.g., HTML, JavaScript). For execution, Web applications depend on Web browsers and include many familiar applications such as online retail sales, online auctions, and webmail [2]. The only requirement is that there is an adequate web browser capable of running the application. Web applications use and serve documents written in one of the standard formats such as HTML and JavaScript. Some advantages of web applications include providing the same version of the application to all users. Historically, web applications have relied on the client-server model. The client-

server model is a type of software architecture model where two parties are involved: service providers, called servers, and service users, called clients. Typically, a server owns some resources that can be retrieved by a client. Web Application is a kind of system that is built using the client-server architecture where all the communication happens through the Internet [3]. Example of client server architecture is shown in Figure 1.

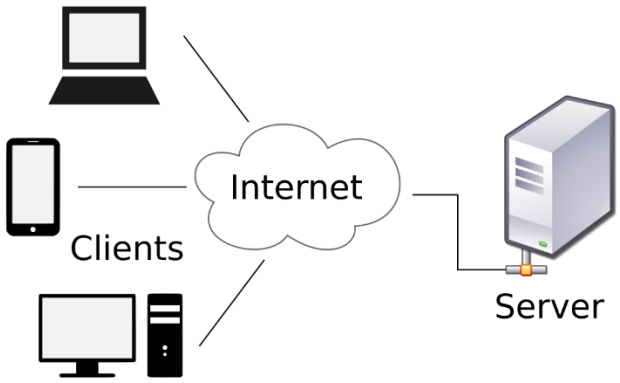


Figure 1: Client server model

3. Accessibility

With the increasing development of web applications in various aspects of human life, software solutions in the form of web applications are becoming increasingly popular. With the increase in the number of users, diverse needs, and user profiles, there arises a need for the client-side of web applications to be implemented in a way that is accessible to both typical users and users with some kind of disabilities. Some of them are:

- Mobility and physical
- Cognitive and neurological
- Visual
- Hearing

There are many ways in which people use web applications, such as keyboards, screen readers and others as shown in Figure 2.

a) Mouth Stick b) Head Wand c) Screen Reader

Figure 2: Examples of ways to use computers

Web accessibility is the practice of making Web sites accessible to all, particularly those with disabilities [4]. More specifically, Web accessibility means that users with specific difficulties and challenges can understand, navigate, and interact with web applications and contribute to the web.

3.1. Accessibility Standards

WCAG (Web Content Accessibility Guidelines) specifies three different conformance levels they are:

- A (lowest)
- AA (middle range)
- AAA (highest)

Level 1 (A) success criteria [5]:

- do not specify how information is presented.
- are reasonably applicable to all Web sites some are machine-testable. Others require human judgment.
- Success criteria that require human testing yield consistent results among multiple testers.

Level 2 (AA) success criteria:

- may require an author to present content in particular ways.
- are reasonably applicable to all Web sites.
- some are machine-testable. Others require human judgment. Success criteria that require human testing yield consistent results among multiple testers.

Level 3 (AAA) success criteria:

- are additional criteria that go beyond Level 1 and 2 that may be applied to make sites accessible to more people with all or particular types of disability.

3.2. Screen Readers

Screen readers convert digital text into synthesized speech. They empower users to hear content and navigate with the keyboard. The technology helps people who are blind or who have low vision to use information technology with the same level of independence and privacy as anyone else. The semantics within an HTML document play a crucial role in web accessibility, especially for users who rely on screen readers. Properly marking elements such as headings, lists, links, and alternative text for images helps screen readers accurately interpret content and deliver it to users with visual impairments or other specific needs. Semantics help ensure that information is conveyed clearly and accurately to users regardless of their abilities. In order for a screen reader to be able to read audio and video content on a website, several steps should be taken to ensure accessibility for users with hearing or vision impairments. It is important to use alternative text for audio and video content, use correct HTML elements such as <audio> and <video>, and provide transcripts. Some of popular screen readers are:

- JAWS
- NVDA
- VoiceOver
- Chrome screen reader

3.3. ARIA

Accessible Rich Internet Applications (ARIA) is a specification being brought forward by the W3C Web Accessibility Initiative's (WAI) Protocols and Formats Working Group. The goal of ARIA is to add additional semantic data into HTML and XHTML to allow assistive technologies to better represent user interface components and dynamic interactions to the user. The specification also addresses providing input focus and full keyboard navigation within the components of an application. The main idea behind ARIA is to add the necessary semantic data into the HTML and XHTML markup. The browser can then interpret this additional semantic data and provide it to the assistive technology via the accessibility API of the platform. Thus, a screen reader can identify a tree control as such. Each tree item is indicated as well as its hierarchy within the tree and its expanded or collapsed state. The ARIA specification defines a standard set of roles and states that can be added into a component [6].

Here are a few key aspects of ARIA:

- Role Attributes: ARIA allows developers to add attributes and role labels to elements on a web page. For example, you can label a specific element as a button or menu. This helps screen readers better understand the function of elements.
- State and Property Attributes: ARIA also enables the addition of attributes that describe the state and properties of elements. For example, you can mark a particular element as hidden or give it a specific value. This helps screen readers inform users about changes in content.
- ARIA allows the addition of attributes that describe actions that can be performed on elements, as well as events associated with those elements. For example, you can label a specific element as an interactive button that can be clicked or as having the functionality to open and close.

4. Technologies

The implementation of the solution in the paper is achieved using JavaScript along with React. These two concepts are described in the following chapters.

4.1. JavaScript and React.js

JavaScript is an interpreted programming language most often used for enhancing webpage interactivity and functionality. It has powerful capabilities to interact with webpage documents and browser windows [7].

Some of the characteristics of JavaScript:

- Client-Side Scripting: JavaScript runs on the client-side (in the user's web browser). This makes it ideal for creating interactive web pages because it can respond to events and actions taken by users on the web page.
- Object-Oriented Language: JavaScript supports object-oriented concepts and allows programmers to organize code in the form of objects, making it easier to manage and maintain code.
- Asynchronous Execution: JavaScript supports asynchronous execution, which means it can perform tasks without blocking the main execution thread. This is especially important for handling long-running operations like server requests without interrupting the interactivity of the web page.

JavaScript is often used with HTML and CSS to create complete web applications. It is also used on the server-side (Node.js) for building server applications.

ReactJS is JavaScript library which is deployed to develop reusable user interface (UI) components. According to React official documentation, following is the definition React is a library for building modular user interfaces. React basically enables development of large and complex web based applications which can change its data without subsequent page refreshes. It is used as the View (V) in the Model-View-Controller(MVC). React abstracts the Document Object Model (DOM), thus offering a simple, performing and robust application development experience. React mostly renders on server side using NodeJS, and support for native mobile apps is offered using React Native. React implements unidirectional data flow thus simplifying the boilerplate and hence proves to be much easier than traditional data binding [8]. React is often used together with other technologies such as Redux for managing application state, React Router for routing, and many other libraries and frameworks in the React ecosystem.

5. Implementation

In this chapter, the implementation and specific parts of the code necessary for the solution will be presented. In Figure 1, a dropdown with keyboard-navigable options following accessibility guidelines is displayed. In the image below, it is possible to see that the implemented dropdown provides the user with the option to navigate through focusable elements using the Tab key. Some of the elements include an "X" button that allows the user to remove an option, and it's also possible to input characters into the

input for easier option filtering. By using arrow keys, users can easily navigate through the options in the popup menu. Additionally, there's functionality implemented to easily remove all selected options. It's important to note that all of these functionalities are designed to be accessible to users who rely solely on keyboard input, ensuring accessibility compliance.

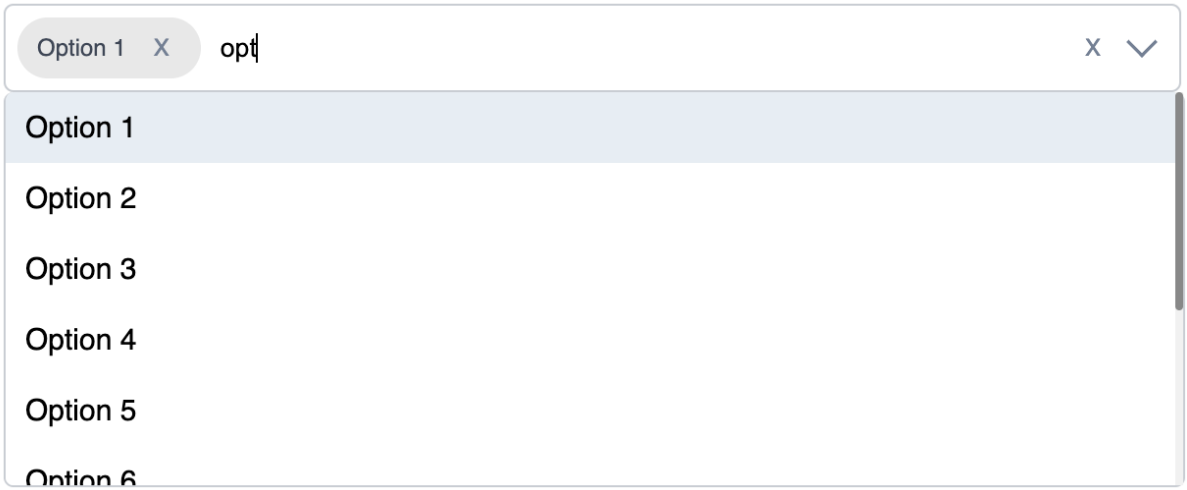


Figure 3: Implemented dropdown

Even though it's not visible in the image, this approach ensures that each element has been given specific semantics using ARIA attributes, allowing screen readers to read and understand their location effectively. More details on implementation of that will be discussed later in the chapter.

```
// Function to navigate to the next or previous item
  const navigateItem = (next: boolean) => {
    if (!liRef.current) {
      liRef.current = firstListItem();
      return liRef.current?.focus();
    }
    liRef.current = next
      ? (liRef.current?.nextElementSibling as T) || liRef.current
      : (liRef.current?.previousElementSibling as T) || liRef.current;
    liRef.current?.focus();
  };

  // Handling keyboard events
  const handleKeyDown = React.useCallback(
    (e: React.KeyboardEvent<T>) => {
      switch (e.key) {
        case KEY_ENTER:
          e.preventDefault();
          handleSelectItem(liRef.current?.innerText);
          liRef.current = null;
          break;
        case KEY_ARROW_DOWN:
          e.preventDefault();
          navigateItem(true);
          break;
        case KEY_ARROW_UP:
          e.preventDefault();
          navigateItem(false);
          break;
        case KEY_TAB:
          onTab?.();
          break;
        default:
          return;
      }
    },
    [handleSelectItem, onTab],
  );
```

Figure 3: Listing 1 – Handlers for keyboard navigation on popup menu

The implementation of keyboard navigation through options within the menu is shown in Figure 3: Listing 1. The navigateItem function is used for navigation between options. It takes an argument next, which determines the direction of navigation within the list, whether to the next element or the previous one. It navigates between sibling options. The handleKeyDown function, shown in the listing, serves the purpose of executing specific logic depending on the key pressed on the keyboard. Within the function, logic is implemented for cases where Enter is pressed (for option selection), arrow keys are pressed (for navigation), and Tab is pressed (for changing focus), among others. Additionally, some built-in functions from the React library were used, such as useCallback for function cache.

On the Figure 4: Listing 2 there is part of the code for the input which is used to search and filter options and buttons responsible for clearing all options and opening the menu with options is shown. Additionally, there are certain ARIA attributes on the input and button, such as:

- aria-autocomplete – which is used to indicate that the autocomplete behavior of the text input is to suggest a list of possible values in a popup and that the suggestions are related to the string that is present in the textbox.
- aria-controls – which is used with provided id of element that serves as a popup.
- aria-expanded – indicator that popup is displayed.
- aria-label – text that is used in screen reader.

These attributes are important for the semantics necessary for a screen reader when navigating through elements.

Figure 5: Listing 3 represents part of the code for the menu options implemented with ARIA attributes to provide necessary semantics for a screen reader. Elements that are not generally intended for use with the Tab key, such as the <li> tag, are enabled using the tabIndex attribute. There are also some additional ARIA attributes to achieve effect of menu dropdown understandable for screen reader with proper semantics. Some of them are:

- aria-multiselectable – indicator that more items can be selected.
- aria-selected – indicator that option is selected for screen reader.
- role – used to identify what can be selected from the user. Combination of listbox with option role.

```
    <input
      ref={inputRef}
      onKeyDown={handleKeyDown}
      className={styles['input-content__input']}
      onChange={(e) => onInputChange(e.target.value)}
      placeholder="Type to search..."
      aria-autocomplete="list"
      aria-controls="options"
      aria-expanded={isOpen ? 'true' : 'false'}
    />
  </div>
  {isLoading && <div className={styles['spinner-icon']} />}
  <button
    aria-label="Clear options"
    onClick={onDeleteAllItems}
    className={styles['x-icon']}
  />
  <button
    aria-label="Open dropdown"
    className={`${styles['arrow-icon']} ${
      isOpen ? styles['arrow-icon--active'] : ''
    }`}
    aria-controls="options"
    aria-expanded={isOpen ? 'true' : 'false'}
  />
</div>
```

Figure 4: Listing 2 – Input for search options with buttons for clear and open menu

```
{isOpen && (
  <ul
    className={styles['dropdown']}
    role="listbox"
    aria-multiselectable={true}
    id="options"
    aria-label="Options to be selected"
  >
    {!isLoading ? (
      dropdownValues.length ? (
        dropdownValues.map((option) => (
          <li
            id="li-elements"
            tabIndex={0}
            onKeyDown={handleKeyDown}
            key={option.id}
            onClick={() => handleSelectItem(option[itemKey])}
            className={`${styles['dropdown__option']} ${
              selectedItems.includes(option[itemKey])
                ? styles['dropdown__option--selected']
                : ''
            }`}
            role="option"
            aria-selected={selectedItems.includes(option[itemKey])}
          >
            {option[itemKey]}
          </li>
        ))
```

Figure 5: Listing 3 – Dropdown menu implementation

Figure 6 presents modal with functionality of trapping focus. It's important to achieve focus while the modal is open for users who rely on keyboards and screen readers.

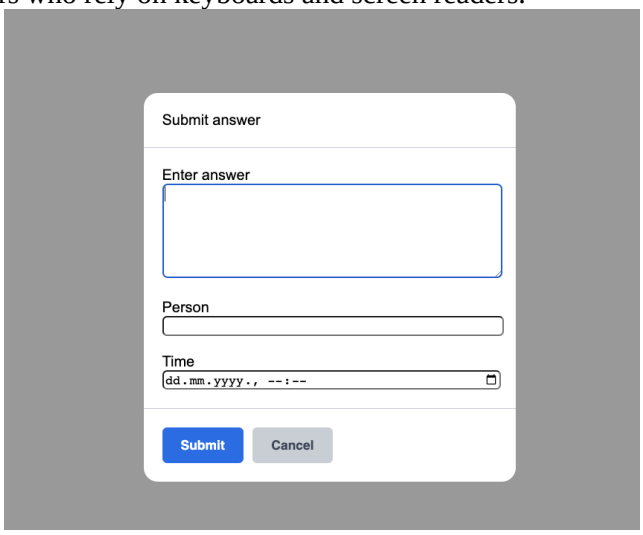


Figure 6: Modal with trapping focus functionality

The importance of such a modal is to navigate using the keyboard only between focusable elements until a certain key is triggered to close the modal or perform a specific action specified by the buttons. Next, the implementation of such a modal with a focus on navigation and handling specific keys will be shown. The modal is flexible and allows dynamic content to be placed inside it, with the responsibility lying on the user to ensure that it meets certain web accessibility standards. In the image above, the content of the modal allows the user to fill in an answer, enter a name, and choose a time interval, all while ensuring web accessibility. Actions for submitting and canceling are enabled, with canceling used to close modal. The user can navigate between each input and button using the tab button.

```tsx
const Modal: React.FC<ModalProps> = ({
  visible,
  destroyOnClose,
  onClose,
  children,
  footer,
  title,
  firstFocusableItemSelector,
}) => {
  /** `indicator` to remove modal */
  const shouldDestroy = !visible && !!destroyOnClose;

  const { contentRef, focusFirstElement } = useModalHandlers({
    visible,
    firstFocusableItemSelector,
    onClose,
  });

  useEffect(() => {
    focusFirstElement();
  }, [focusFirstElement]);

  if (shouldDestroy) return null;

  return (
    <div className={styles['modal-container']}>
      <div className={styles['content']} ref={contentRef}>
        <div className={styles['modal-title']}>{title}</div>
        <div className={styles['main-content']}>{children}</div>
        {footer && <div className={styles['footer']}>{footer}</div>}
      </div>
      <button
        aria-label="Close modal"
        className={`${styles['close-button']} ${styles['x-icon']}`}
        onClick={onClose}
        onBlur={focusFirstElement}
      />
    </div>
  );
};
```

7a)

```tsx
interface UseModalHandlerArgs {
  visible?: boolean;
  onClose: () => void;
  firstFocusableItemSelector?: () => HTMLElement;
}

function useModalHandlers({
  visible,
  onClose,
  firstFocusableItemSelector,
}: UseModalHandlerArgs) {
  // #region used for keyboard events on modal
  useEffect(() => {
    const handleKeyboard = (e: KeyboardEvent) => {
      if (!visible) return;
      if (e.key === KEY_ESC) {
        onClose();
        return;
      }
    };
    document.addEventListener('keyup', handleKeyboard);
    return () => {
      document.removeEventListener('keyup', handleKeyboard);
    };
  }, [visible, onClose]);

  // #endregion

  // #region used for outside click of modal
  const contentRef = React.useRef<HTMLDivElement | null>(null);

  useClickOutside(contentRef, () => onClose());
  // #endregion

  const focusFirstElement = React.useCallback(() => {
    if (firstFocusableItemSelector && visible) {
      const firstElement = firstFocusableItemSelector() as HTMLElement;
      firstElement.focus();
    }
  }, [firstFocusableItemSelector, visible]);

  return {
    focusFirstElement,
    contentRef,
  };
}
```

7b)

Figure 7a): Listing 5 - Implementation of Modal; 7b): Listing 6 - Modal handler implementation

Figure 7a presents implementation of modal and it is using also code from implementation presented in Figure 7b. The useModalHandlers function is used to enable keyboard interaction with the modal. In case the ESC key is pressed, the modal is closed. It uses React's useEffect hook, which is triggered when the modal becomes visible or hidden. Additionally, in the cleanup phase, if the modal is unmounted and no longer in use, it removes the keyboard event listener. Furthermore, useModalHandlers returns the focusFirstElement function as its return value, which is a reference to a function that can focus on the first focusable element within the modal when it is opened. The useClickOutside is used to provide when it's clicked outside of modal content to close a modal. In Figure 7a, the arguments children, footer, and title allow the modal to receive dynamic content and display it using React. This provides flexibility to the modal for use in different scenarios. The firstFocusableItemSelector represents a function for retrieving element that should be the first one to get focus within the modal. The close button of the modal with the ARIA label "Close modal" is used to close the modal. It's important to emphasize that since it's the last focused element within the modal and is on the same level as the content, on blur, it will trigger a re-focus on the first element within the modal unless clicked, in that case modal will be closed.

6. Conclusions

Accessibility is becoming an essential item in the development of web applications due to the increasing number of users with various needs and abilities. As more people use the internet for work, entertainment, and everyday activities, the need for web applications to be accessible to everyone has grown. The goal of accessibility is to enable everyone to access information and use web applications regardless of their physical or cognitive abilities. This paper also demonstrated the implementation of a UI components that meet accessibility criteria using JavaScript and the React library. The implemented solutions allow users to efficiently use the web application via keyboard and screen readers. It described the importance of HTML semantics and the use of ARIA attributes for proper element labeling, enabling screen readers to accurately interpret content. Overall, web accessibility has

become an important aspect of development because all users deserve access to the internet and web content regardless of their individual needs and abilities.

References:

[1] Richards, John T., and Vicki L. Hanson. "Web accessibility: a broader view." Proceedings of the 13th international conference on World Wide Web. 2004.
[2] Al-Fedaghi, Sabah. "Developing web applications." International journal of software engineering and its applications 5.2 (2011): 57-68.
[3] Pavlenko, Andrey, et al. "Micro-frontends: application of microservices to web front-ends." J. Internet Serv. Inf. Secur. 10.2 (2020): 49-66.
[4] Bradbard, David A., and Cara Peters. "Web accessibility theory and practice: An introduction for university faculty." (2010).
[5] Caldwell, Ben, et al. "Web content accessibility guidelines 2.0." W3C Recommendation 11 (2008).
[6] Gibson, Becky. "Enabling an accessible web 2.0." Proceedings of the 2007 international cross-disciplinary conference on Web accessibility (W4A). 2007.
[7] Yue, Chuan, and Haining Wang. "Characterizing insecure JavaScript practices on the web." Proceedings of the 18th international conference on World wide web. 2009.
[8] Aggarwal, Sanchit. "Modern web-development using reactjs." International Journal of Recent Research Aspects 5.1 (2018): 133-137.
[9] Section 508: The Road to Accessibility. https://www.section508.gov/, last accessed on 10.09.2023.

DIRA Learning Platform as a Learning Management System (LMS) for Roma Adults Gaining Knowledge and Skills in E-Services

Nikola Rendevski[1], Dimitar Veljanovski[2], Andrijana Bocevska[3], Monika Markovska[4] Prakash Dhakal[5] and Anca Enache[6]

[1,2,3,4] *University St Kliment Ohridski, 1 Maj nn, 7000 Bitola, R. North Macedonia*
[5,6] *The Deaconess Foundation, Alppikatu 2, FI-00530 Helsinki, Finland*

nikola.rendevski@uklo.edu.mk; veljanovski.dimitar20@uklo.edu.mk; andrijana.bocevska@uklo.edu.mk monika.markovska@uklo.edu.mk, prakash.dhakal@rinnekodit.fi, anca.enache@hdl.fi

Abstract:
Romas' ability to use e-Services is very important. With skills in using digital equipment and e-Services platforms, the Roma are better included in the society. Social inclusion will help Roma adults at risk of social exclusion gain opportunities and resources necessary for full participation in socio-economic and cultural life, and for achieving a decent socially acceptable standard of living and well-being. DIRA project's aim is to achieve inclusive societies in which Roma adults enjoy equal rights and access to services and knowledge. This paper presents the developed DIRA learning platform with an appropriate content for education of adults in usage of the e-Services. Translated in six languages (Finnish, Swedish, Italian, Macedonian, Serbian and English), the outputs will be ready to reach a wider audience impact through the partners' membership in European and international networks. The DIRA learning platform may be used not only by the Roma, but also by other vulnerable groups at an international level (e.g. ethnic or socially excluded minorities and refugees). As the learning platform and materials will be open access and placed in the project's website and EPALE, anyone can freely use the platform in a resource-wise and sustainable manner.

Keywords:
DIRA platform, e-Services, Roma adults

1. Introduction

The Roma are Europe's largest ethnic minority. According to the European Commission out of an estimated 10 to 12 million Roma living in Europe, approximately 6 million are citizens or residents of the EU. Many EU Roma are still victims of prejudice and social exclusion, despite the discrimination ban across EU Member States [1]. The education of the Roma as a marginalized group characterizes with low level of knowledge quality and quantity. The general dimensions of the social exclusion like poverty, unemployment, endangered health, illiteracy etc. which are more present at the Roma population, can be noted the most in their level of education. The educational system is relatively and strictly linear, which means the lack of primary and secondary education among the adult Roma is in correlation with the low level of education in the adult years. This education reality unravels the opportunity to identify possibilities for adult education, which models the project idea to create space for completing Roma adults' education with modern ways of life, especially the digital skills.

The modern information and communication technologies impose the need to modernize the formal and non-formal education of Roma, thus opening up space for implementing information concepts, which enable the completion of their education and gaining skills that make their inclusion in the society more effective. From the aspect of non-formal education, studies show there is no offer for gaining knowledge about the mentioned technologies for the Roma population, opening space for introducing

non-formal educational programs based on the acquisition of digital skills and competences. This type of education represents an important tool for investment in intellectual capital and inclusion of the Roma in the socio-economic structures of society.

The needs assessment study carried out in 2022, done in the project partner countries by the consortium, showed that Roma adults are hardly using available IT tools for e-Services provided by public and private institutions [2]. These services include checking out medical appointments, filling online forms (e.g., tax forms and social services forms), using electronic banking systems, performing online payments, filling service requests, online shopping etc. The vast of adult population and the socially excluded persons are not acquainted with new technologies that nowadays are used for most of the service delivery. Those who are not in everyday contact with technologies consider the fast-growing development of the IT sector, especially in the usage of applications and online forms, difficult to follow. Thus, every update in e-Services needs new skills creating gaps for their successful usage by socially excluded people. Problems and gaps deepen with time and this situation can be decisive for people's inclusion in the system. Although in the partner countries there is generally a developed e-culture among most of the population, there are groups of socially excluded persons, which are not part of this development. There is difference among the countries and a need for sharing the experience and common approach to the problem. Some of the countries have high e-Service culture among people due to the quality of the e-Services and standards of living. But the countries with lower living standards have lower culture, which affects poor and socially excluded persons the most.

The need for abilities and skills related to e-Services usage is increasing as most of the public services are online, or if they are not, governmental institutions invest to make them online. This negatively affects service users who do not have access to IT tools, internet and mobile phone services. Another group of service users affected are adults and socially excluded groups. Also, the covid-19 pandemic made the usage of e-Services a requirement to prevent long lines in front of public service offices, banks and shops for social distancing. However, due to the lack of IT skills and use of e-Services, in many countries we see long lines of Roma adults waiting to withdraw their salaries and pensions or buy groceries and other necessary products, which increases the possibility of covid-19 infection and makes them more vulnerable than those effectively using e-Services. There is an evident need for bringing e-Services closer to Roma adults and other vulnerable groups.

This paper represents the design/features of the developed DIRA web learning platform and modules from which the platform is built for the purpose of the project. It is intended for learning and training Roma adults to increase their use of e-Services and improve their digital and language skills.

The paper is organized as follows. Next section states the objective, previous work is described in Section 3, whereas Section 4 discusses the technologies used for the development of the DIRA platform. Description as well as screens of the developed web platform are given in Section 5. Last section gives a brief review of the research, providing concluding remarks and directions for further work.

2. Objective

The main objective of the creation of the learning platform is the improvement of the socio-economic state of Roma adults through IT education and development of digital skills and competences to enable the adult population and socially excluded people to start using electronic services available in their surroundings, and to become better included, better informed and empowered in their communities and societies.

3. Previous work

There is an evident need for bringing e-Services closer to Roma adults and other vulnerable groups. This will result with benefit from upskilling their knowledge and skills on using e-Services to being better included in their communities and societies.

The digital educational method, contents and partial results of the ongoing project Head in the Clouds: Digital Learning to Overcome School Failure - an EU-Erasmus+ strategic partnership aims at providing quality educational materials for students from Roma communities in order to help participants develop the digital and transversal skills required to overcome existing boundaries to access

(higher) education, employment as well as economic stability [3]. The findings presented in the article by the authors of [4] illustrate that the use of ICTs can contribute to empowering Roma/Gitano adolescents to improve the position they occupy as a group in the social structure. The paper [5] discusses how potential long-term impacts of ICT enabled self-organized learning environments, creating more inclusive educational programs and societies, and the potential contribution of inclusive and IT supported learning environments towards the achievement of the United Nations Sustainable Development Goals (SDGs). Based on the outcomes of the study that included surveys among Roma assistants and teachers, the article [6] analyses the technical conditions available to Roma pupils for remote learning during lockdown and provides a comparative perspective of the position of Roma and other pupils and draws attention to the possible long-term consequences of remote learning for Roma pupils. The authors of [7] introduce the CloudLearning project that represents an alternative and innovative educational method: the way of the SOLE method implemented in their education. The results of the general survey presented in [2] and the conclusions of the national reports show that the usage of digital devices and the access to digital opportunities are severely limited for Roma people regardless of the country - of residence and of origin. This result reinforces the approach of the DIRA project to devise common actions and initiatives to introduce and foster access to digitalization among Roma. Exploring barriers, for example languages and cultural sensitivity, influencing the effectiveness of the learning platform, a strategic model for the effective implementation of the platform was proposed via trainer's training. The trainings provided educators and participants with reinforced digital skills needed at different levels in the community. Paper [8] explores how digital financial services have the potential to support the economic inclusion of poor Roma families if such services are implemented in ways that comprehensively take all five dimensions of access into account. The focus of [9] is the issue of discrimination against Roma communities in the use of educational online platforms on account of failure to provide digital consent. It aimed to foster the discussion of discrimination against marginalized Roma communities in accessing education through online platforms and to point out the inadequacy of national legislation on the provision of digital consent.

4. Used technologies

The developed DIRA web platform aims to help Roma adults to learn smoothly how to use the IT tools. The platform is designed in a simple way, adapted to their low level of knowledge of IT technologies, so that they can remember and practice using e-Services from different institutions and in different conditions and situations. The development and design of such a learning system for Roma adults will contribute to the development of digital skills and competencies, which is in direct correlation with the overall goal of the project.

Towards fulfilling this goal, several different technologies were used for the development of the platform. The backend was build using Flask and a MySQL Server as a relational database management system (RDBMS). Flask is a web application framework written in Python, which was chosen due to its independence upon external libraries, its flexibility and simplicity. Being free and open source while offering data security, high efficiency and scalability on demand, MySQL is one of the most popular and most used RDBMSs worldwide. Phusion Passenger was used as an application server, whereas nginx was used as a web server. Phusion Passenger is an open source and extremely memory efficient web application server which is also designed to integrate into the nginx web server. The nginx web server offers various advantages such as scalability, handling of concurrent requests, load balancing, performance speed up and an overall good browsing experience for the users.

The frontend of the platform was developed using React.js. There are many compelling benefits of using React.js, such as component-based architecture, stable code structure, simplified scripting, as well as easy to learn and use.

5. Description of the developed DIRA platform

The DIRA Project Learning Platform is an innovative online learning platform that aims to provide a comprehensive and customizable solution for creating and delivering courses. It offers a range of features that enable administrators to create courses with individual lectures, each containing diverse content types such as text, video, images, and documents. With its user-friendly interface and customizable language options, the platform aims to make the learning experience accessible and tailored to the needs of users worldwide in a sustainable manner. Consequently, identifying the knowledge gaps, the platform offers additional insights to learning barriers such as languages and cultural sensitivity to improve performance of the platform at an international level.

Administrator users can only be registered by other administrators. They have access to the following features:

1. User Management module - create, edit and delete users of all types,
2. Language Management module - create, edit and archive languages,
3. Category Management module - create and edit category list.

Users Management module, Figure 1, has a search bar (search for a user…) to search for a specific user by their full name. Right under that, there is a green plus button next to which there is the text "CREATE NEW USER" for redirecting to a form for creating new users. Also, there is a table with all the users in the system.

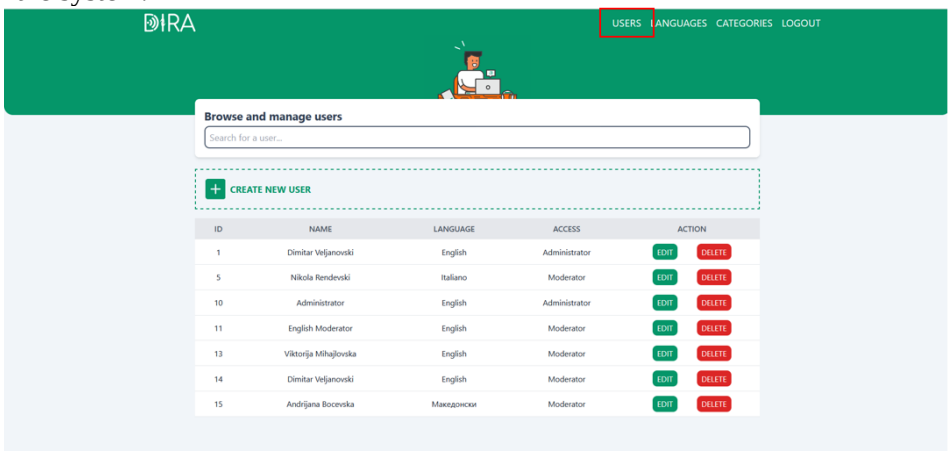


Figure 1: User Management - Navigate to screen

The Language Management module, as illustrated in Figure 2, provides the ability to create, edit and archive different languages. The main difference between Archiving and Deleting is that if you archive a language, you can retrieve it. This is done with the goal not to lose any data for that language.

Similar to the previous two modules, the Category Management module, Figure 3, provides capabilities of creating, browsing and updating categories in the system. By clicking on the name of one of the categories, you will be redirected to a page where you can edit the name and description of that specific category.

Moderator users can only be registered by administrators. They have access to the following features:

1. Course Management module - create, edit courses,
2. Translation Management module,
3. Profile page - Edit user profile.

Courses Management page, Figure 4, has a search bar to search for a specific course by its name. Also there is a green plus button next to which there is the text "CREATE COURSE" for creating new courses. At the very bottom of the screen each created course can be found. By clicking on the name of one of the courses, users will be redirected to a page for editing the details and lectures of the specific course, Figure 5.

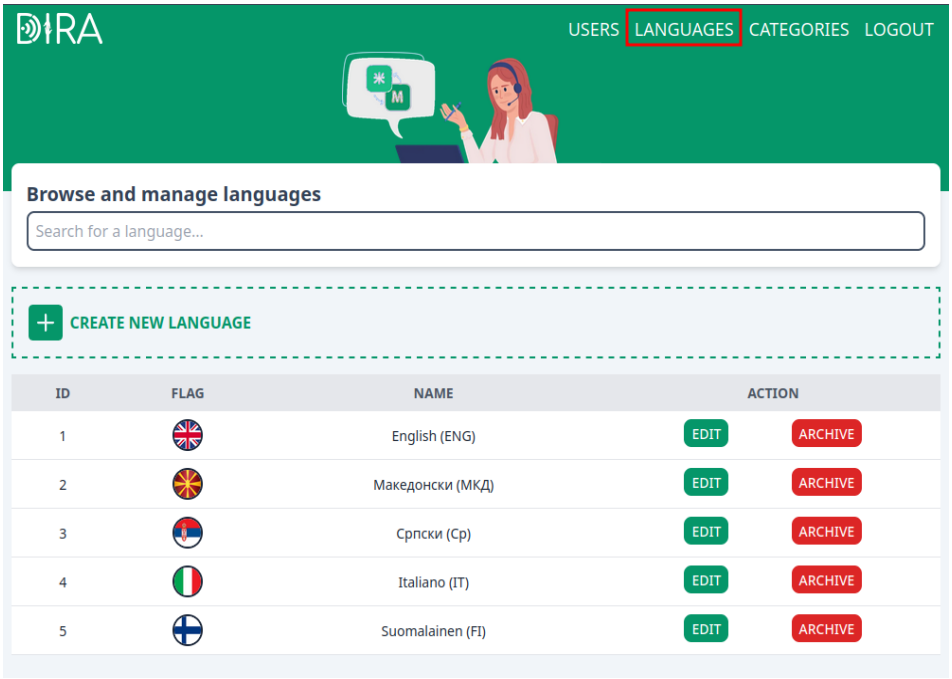


Figure 2: Language Management - Navigate to screen

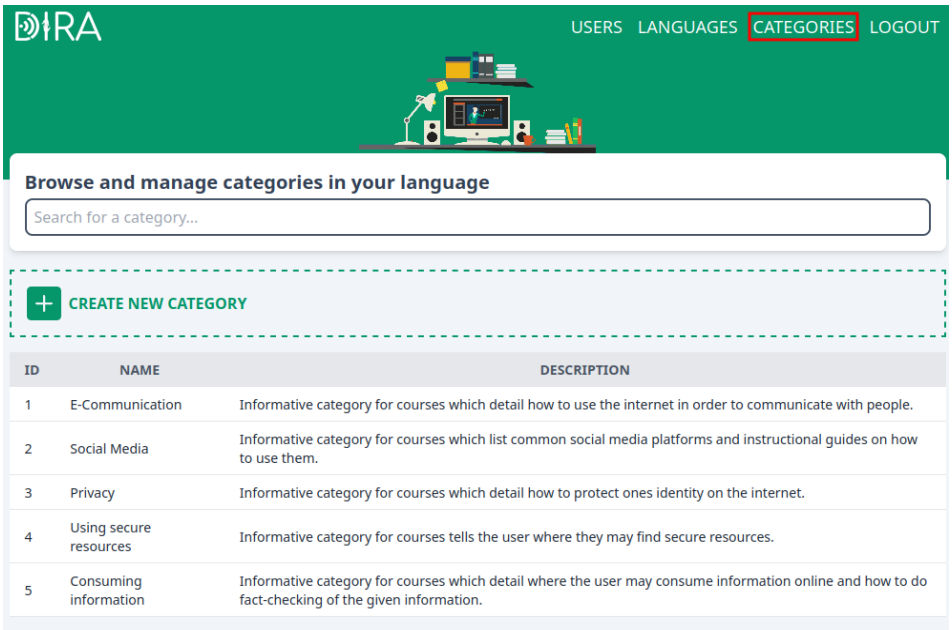


Figure 3: Categories Management - Navigate to screen

The course name, description, categories list and cover photo can be edited. For adding a new lecture, the plus button can be used. Next to each lecture, there are arrows which are used to change the order of the lectures in the course. By clicking on the up-arrow, the lecture will be moved one place up in the list of lectures. Conversely, by clicking on the down arrow, the lecture will be moved one place down in the list of lectures. On the far right, there is an edit button for redirecting to the Edit Lecture page. On that page, the specific lecture itself and its content can be updated. There are two modes of view on this screen. On the top-right of the Browse Lecture Content section, there is a button to toggle between the two modes of view. The view mode is used to see roughly how the content will appear for viewing by the user. The edit mode is used to edit the order and data of existing contents in the lecture or delete the content.

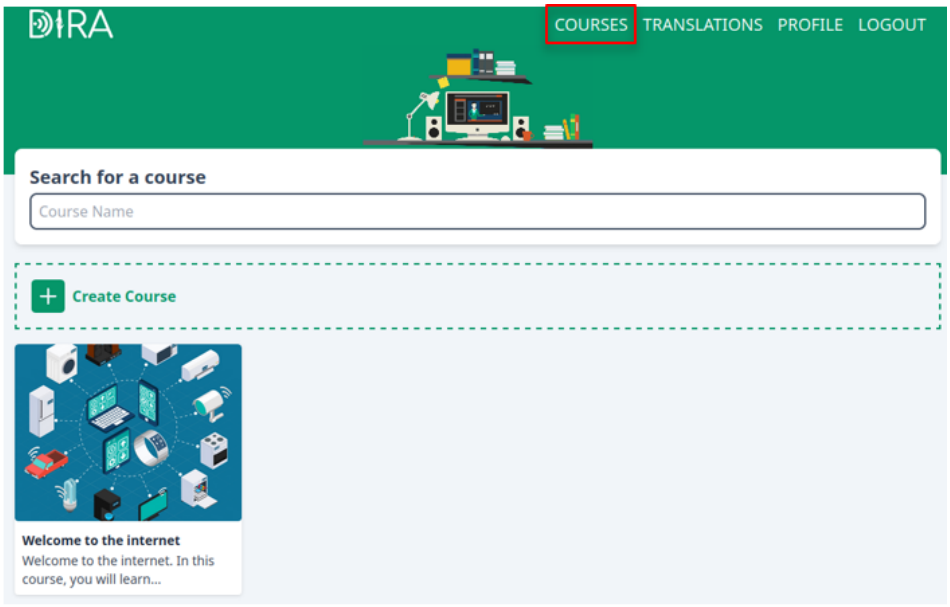


Figure 4: Courses Management - Navigate to screen

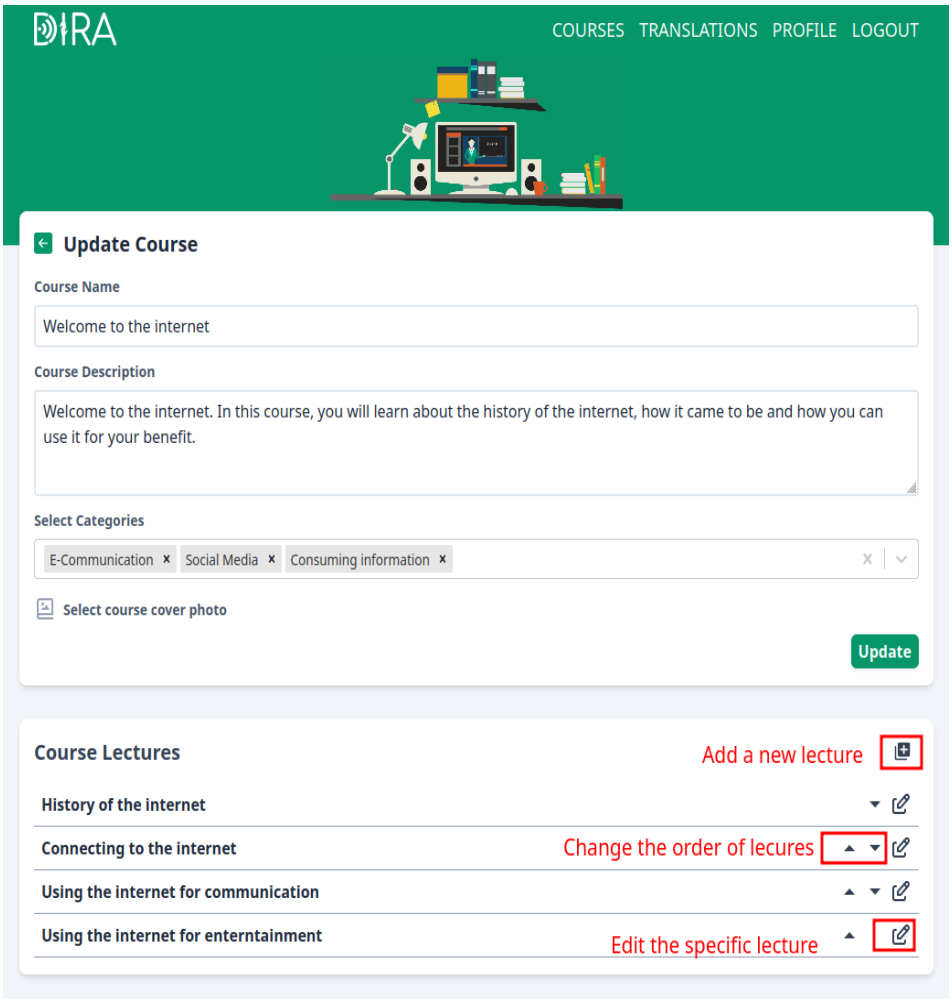


Figure 5: Course Management - Edit course and manage lectures

The Translations Module is used to translate the UI in the language you are assigned to moderate. The translation page consists of a form with multiple inputs, of which each has a label that describes the phrase you need to translate from English to your language.

Profile Page is a screen which enables the user to change their individual profile information - name, surname, email and update their password.

Regular users do not require an account to log-in. When they open the DIRA Learning Platform, they will be greeted with the Courses page. By default, this page displays the courses for the English language. The user can select a language by clicking on the language select button in the navigation bar, as shown in Figure 6.

On the same screen, the user can filter the courses by name or category. By clicking on a course, the user will be redirected to the content page for that specific course, Figure 7. On this page, the user can browse through the different sections in the course and read through the content. By hovering over an image, or pressing on it when using a mobile device, a button will appear. Upon clicking it, the image will expand to fit almost the entire width of the screen.

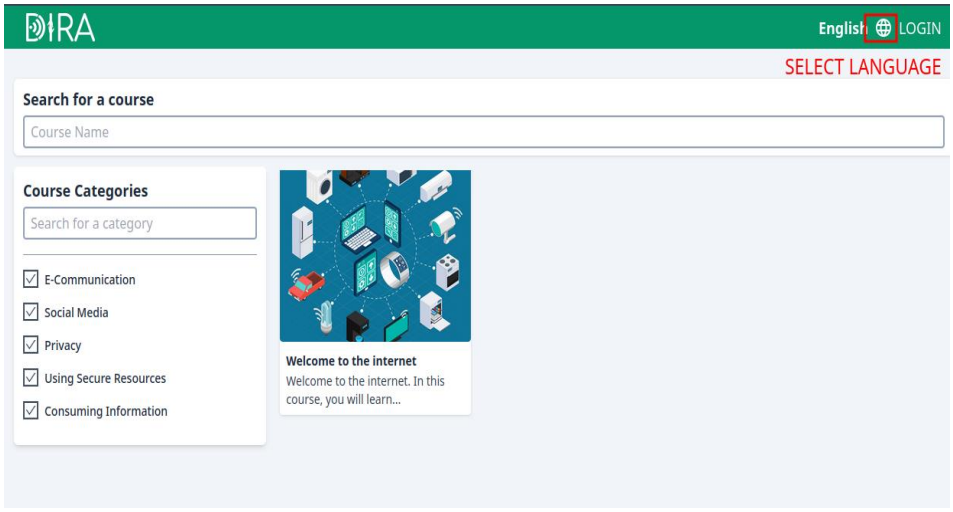


Figure 6: Language button

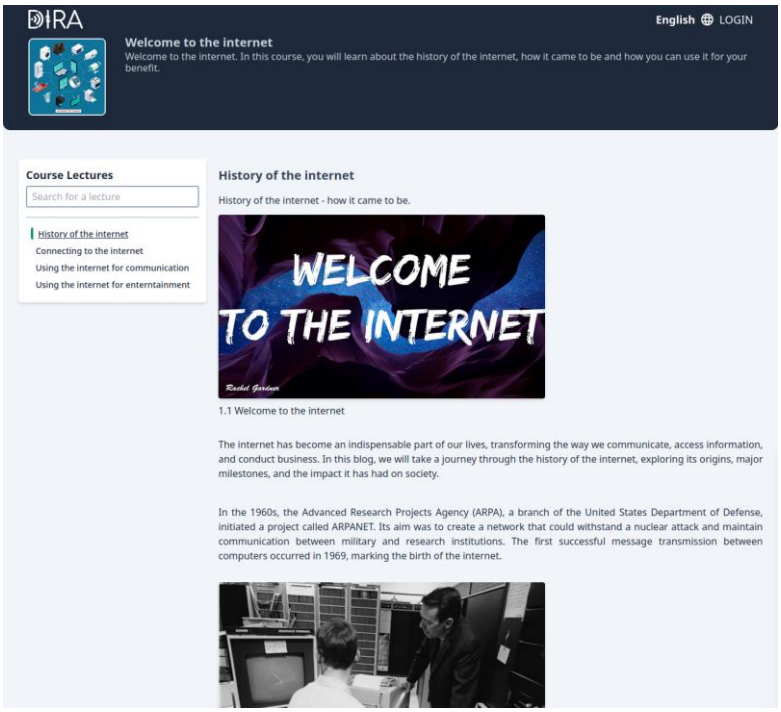


Figure 7: Course Content Page

6. Conclusions

The DIRA project for development of digital skills among the Roma population is an incremental innovation that is computer-generated, based on the principle of learning by doing. The development of digital skills through the creation of IT tools and online platforms is a technological innovation that enables modernization of the knowledge acquisition process. The paper provides basis for improvements to the organizations and institutions where the systematic problem of e-exclusion exists and it will enable them to use the project's outputs and capacities to improve their e-inclusion functioning. The developed DIRA web learning platform for learning and training for the project purpose is intended for upskilling the knowledge and skills of Roma adults on using e-Services in order to become better included in their communities and societies. With its user-friendly interface and customizable language options, the platform aims to make the learning experience accessible and tailored to the needs of users worldwide. The modules and materials from which the web platform is built are tailored through the baseline study and evaluation survey done in the project partner countries. The platform is translated in six languages (Finnish, Swedish, Italian, Macedonian, Serbian and English) with the intention of "achieving inclusive societies in which Roma adults enjoy equal rights and access to services and knowledge". As the learning platform and materials will be open access and placed in the project's website and EPALE, anyone can use them. As a direction for further work, the developed DIRA platform can be extended with new categories, as well as courses and content for these categories. Some aspects of the platform such as for example efficiency, user experience and impact will be evaluated using suitable methods. Moreover, sustainability measures of the platform will be ensured.

Acknowledgment:

This project received funding from the Erasmus+ KA220-ADU - Cooperation partnerships in adult education. In addition, we want to thank our partners DEACONESS FOUNDATION (Finland), ROMNI -APS (Italy), ASSOCIATION FOR ROMA COMMUNITY DEVELOPMENT SUMNAL (THE REPUBLIC OF NORTH MACEDONIA), UDRUZENJE GRADJANA DJURDJEVDAN (Serbia) for their dedication, efforts and continuous support in the framework of the Digital Inclusion for Roma Adults: Gaining Knowledge and Skills in e-Services for upskilling Roma adults knowledge and skills on using e-Services towards being better included in their communities and societies.

References:

[1] Roma equality, inclusion and participation in the EU, European Commission, 2020, https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combatting-discrimination/roma-eu/roma-equality-inclusion-and-participation-eu_en

[2] DIRA Project Report "Digital skills in the Roma community", Results of the survey implemented by DIRA Project in Finland, Italy, Serbia and North Macedonia, https://evermade-hdl.s3.eu-central-1.amazonaws.com/wp-content/uploads/2023/03/28151019/DIRA_SurveyReport_final.pdf

[3] N. M. Novak, M. Rabiee, A. M. Tjoa, Head in the Clouds: An Initiative for Digital Learning among Roma Communities in Europe, In Proceedings of the 10th International Conference on Computer Supported Education (CSEDU 2018), pages 384-390, ISBN: 978-989-758-291-2, https://www.scitepress.org/papers/2018/68068/pdf/index.html

[4] M. Garmendia, I. Karrera, ICT Use and Digital Inclusion among Roma/Gitano Adolescents, Media and Communication (ISSN: 2183–2439) 2019, Volume 7, Issue 1, Pages 22–31, https://core.ac.uk/download/pdf/229941059.pdf

[5] N. M. Novak, M. Rabiee and A. M. Tjoa, "ICTs for Education: An Inclusive Approach to Addressing Challenges Faced by Roma Communities in Europe," 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2019, pp. 1355-1361, https://ieeexplore.ieee.org/document/8757108

[6] R. Bešter, J. Pirc, Impact of Remote Learning during the Covid-19 Lockdown on Roma Pupils in Slovenia, Treatises And Documents Journal Of Ethnic Studies Razprave In Gradivo Revija Za Narodnostna Vprašanja 85/2020, p.139–164, https://rig-td.si/wp-content/uploads/2020/12/Rig-85-Bester-Pirc.pdf

[7] D. Pal'ová, N. M. Novak and V. Weidinger, "Digital learning as a tool to overcome school failure in minority groups," 2017 40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 2017, pp. 767-772, https://ieeexplore.ieee.org/document/7973525?reload=true

[8] C. I. Ravnbøl, "Accessing Cash(lessness): Cash Dependency, Debt, and DigitalFinance in a Marginalized Roma Neighborhood." Economic Anthropology10(1), 2023, 44–54, https://anthrosource.onlinelibrary.wiley.com/doi/10.1002/sea2.12265

[9] D. Nevická, D., M. Mesarčík, Why are you offline? The issue of digital consent and discrimination of Roma communities during pandemic in Slovakia. International Journal of Discrimination and the Law, 22(2), 2022, 172–191. https://journals.sagepub.com/doi/10.1177/13582291221096615

Selection of Key Functionalities for Website Development with a Real Example

Predrag Novokmet[1], Vesna Makitan [1], Dragana Glušac[1], Eleonora Brtka[1] Mila Kavalić[1], and Siniša Mitić[2]

[1]*University of Novi Sad, Technical Faculty "Mihajlo Pupin" Djure Djakovica bb, Zrenjanin, Republic of Serbia*
[2] *Novi Sad School of Business, Vladimira Perića – Valtera 4, Novi Sad, Republic of Serbia*

pedja.novokmet023@gmail.com; vesna.makitan@tfzr.rs; dragana.glusac@tfzr.rs
eleonora.brtka@tfzr.rs; mila.kavalic@tfzr.rs; smitic@uns.ac.rs

Abstract:
Considering the fact that internet marketing and e-commerce play a key role in the success of organizations' business, websites are essential tools for promoting products and services, as well as for establishing interaction with consumers.
This paper aims to provide a thorough insight into the requirements and standards behind the development of websites in accordance with the latest e-marketing trends. Through the analysis of this website its functionality, design, use of interactive elements, search engine optimization, and other key factors that influence the success of websites in the digital environment will be explored.
This study will provide valuable insight into the complexity and requirements behind the development of modern websites, as well as their connection to the latest e-marketing standards.
Keywords:
Website development, e-marketing, CMS

1. Introduction

In today's digital age, internet marketing and e-commerce play a key role in the success of organizations' business. Websites are becoming essential tools for promoting products and services, as well as for establishing interaction with consumers. [1] Accordingly, the analysis of requirements and selection of key functionalities for the development of modern websites become crucial in the process of planning and implementing effective online presences.

This paper aims to provide a thorough insight into the requirements and standards behind the development of websites in accordance with the latest e-marketing trends. In the following sections, the paper will analyze in detail a website that represents an example of good practice in the field of e-marketing and modern website design. Through the analysis of this website its functionality, design, use of interactive elements, search engine optimization, and other key factors that influence the success of websites in the digital environment will be explored. In addition, marketing strategies, including the use of social media, digital advertising, and conversational marketing will be examined. [2] This study will provide valuable insight into the complexity and requirements behind the development of modern websites, as well as their connection to the latest e-marketing standards.

It is important to note that the selection of key elements for successful websites is primarily the result of the first author's three-year experience in developing various websites. Therefore, the proposal of key functionalities is an empirical result, which as such has been confirmed in practice.

2. Research goal and methodology

The main goal of this research is the analysis of the requirements for the development of modern websites and their application to the selected case study. Through this case study, the specific requirements of particular industry will be explored and implemented through the WordPress platform. The purpose of the research is to provide a deeper insight into the website development

process and to identify the key factors that will enable success in their usage. This study will use a combination of qualitative and quantitative research methods. Qualitative methods will be used to analyze requirements and evaluate user experience, while quantitative methods will be used to evaluate website performance. Case study analysis will be performed through the process of website design, implementation, testing, and optimization.

3. Related work

There is much research addressing the issues related to website elements, Content Management Systems (CMS), and WordPress. The researsers in [3] recognized the necessity of including education in designing e-shops using CMS, such as WordPress in e-business/e-commerce classes. This research was extended in [4] and showed that the introduction of an innovative way of teaching this subject, which involves the design of websites and web applications (based on the CMS platform WordPress) is needed, appreciated by students and the market, and worth continuing in future years. The purpose of research [5] is to theoretically identify the salient elements of e-retailing websites and empirically verify them. The researchers in [6] developed a theory-based model of utilitarian and hedonic website features, customer commitment, trust, and e-loyalty in an online hotel booking context. In [7] it was noted that Content Management Systems (CMS), such as WordPress, are built on top of plugin architectures and that the combined activation of multiple plugins in a CMS website will produce unexpected behavior. The researchers in [7] proposed a technique to detect conflicts in large sets of plugins as those present in plugin market places. The study [8] aimed to describe and analyze the global status and trends in the application of Web Content Management Systems (WCMS) in academic library websites over the last decade. One of the conclusions was that there are a significant geographical differences in the adoption of WCMS in different countries/regions. For example, university libraries in Europe prefer to use open-source WCMS to manage their websites, such as Drupal and WordPress.

4. Setting up the WordPress environment

Within the development of modern websites, WordPress stands out as one of the most popular Content Management System systems [3,8,9]. Content Management System is a software platform that enables the creation, editing, and management of digital content on a website, without the need for programming or advanced technical skills [10]. When creating websites with WordPress, the choice of theme plays a key role in defining the appearance and functionality of the site. However, in this case, templates were not used, but Elementor, which allows building sites "from scratch" [11]. It allows complete freedom in creating a unique design and structure of websites that match the specific requirements and wishes of the clients. The steps in setting up WordPress environment to work with Elementor are:

- Installing WordPress and choosing and installing an Elementor-compatible theme;
- Installation of the Elementor plugin – Elementor is a plugin for WordPress that enables visual editing of websites. It can be downloaded and installed from the WordPress admin panel;
- Adjusting Elementor settings;
- Creating and editing web pages in Elementor.

In addition, the following plugins were used to provide additional features and improvements:

- Rank Math SEO (search engine optimization) – a search engine optimization plugin that provides tools and functionality to improve the SEO performance of websites;
- WordPress File Manager – a plugin that allows you to manage files and folders directly from the WordPress admin panel;
- WPvivid Backup Plugin – a plugin for creating backup copies of the website;
- MouseWheel Smooth Scroll – a plugin that provides smooth scrolling when using the mouse on a website.

While working on the website, the needs and goals of the company were studied, and personal experience, knowledge, and skills were used in order to adapt the website to the visual identity of the

company, its industry, and target group. Although no template themes were used to create the site, the principles of web page design such as visibility, ease of navigation, responsiveness, and fast-loading were respected. The aforementioned made it possible for the website to be unique and adapted to the company's needs, which helps it stand out on the Internet and achieve success in its field.

5. Case study: Creating a WordPress website for a hostel

One of the main goals of the hostel is to attract travelers who are looking for comfortable accommodation and an authentic experience. When analyzing the requirements for the hostel website, the following guidelines were taken into account in relation to each of the elements of the site:

- Website design: The key is to create a visual identity that reflects the spirit of the hostel and attracts travelers. The focus should be on a modern, fresh, and attractive design that emphasizes the comfort and hospitality of the hostel. Also, high-quality photos should be used that show the accommodation units, common areas, and the surroundings of the hostel.
- Site structure: The site should be well-organized and easy to navigate. It is important to have a clearly marked menu that allows quick access to key sections, such as accommodation units, common areas, services, and hostel events. Also, information about the location of the hostel, proximity to attractions, and transportation options should be highlighted.
- Reservation and contact: The site should provide a simple and secure system for booking accommodation. Also, contact information and an inquiry form should be readily available for visitors who wish to ask questions or request additional information.
- Photo gallery: In order to attract the attention of travelers and to convey the atmosphere of the hostel, the site should contain a photo gallery. Photos should show the comfort of the accommodation, modern common areas, social activities, and proximity to attractions. These photos will help travelers gain a better insight into what the hostel offers and awaits them during their stay.
- Social media integration: In order to increase the hostel's visibility and interaction with travelers, it is important to have social media integration. The site should allow visitors to follow the hostel on social networks and share relevant content through their profiles. Also, links to hostel profiles on popular platforms should be available.

This analysis of hostel website requirements serves as the basis for the planning and development of this website.

5.1. Site testing and optimization

After the successful implementation of the functionality on the website of the "Garni Ideal" hostel, the next step was to test and optimize the site to ensure that all functionalities work properly and that visitors have a quality user experience. All interactive elements on the site were tested, including reservation forms, contact form, and photo gallery. The correctness of the display of all data, the success of sending reservations, and messages, as well as the correct display and the possibility of viewing photos were checked. Also, field validation is checked to prevent incorrect data from being entered. After that, the site's responsiveness was tested on different devices and browsers. It is ensured that the site is displayed correctly on desktop computers, tablets, and mobile phones, as well as that all elements and content are clearly visible and adapted to each device. In parallel with the testing, the loading speed of the site was optimized to ensure fast page loading and avoid long waiting times for visitors. This is made possible by optimizing images, minimizing CSS and JavaScript files, as well as using caching and other techniques to speed up loading.

In the site optimization process, Rank Math SEO plugin was used to optimize meta data, page titles, descriptions, and keywords. Also, the structure of the URLs has been checked and ensured that they are friendly to search engines. This improved the visibility of the site on search engines and enabled better ranking in the search results. All changes and optimizations on the website of the "Garni Ideal" hostel were aimed at improving the user experience, increasing visibility on search engines and achieving better site performance.

5.2. Evaluation of results and user experience

After the creation of the "Garni Ideal" hostel website, an evaluation of the results and an analysis of the user experience was carried out in order to assess the success of the project and identify opportunities for further improvement.

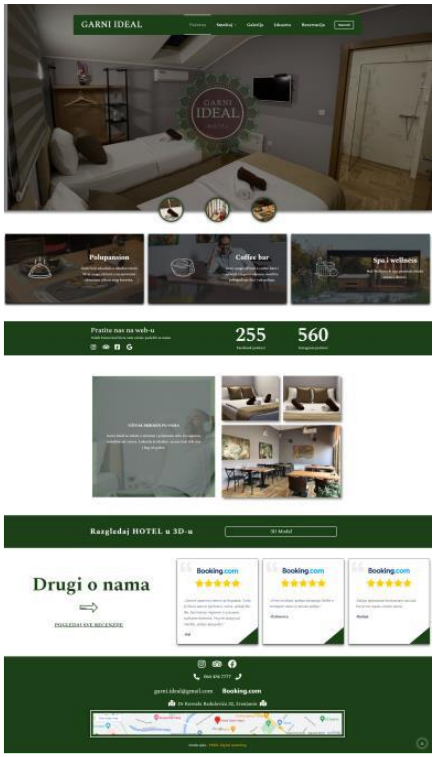


Figure 1: *Index* page of hostel website.

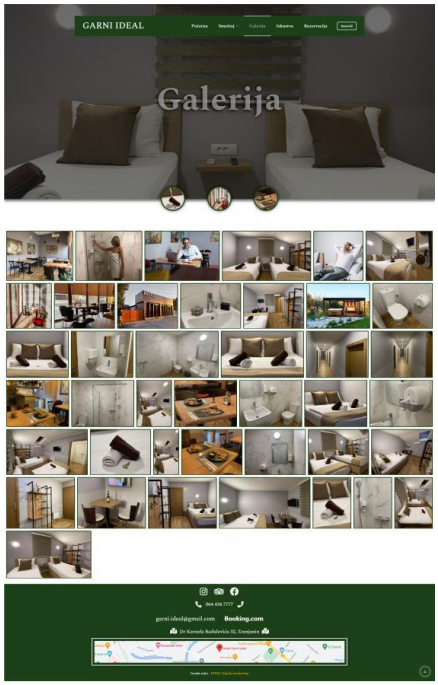


Figure 2: *Gallery* page of hostel website.

Figure 3: *Reservation* page of hostel website.

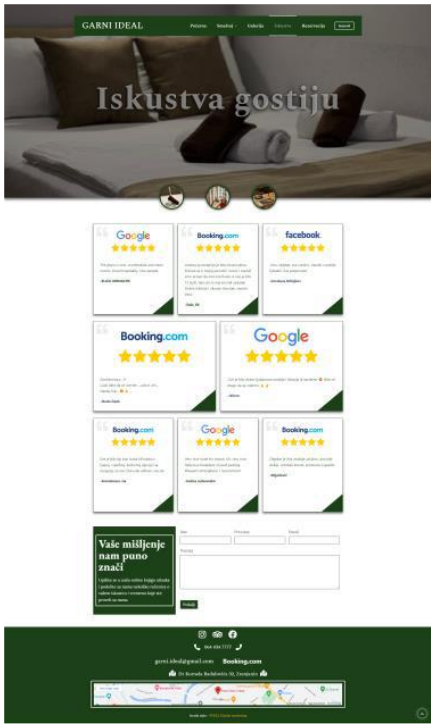


Figure 4: *Recommendations* page of hostel website.

One way to measure the success of the site was to collect feedback from visitors. There is an option for comments and ratings on individual pages so that users can express their impressions and suggestions. Also, analytical tools were used to monitor the number of visits, average time spent on the site, and other relevant metrics.

As part of the evaluation, screenshots of the site pages were used. A few selected screenshots illustrating the design, functionality, and interaction of the hostel website are shown in Figures 1-4.

The charts in figures 5-8 represent user experience about content overview, ease of navigation, intuitive design, and usefulness of booking forms and photo galleries.

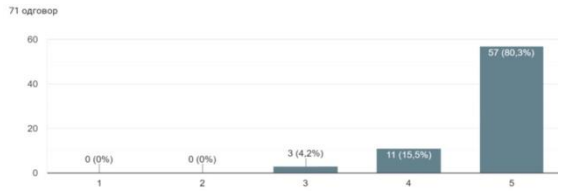


Figure 5: User experience about content overview.

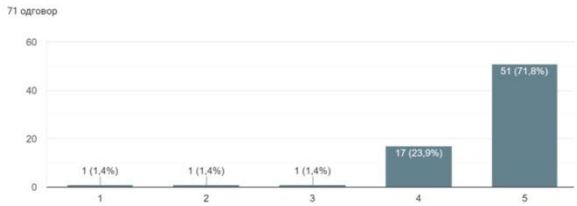


Figure 6: User experience about ease of navigation.

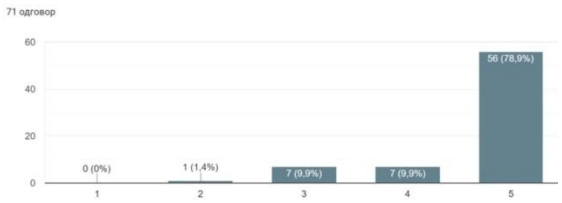


Figure 7: User experience about intuitive design.

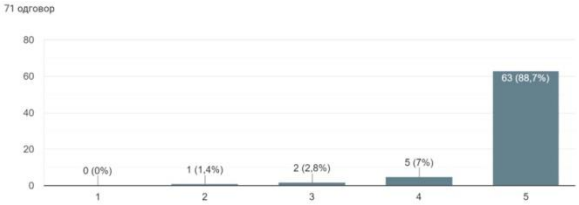


Figure 8: User experience about interaction.

Taking into account the feedback and analytical data, it was determined that the website of the "Garni Ideal" hostel has a positive response from visitors. Users expressed satisfaction with the content overview, ease of navigation, and intuitive design. They also highlighted the usefulness of booking forms and photo galleries for investigating the hostel's offer.

6. Analysis of requirements for modern websites

Regardless of the type of business, websites have become an essential tool for success in today's digital environment and play a key role in business by enabling companies to: have an online presence, communicate with potential customers, build a brand, promote products and services, generate sales, and track performance.

Based on the experience in creating websites, it can be concluded that there are several key factors for the development of successful sites that meet modern e-marketing standards:

1. Attractive design and well-organized site structure – attractive design with modern visual elements and a logically organized site structure improves user experience and attract visitors.
2. Site loading speed – it is a fact that users have less and less patience for slow-loading web pages. Therefore, site performance optimization, such as image optimization, caching, and server optimization, is necessary to ensure fast and efficient content delivery.
3. Responsive Design – considering the increasing use of mobile devices, responsive design is vital. Adapting a website to different screens and devices ensures an optimal user experience [1].
4. Relevant and high-quality content – users appreciate high-quality content that is original, relevant and provides value. Integrating different types of content, such as text articles, blogs, images, and videos, contributes to user engagement and improves SEO results.
5. Site security – it is important to ensure data security and protect user information. The implementation of security measures such as SSL certificates, data encryption, and regular updates of the platform and plugins ensure that the site is safe for users.

Therefore, a combination of the above key factors ensures the success of a website. The integration of attractive design, loading speed, responsive design, quality content, and site security enables users to have an outstanding user experience and achieve the business goals of the enterprise.

7. Conclusions

This paper provided a thorough insight into the requirements and standards behind the development of websites in accordance with the latest e-marketing trends. Also, it provided valuable insight into the complexity and requirements behind the development of modern websites, as well as their connection to the latest e-marketing standards.

The website of the "Garni Ideal" hostel was used as a case study. It was established that the key elements for successful websites are: attractive design and well-organized site structure, site loading speed, responsive design, relevant and high-quality content, and site security.

Through an attractive design, relevant information, and functionality, this kind of site is a powerful tool for attracting visitors, providing information, and achieving the goals of business entities.

Acknowledgment:

This research is financially supported by Ministry of Education and Science of the Republic of Serbia under the project number TR32044 “The development of software tools for business process analysis and improvement”, 2011-2023.

References:

[1] M. Vuković, M. Đorđević, Digitalni marketing: Strategije i taktike, Data Status, 2020.
[2] A. Macarthy, 500 Social Media Marketing Tips, Independently published, 2018.
[3] K. Krupcała, A. Januszewski, Website and e-shop Development as an e business Teaching Programme Innovation in Management Education. 24th International Conference on Knowledge-Based and Intelligent Information & Engineering, Systems Science Direct, 2020.
[4] K. Krupcała, Developing e-business competencies among business students – research and opportunities, Procedia Computer Science 192, pp. 4711–4720, 2021.
[5] I. Alnawas, A. A. Khateeb, Developing and validating a multidisciplinary scale of E-retailing website elements, Journal of Retailing and Consumer Services 66, 102905, 2022.
[6] A. Bilgihan, M. Bujisic, The effect of website features in online relationship marketing: A case of online hotel booking, Electronic Commerce Research and Applications 14, pp. 222–232, 2015.
[7] I. Lima, J. Cândido, M. d’Amorim, Practical detection of CMS plugin conflicts in large plugin sets, Information and Software Technology, 118, 106212, 2020.
[8] X. He, Y. Huang, Web Content Management Systems as a Support Service in Academic Library Websites: An Investigation of the World-class Universities in 2012–2022, Elsevier: The Journal of Academic Librarianship 49(3), 102716, 2023.
[9] W3Techs, Usage statistics of content management systems, 2023. URL: https://w3techs.com/technologies/overview/content_management
[10] Kinsta, What Is a Content Management System (CMS), 2019. URL: https://kinsta.com/knowledgebase/content-management-system/
[11] Elementor, Create Websites with Elementor's WordPress Customizer, 2023. URL: https://elementor.com/features/wordpress/

Use of Python and OpenCV in Thermal Image Processing

Sinisa Mihajlović[1], Dragan Ivetić[2], Ivana Berković[1], Dalibor Dobrilović[1]

[1] *University of Novi Sad / Technical Faculty "Mihajlo Pupin", Đure Đakovića St., 23000 Zrenjanin, Serbia*
[2] *Faculty of Technical Sciences, Novi Sad, Republic of Serbia*

mihajlovicsinisa25@gmail.com; ivetic@uns.ac.rs; ivana.berkovic62@gmail.com; dalibor.dobrilovic@tfzr.rs

Abstract:
This paper presents the processing of thermal images using Python programming language and OpenCV library. Computer vision deals with image processing such as filtering and edge detection, color detection, etc. The process of identifying colors is called color detection. Red, blue, and green are the basic colors of computer vision. For this research, images of the platform based on Arduino UNO microcontroller board created with a thermal camera Testo 882 were used. This platform is designed to collect solar radiation data and therefore planned to be exposed to direct sunlight. The thermal images with color detection can be used to identify and monitor the heating of the platform components, as well as to analyze the impact of the temperature of the solar panel on its efficiency. In this paper, original thermal images of the platform are shown, as well as filtered images that show areas with defined temperatures.

Keywords:
thermal images, image processing, Python, OpenCV, Arduino

1. Introduction

OpenCV is Open Computer Vision Library launched by Intel 1999 The library is written in programming languages such as C and C+ and can run on Windows and Linux. This library can be used with other programming languages like Python, MATLAB, Ruby etc. NumPy and Python are powerful tools for image processing. The field of computer vision has image processing abilities that include filtering, edge detection, corner detection, sampling and exclamation, color conversion, morphological operations, histograms, and many more [4].

Color detection is the process of identifying the color of targeted pixels in images. Humans perform this action naturally, while it is not the case for computers. Human eyes and brain work in coordination to translate light into color. The signal to the brain that recognizes color is transmitted by light receptors present in the eyes. The basic colors of computer vision are red, green, and blue. This helps in recognizing colors and in robotics. Color detection has its applications in driverless cars. This system is useful in detecting traffic and vehicle backlights and in making decisions to stop, start, and continue driving. It can be used in industry to pick and place different colored objects with the robotic arm. Color detection is also used as a tool in various image editing and drawing apps [5].

The statistical machine learning libraries used by OpenCV are Decision tree learning, Gradient boosting trees, Boosting, Deep neural networks, Convolutional neural networks, Support vector machine, Random Forest, Artificial neural networks, Naive Bayes classifier, K-nearest neighbor algorithm, and Expectation-maximization algorithm [6].

Python is a powerful, procedural, object-oriented, and functional programming language, created by Guido Van Rossum in the late 1980s. Python is used for computer GUI development, software development, web development, education, and scientific applications. Removing the brackets makes the code shorter. Learning advanced functions is a bit more complex, while some tasks are quite simple.

Python project is easy to understand, and Python code is concise, efficient, understandable, and manageable. Python's standard library is huge and there are all the functions to solve various tasks. The Python package manager (pip) makes it easy to import other packages from the Python package index (PyPi). A Python application can be written on one platform and used on other platforms [12].

This paper presents an approach in using Python and OpenCV library for thermal image processing. The thermal images of the platform based on an Arduino UNO microcontroller board created with a thermal camera Testo 882 were used. This platform is designed to collect solar radiation data and therefore planned to be exposed to direct sunlight [1, 2, 3]. The thermal images with color detection can be used to identify and monitor the heating of the platform components over the time of its utilization. Also, to to analyze the impact of the temperature of the solar panel on its efficiency by comparing the change of solar panel temperature and solar panel output voltage in combination with other parameters collected by the sensors.

2. Related work

Computer vision can recognize the image characteristics and deals with image processing, such as color detection, edge detection, shape detection, etc. Using OpenCV can improve image recognition, and together with Python it can be used for real-time video recognition. [6].

Other usage of image recognition techniques can be various. Documents and certificates are signed daily, and those documents can be scanned. One example is signing a contract for a mobile SIM card. Fraudsters often take original documents and scan them. Scanned documents can be easily modified through various digital tools. One of the most used tools is Adobe Photoshop. Fake documents, i.e., counterfeits can be used for various scams. Code written in Python and OpenCV based on FLANN (Fast Library for Approximate Nearest Neighbors) evaluates whether the image is original or fake. OpenCV has huge algorithms support for extracting features in images and videos [6].

With the help of OpenCV it is possible to extract colors from images using the KMeans algorithm and filter images from image collections based on RGB color values. Matplotlib.pyplot and numpy libraries should be also installed as well as the KMeans algorithm which is part of the sklearn cluster subpackage. To compare the colors, they must be converted using rgb2lab, and then the similarity is calculated [7]. In the paper [8], an application for tracking and detecting faces in videos and cameras was developed, which can have multiple purposes. Face detection is analyzed using OpenCV.

Artificial intelligence and machine learning algorithms help in object detection. OpenCV can be used to identify and track objects in real time. Image identification uses object detection, recognition, and segmentation techniques. The use of artificial intelligence and machine learning improves the speed of data processing and the maintenance of results [6]. To ensure that the wire meets the production requirements, the exact number of copper cores in the wire must be known. An edge detection method based on OpenCV with computer vision and image processing algorithms and functions can help for this purpose. High-resolution cameras are used to record the interior of the wire structure [13].

In urban areas, it is necessary to optimize the efficiency of traffic flow. One of the reasons for traffic congestion is due to red light delays. In order to avoid the problem of traffic control, an Adaptive Intelligent Traffic Light control system (AITLCS) based on OpenCV, and image processing technique is proposed. The proposed system is designed to provide efficient traffic flow for everyday life. OpenCV handles signals. The system consists of a lane-facing camera that takes a picture of the route being traveled, takes the density of the sand and the vehicle, and compares it to each image using image processing. The images are effectively processed to know the traffic density [14].

The smartphone industry is growing rapidly, so smartphone apps need to use less power. In [15] two Android applications based on video processing methods are presented, one uses the OpenCV library and the other uses the Android library with the CamTest algorithm. Eight methods are applied to each frame of recorded video. Efficiency and energy consumption, as well as processing speed, are compared. Out of eight image processing methods, six methods using the OpenCV library are faster than CamTest.

3. Processing of thermal images

OpenCV is a Python library and is used to solve computer vision problems, i.e., analyzing and manipulating digital images, as well as image processing. OpenCV is often used to recognize human faces, objects, handwriting, and more. OpenCV must be installed before using it.

It is installed through the command prompt, with the following command:

```
pip install opencv-python
```

After entering the command, the OpenCV library package will be downloaded and installed [10]. In the previous research [11], a thermal camera Testo 882 was used to create thermal images. The thermal camera Testo 882 was fixed on a tripod, and a timer was set to take images every 30 minutes. Thermal images have a resolution of 640x480 pixels. Images are stored locally on the camera's memory card, and after a trial period, the images are transferred to a computer. After that, thermal images are analyzed and processed.

Easy processing and precise analysis of infrared images on a computer is made possible by IRSoft thermography software. IRSoft is used for analysis, processing, and reporting for clear presentation of data. The display of critical temperature points on the image can be highlighted by software. Hot/cold spots can be determined through thermographic application. Processing of thermal images is used to determine the temperature exposure of the platform, which is based on the Arduino UNO microcontroller and open-source platform. The obtained images can be used to identify and monitor the heating of the platform components, as well as to collect (acquisition) data on the influence of the temperature of the solar panel on its efficiency.

Processing images with the IRSoft is sufficient in the case when we have several images to process. In the case when the batch processing of multiple images is needed it is much better to develop our own tool for image processing. This can be achieved with Python and OpenCV library.

4. Results

A prototype application was created using the Python programming language and OpenCV and the working environment Spyder. Spyder is a free, open-source environment written in Python for Python. The created application enables image processing. For further testing of the application, thermal images created by the Testo 882 thermal camera were used. Testo thermal camera stores thermal images in .BMT format which can be opened in IRSoft. The preprocessing of the images in IRSoft includes changing of palette to Hot/Cold, and exporting the image in .PNG format. The size of exported images is 1000x1633.

In the IRSoft program, a thermal image obtained by a thermal camera is opened. Thermal images can be exported in different image formats. For further testing of application, a thermal image in .PNG format is used.

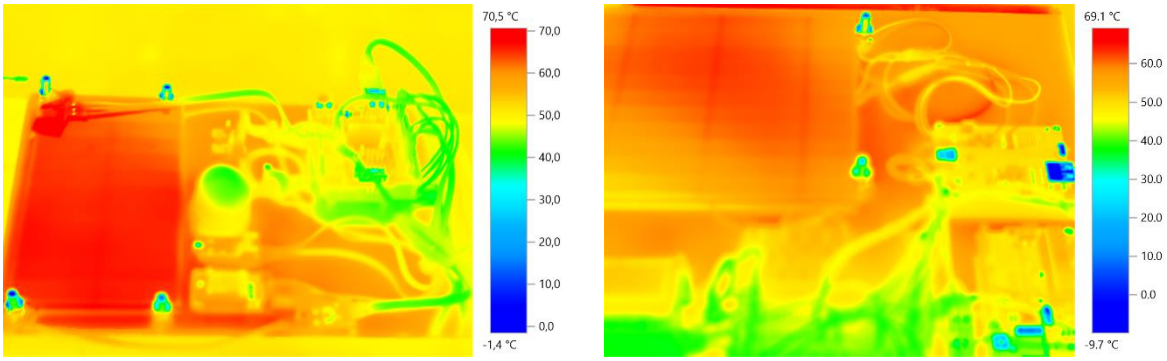


Figure 1. Two thermal images of the platform based on the Arduino UNO microcontroller used for processing.

Two images shown in Figure 1 are used for testing the application. Both images show a platform based on the Arduino UNO microcontroller and an open-source platform. In Figure 1, in addition to the display of the platform, the temperature scale is also shown. The temperature scale ranges from -1.4°C to 70.5°C on the left image, and from -9.7°C to 69.1°C on the right image. The temperature scale contains colors that show the temperature from the coldest to the hottest.

In the first phase, we wanted to separate parts from the thermal image, with a temperature range from 60°C to 70.5°C. Colors are displayed in HSV format. To detect the color from the temperature scale, the GIMP software and color picker tool are used, by manually clicking to the targeted pixel in a color scale. This can be automated in the Python script, and it will be included in the next phase of developing software. With the color picker tool, the first color on the scale at 60°C is selected, and in

HSV code this color is 34, 100, 100, and the second color for 70.5°C is 0, 100, 100. Those values are used for both images.

For now, the application can only load one image at a time. Before the start of processing the temperature is shown in specific colors, red for hot and blue for cold, as can be seen in Fig. 1 in color scales. After starting the application, the image is filtered to show only part of the images with color indicating temperature in the range from 70.5°C to 60°C. the original image is displayed, then a black and white mask, where the white color represents the result. The result of the range between 60°C and 70.5°C is shown in Figure 2 in the form of a colored mask.

Figure 2. The result of the range between 60°C and 70.5°C shown in the form of a colored mask.

The same filter can be shown in grayscale masked from as it is shown in Fig. 3.

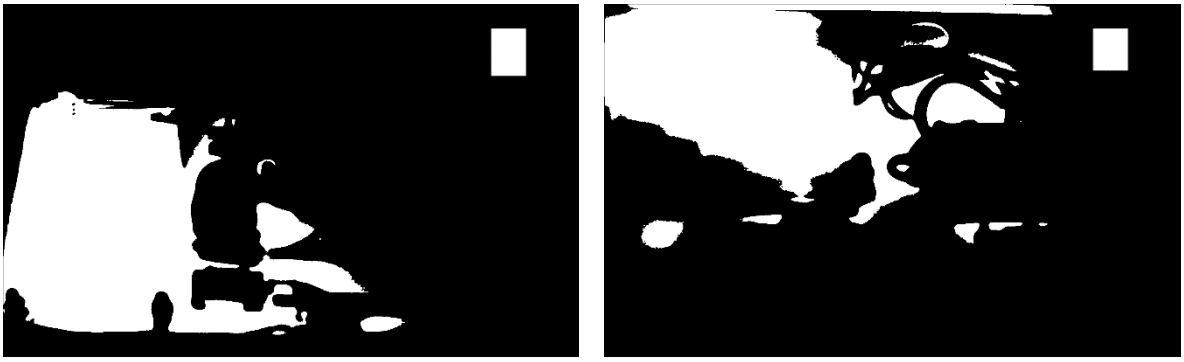


Figure 3. The result of the range between 60°C and 70.5°C shown in the form of a greyscale mask.

Another example is to extract colors from a thermal image that are in the range of 30°C to 40°C. Color codes in HSV format for 30°C are 184, 100, 100, while for 40°C it is 125, 100, 100. Figure 4 shows the result of the range between 30°C and 40°C, in the form of a colored mask.

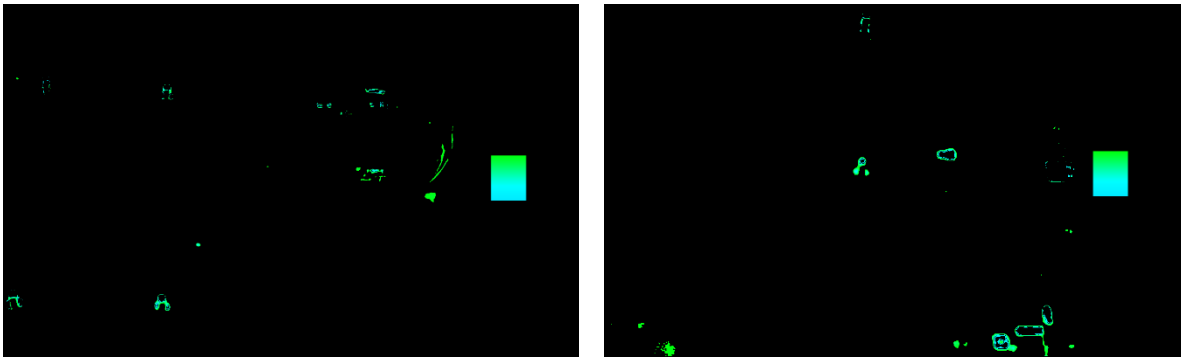


Figure 4. The result of the range between 30°C and 40°C shown in the form of colored mask.

5. Conclusion and further work

This paper presents the processing of thermal images using Python and OpenCV. Thermal images were obtained with a thermal camera Testo 882. The thermal camera was fixed, and every 30 minutes it created one photo of the platform based on Arduino UNO microcontrollers and open-source platform. Thermal images of the platform are used for further research and analysis, where it is important to monitor the heating of the platform components, as well as to collect data on the impact of the temperature of the solar panel on its efficiency.

An application prototype was created in the Spyder environment, using Python and OpenCV. The application currently has the ability to open and process one image at a time. After that, it is necessary to determine the range of colors that are required in the image. In the thermal image, shades of color are shown, and they represent temperature. Two examples are made in the paper with two images. One example is finding a temperature in the range of 60°C to 70.5°C, and another example in the range of 30°C to 40°C.

Further development will include the expansion of applications for multiple (batch) processing and the use of convolutional neural networks for detailed analysis of thermal images in combination with other data collected with the sensors. Thus, recognizing the critical parts of the platform (i.e., high-temperature points and component overheating) and its behavior during exposure to direct solar radiation, will be enhanced with convolutional neural networks.

Acknowledgment:

This research was funded through the project "Creating laboratory conditions for research, development, and education in the field of the use of solar resources in the Internet of Things", at the Technical Faculty “Mihajlo Pupin” Zrenjanin, financed by the Provincial Secretariat for Higher Education and Scientific Research, Republic of Serbia, Autonomous Province of Vojvodina, Project number 142-451-2684/2021-01/02.

References:

1. D. Dobrilovic, J. Pekez, E. Desnica, Lj. Radovanovic, I. Palinkas, M. Mazalica, L. Djordjević, S. Mihajlovic, Data Acquisition for Estimating Energy-Efficient Solar-Powered Sensor Node Performance for Usage in Industrial IoT, Sustainability 2023, 15, 7440. https://doi.org/10.3390/su15097440
2. D. Dobrilovic, Approach in Using Open-Source Hardware in Teaching Solar-Powered IoT Node Development, In proceedings Proceedings of the International Scientific Conference “Informatics, Mathematics, Education and their Applications” - IMEA’2022, pp 13-26, 23-25 November 2022, Pamporovo, Bulgaria.
3. D. Dobrilovic, Implementing AI on Microcontrollers in Fog and Edge Architectures, Proceedings of 4th Annual International Conference on Data Science, Machine Learning and Blockchain Technology - AICDMB 2023 (In Press), 16th-17th March, Mysuru, India.
4. Colour detection using Python, Jatin Mutneja, Rohan Vesuwala, supervision Dr. Jagpreet Sidhu, Department of Computer Science & Engineering/Information Technology, December, 2020.
5. Raguraman P., Meghana, A., Navya Y., Karishma Sk., Iswarya S., Color Detection of RGB Images Using Python and OpenCv, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, pp. 109-112, 2021, doi: 10.32628/CSEIT217119.
6. Open Source For You, URL: https://www.opensourceforu.com/2018/09/using-opencv-for-machine-learning-in-real-time-computer-vision-and-image-processing/, accessed: 20.09.2023.
7. Towards Data Science, URL: https://towardsdatascience.com/color-identification-in-images-machine-learning-application-b26e770c4c71, accessed: 20.09.2023.
8. K. Goyal, K. Agarwal, and R. Kumar, Face detection and tracking: Using OpenCV, 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2017, pp. 474-478, doi: 10.1109/ICECA.2017.8203730.
9. A. Sharma, J. Pathak, M. Prakash, and J. N. Singh, "Object Detection using OpenCV and Python," 2021 3rd International Conference on Advances in Computing, Communication Control and

Networking (ICAC3N), Greater Noida, India, 2021, pp. 501-505, doi: 10.1109/ICAC3N53548.2021.9725638.
10. Tutorials Point, URL: https://www.tutorialspoint.com/how-to-install-opencv-in-python, accessed: 20.09.2023.
11. Mihajlović S., Ivetić D., Berković I., Dobrilović D.: Thermal Images Analyzes of Platform for Solar Radiation Data Acquisition, XII International conference on Applied Internet and Information Technologies – AIIT, 14. Oktobar, 2022, Zrenjanin, pp. 269-274, ISBN 978-86-7672-361-4, UDK 004(082) (0.034.4)
12. Mihajlović S., Kupusinac A., Ivetić D., Berković I.: The Use of Python in the field of Artifical Intelligence, XI International Conference of Information Technology and Development of Education – ITRO, Zrenjanin, Oktobar, 2020, pp. 136-140, ISBN 978-86-7672-341-6, UDK: 37.01:004(082)
13. Guobo X. and Wen L., Image edge detection based on opencv, International Journal of Electronics and Electrical Engineering 1.2 (2013): 104-106.
14. Suning Gong, Rakesh Kumar, and D. Kumutha, Design of Lighting Intelligent Control System Based on OpenCV Image Processing Technology, International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, Vol. 29, No. Supp01, pp. 119-139, 2021, https://doi.org/10.1142/S0218488521400079
15. Ammar Anuar, Khairul Muzzammil Saipullah, Nurul Atiqah Ismail, Yewguan Soo, OpenCV Based Real-Time Video Processing Using Android Smartphone, International Journal of Computer Technology and Electronics Engineering (IJCTEE) Volume 1, Issue 3, ISSN 2249-6343

Light Sensor Analyses for Usage in Open-Source Hardware Platforms for Solar Data Acquisition

Sinisa Mihajlovic, Milica Mazalica, Jovana Borovina, Dalibor Dobrilovic, Jasmina Pekez

University of Novi Sad / Technical Faculty "Mihajlo Pupin", Đure Đakovića St., 23000 Zrenjanin, Serbia

sinisa.mihajlovic@tfzr.rs; milica.mazalica@tfzr.rs; jovana.borovina@tfzr.rs; dalibor.dobrilovic@tfzr.rs; jasmina.pekez@tfzr.rs

Abstract:
The energy received from the sun in the form of electromagnetic radiation is solar radiation. A device that converts light energy into an output electrical signal is a light sensor. Different types of light sensors are used to measure brightness, react to changes in the amount of light received, or convert light into electricity. This paper presents an exploratory analysis of light sensors for use with Arduino-based solar data collection platforms. This paper provides an overview of available low-cost light sensors with the ability to interface with Arduino. Light sensors can be analog or digital. They can operate in the IR, UV, or visible light spectrum, or all three. Different types of sensors are presented in the paper. The paper shows examples of sample code and examples of wiring diagrams. At the end of the paper, the comparison of the described sensors is given with the proposal of expanding the existing platform for solar data collection based on Arduino UNO and described sensors.

Keywords:
Open-source hardware, Arduino UNO, light sensors, solar radiation, solar radiation data collection

1. Introduction

Broadly speaking, solar radiation is the transfer of energy from the Sun to all directions in space. Sunlight, or solar light, is the light produced by the sun. The sun produces energy through thermonuclear reactions on its surface, which are manifested, among other things, by the radiation of photons to the planets in the solar system. On its way to the Earth's surface, sunlight passes through a dense layer of the atmosphere, which filters its ultraviolet radiation and thus removes its effects, which are harmful to living beings on Earth [1].

The light sensor is a passive device that converts the light energy into an electrical signal output. Light sensors are more commonly known as photoelectric devices or photo sensors because they convert light energy (photons) into electronic signals (electrons). The work of those sensors is based on internal photoelectrical effects resulting in electron flow or electric current. Light sensors are a type of photodetector (or photosensors) that detect light. Different types of light sensors can be used to measure illuminance, respond to changes in the amount of light received, or convert light to electricity [2].

The most common types of light sensors are photodiodes, photoresistors, phototransistors, and photovoltaic light sensors. These components can be used in applications such as light sensing in mobile devices, automatic outdoor lighting, proximity sensors, and renewable energy. There are several types of light sensors based on the method of conversion of the measured physical quantity into measurement information: sensors with analog conversion, sensors with analog signal conversion in signal with variable period or frequency, and sensors with analog-digital conversion.

There are several types of sensors with analog signal: sensors with photoelectronic emission (external photo effect), sensors with photoconductivity effect (internal photo effect), sensors with photovoltaic effect (photocells), sensors with change in resistance of p-n junction and sensors with pyroelectric effect. The output signal of optical sensors with periodic or frequency output is a pulse signal with variable duration or with variable frequency.

In the case of optical sensors with analog-digital conversion, the output signal is a code that is proportional to the change in the analog input signal. Sensors with A/D conversion are the most promising optical sensors. They are divided into two large groups. The first group includes sensors with sequential coding, and the second group includes sensors with parallel coding [3].

This paper deals with the usage of open-source hardware and low-cost light (photo) sensors for building an Arduino-based platform for solar radiation data collection. Data collected with the platform are ambient temperature and humidity, visible and UV light intensity, solar panel temperature and voltage output as it was described in [4, 5, 6]. All collected data are used to analyze the impact of various ambient parameters on solar the panel performance. The contribution of this paper is a survey and analysis of available low-cost light sensors, with a special focus on their characteristics, and the possibility to integrate with Arduino platforms. The paper presents the plan for platform expansion based on research findings, and on the author's experience with the already created and tested platform, different sensors, and their performances as published in previous works [4, 5, 6]. The planning of the expansion of the existing platform takes into the consideration performances of the development boards and their limitation with the goal to find the optimal combination of sensors and board capabilities. For example, Arduino UNO and MEGA have enough analog pins, but no integrated communication modules. The ESP8266 boards (NodeMCU, Wemos D1 and clones) have only one analog pin, but integrated communication module (Wi-Fi). The ESP32 boards have enough analog pins and integrated communication modules (Wi-Fi, Bluetooth). The communication nodules are important for integration of this data acquisition platform in the wireless sensor networks. The research findings are used to find the efficient and budget friendly expansion of the existing platforms which are mainly based on Arduino UNO and ESP8266. So far, in these platforms, UV analog sensors and BH1750 have been used. The goal of this expansion is to build a platform for more effective data collection. This paper is structured as follows. After the introduction, the related work is presented. A detailed overview of available light sensors is given in the third section. In the fourth section with sensor comparison, the proposed platform expansion is presented. In the end, the conclusion of the research findings is given.

2. Related work

Arduino is an open-source hardware electronic platform based on hardware and software components. Arduino boards can read inputs such as light sensing, a button press, or Twitter messages and convert these inputs into outputs, for example, to activate a motor, turn on an LED, or perform tasks on a network. Data collected with Arduino products can be stored on an SD card, transferred with Wi-Fi to the cloud, or via serial port directly to a PC.

The Arduino UNO and clone boards are often used for building prototypes and for system design for various purposes. In this section will be presented several of these projects. The paper [7] presents an approach that integrates PLCs and the Arduino platform into a common SCADA system using connectivity provided by Open Platform Communications (OPC). In the paper [8], a prototype was created based on affordable open-source hardware and computer vision to test control algorithms developed in Mathematica and Simulink. The aim of developing this prototype is to facilitate learning about solar energy. It helps in understanding the fundamentals of solar systems and offers the possibility of working in other fields connected to solar concentration systems.

The [9] provides an overview of several light sensors: a photo-resistor, UV/IR/Visible light (SI1145), high dynamic range LUX(TSL2591), digital luminosity LUX(TSL2561), and an analog light (GA1A1S202WP). It examines a subset of sensors suitable for specific microcontrollers and evaluates their qualities. It is assessed that the most crucial quality of these sensors is their light sensitivity and analog sensors are preferred by this author for basic-lever applications.

The paper [10] proposes a data acquisition system (DAS) with the capability to collect voltage and current data in real-time. The system is cost-effective, with analog voltage and current sensors connected to an open-source Arduino platform for real-time data storage on an SD card. As a result, power losses are minimized. In this paper [11] it is described a method for reducing electrical energy consumption using light-dependent resistors (LDR) sensors. A controller for reading the sensors was tested in a small room with two LED lights. Two LDR sensors were used to measure one LED bulb and

the system was tested during the day. It has been demonstrated that this system can effectively reduce electricity consumption during daylight hours.

3. Type of light sensors

Generally, considering the connectivity to Arduino UNO and clone boards, authors can classify sensors into two major groups: analog, and digital sensors (I^2C bus). In the following section, the most popular types of low-cost light sensors of both types used for interfacing with Arduino UNO and clone boards will be presented. Besides the specification of the sensors in the paper will be presented the wiring schemes with Arduino UNO and clone boards, as well as a couple of simple code examples for Arduino UNO.

3.1. Analog light sensors

LDR (light-dependent resistor) is the most common and very cheap analog light sensor module that uses a GL5528 photoresistor to detect the ambient light intensity. The resistance of the sensor decreases when the ambient light intensity increases. An LM358 chip is used as a voltage follower to allow accurate readings [12]. This sensor has low power consumption, and it is used in a variety of applications such as electronic toys, light-control switch monitors, etc. It is very easy to be implanted in prototypes and projects, especially for beginners. The simple Arduino code for the LDR sensor is given in Listing 1, and the wiring with the Arduino is Shown in Fig. 1. It can be connected with Arduino in combination with a 10 kΩ resistor, or with a breakout board with an integrated resistor. For better understanding the code, the variable *rawRange* represents the maximal number of values on the Arduino UNO analog pin. Arduino UNO has 10-bit analog to digital converter, giving the values on analog pins in the range 0-1023. The variable longRange is variable that represents 5V operating voltage of Arduino UNO microcontroller board, while operating voltage of NodeMCU is 3.3.V. Both values are needed to convert Arduino UNO analog read values to lux values.

Another example of an analog sensor is UV light sensor - GUVA-S12SD [13]. It uses a UV photodiode, which can detect the 240-370nm range of light (which covers UVB and most of the UVA spectrum). Similar to the LDR sensor it is connected with three wires to the Arduino or any other microcontroller board: one to 2.7-5.5VDC, one to GND, and one to read the analog signal from the OUT pin. Fig. 1 (right side of the image) shows the wiring of the analog UV sensor to NodeMCU with ESP8266 chip.

```
float rawRange = 1024;
float logRange = 5.0;

void setup() {
 Serial.begin(9600);
 pinMode(2, OUTPUT);
}

void loop() {
 int sensorValue = analogRead(2);
 float luxValue = sensorValue * logRange / rawRange;
 float realValue = pow(10, luxValue);
 Serial.print("Light intensity: ");
 Serial.print(sensorValue);
 Serial.print(", ");
 Serial.print("Light intensity: ");
 Serial.println(sensorValue);
 delay(500);
}
```

Listing 1. Sample code for LDR light sensor

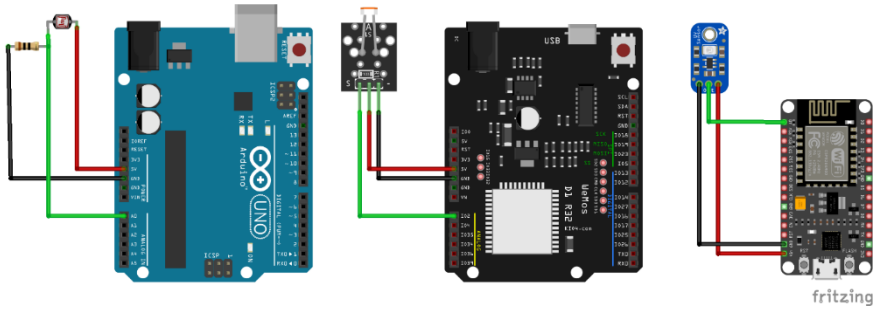


Figure 1. Wiring diagram for Arduino UNO and clone boards with analog light sensors. From left to right: LDR sensor with 10kΩ resistor and Arduino UNO board, LDR sensor with breakout board and Wemos D1 R32 board with ESP32 module, and UV analog sensor with NodeMCU ESP8266 board. Image is generated with Fritzing [14].

CJMCU 101 sensor can be used instead of the LDR sensor. This sensor module is created using the OPT101 light sensor chip. This chip includes a monolithic photodiode and a trans-impedance amplifier. One of the advantages of the sensor is that the amplifier is single or dual designed for power supply, which can run on battery equipment. When the photodiode operates in the photoconductive mode, it has good linearity and low dark current. Through this sensor, the light intensity can be obtained as an analog value. This chip requires a potential of 2.7V to 36V to operate and can be operated from a single or dual supply. This sensor is used for barcode scanners, money changers, smoke detectors, proximity sensors, medical instruments, laboratory instruments, and photographic analyzers. The main features of the sensor are as follows: size of photodiode: 2.29mm x 2.29mm, internal feedback resistor: 1MΩ, low quiescent current: 120uA, high responsivity: 0.45A/W, and working temperature: 0-70°C [15, 16].

3.2. Digital light sensors

The ideal sensor for use in a wide range of lighting situations is the *TSL2561 Advanced Digital Light Sensor*. This sensor is more accurate, compared to cheap LDR sensors, allows accurate lux calculations, and can be configured for different gain/time signal ranges to detect light ranges from 0.1 – 40,000+ lux. The sensor contains infrared diodes and full-spectrum diodes. The sensor can measure infrared light, full-spectrum light, or human-visible light separately. Most sensors can only detect one or the other, which does not accurately represent what the human eye sees. IR light detected by most photo diodes can't be detected by humans. There is a version with a 3.3V regulator and level shifter circuit so it can be used with any 3-5V power/logic microcontroller. The sensor has a digital (I^2C) interface. It can select one of three addresses so it can have up to three sensors on one board – each with a different I^2C address. The built-in ADC means it can be used with any microcontroller, even if it doesn't have an analog input. Current consumption is extremely low, making it great for low-power data logging systems, around 0.5mA when actively sensing, and less than 15µA when it shutdown mode [16]. The example of wiring TLS2561 with Arduino UNO is shown in Fig. 2 (left).

The main features of the sensor are as follows: approximates human eye response, temperature range: -30 to 80°C, dynamic range (Lux): 0.1 to 40,000 lux, voltage range: 2.7-3.6V, Interface: I2C with I2C 7-bit addresses 0x39, 0x29, 0x49, selectable with jumpers [17].

The *VEML7700* is another popular sensor. It has a 16-bit dynamic range for ambient light detection from 0 lux to 120 lux with a resolution of up to 0.0036 lx/ct, with software adjustable gain and integration time. The connection is simple, the sensor uses a simple universal I^2C. This sensor is mounted on a breakout board with a 3.3V regulator and logic level shifter so it can be used with 3.3V or 5V power/logic microcontrollers. A library has been written for Arduino (C/C++) as well as CircuitPython (Python 3). This sensor can be used with any type of device, even a Raspberry Pi [18]. The example of wiring VEML7700 with ESP32 board is shown in Fig. 2 (right).

Grove – Sunlight Sensor is a multi-channel digital light sensor, which can detect UV light, visible light, and infrared light. This device is based on the SI1151, a new sensor from SiLabs. The SI1151 is a low-power, reflectance-based infrared proximity, UV index, and ambient light sensor with an I^2C digital interface and a programmable event interrupt output. It offers excellent performance in a wide

dynamic range and various light sources including direct sunlight. The Grove – Sunlight Sensor includes a built-in Grove connector, which allows connection to the Arduino. It can also be used with a Raspberry Pi weather station or a smart irrigation system using Arduino if it needs to monitor the visible spectrum. Features of the sensor are as follows: a multi-channel digital light sensor that can detect UV light, visible light, and infrared light, wide spectrum detection range: 280-950nm, I2C Interface (7-bit), compatible with Grove port plug-and-play, low power consumption: 3.3/5V Supply, suitable for many microcontrollers and SBCs (Single Board Computers) [19].

The *Avago/Broadcom APDS 9930* can be easily used in various projects together with SBCs and MCUs. Those projects may need a touchless motion sensor for controlling a computer, microcontroller, robots, or home devices, the automotive industry motion switching with gestures, or measuring ambient light and color and proximity detection. The APDS 9930 provides a digital ambient light sensor, an IR LED, and a complete proximity detection system in one chip. The proximity function offers plug-and-play detection up to 100mm without front glass. The proximity detection function works well from bright sunlight to dark areas. The wide dynamic range allows operation in the ability to put the device in low power mode between ALS (Ambient Light Sense) and proximity measurement. The APDS 9930 is useful for screen management to extend battery life and ensure optimal viewing in different lighting conditions [20].

The Adafruit *BH1750 Ambient Light Sensor* is a 16-bit I^2C ambient light sensor designed by Rohm Semiconductor. It is a small, capable, and inexpensive light sensor used for light detection and measurement. The BH1750 can measure from 0 to 65K+ lux. With calibration and advanced timing settings, it can even measure up to 100,000 lux. It comes integrated with a voltage regulator and level-shifting circuit to allow it to be used with 3.3V devices such as the Feather M4 or Raspberry Pi, or 5V devices such as the Arduino. The printed circuit board it's packaged on breakout board with pins on a standard 0.1 inch/2.54mm pitch header. The library used is compatible with CircuitPython devices, as well as Raspberry Pi, by installing PyPi [21]. The example of wiring BH1750 with an Arduino UNO board is shown in Fig. 2 (center). The example Arduino code for BH1750 is given in Listing 2. [22]. For better understanding the code, BH1750 library (`#include <BH1750.h>`) is used for reading sensor measured values (`lightMeter.readLightLevel()`) in lux (lx) units of illuminance.

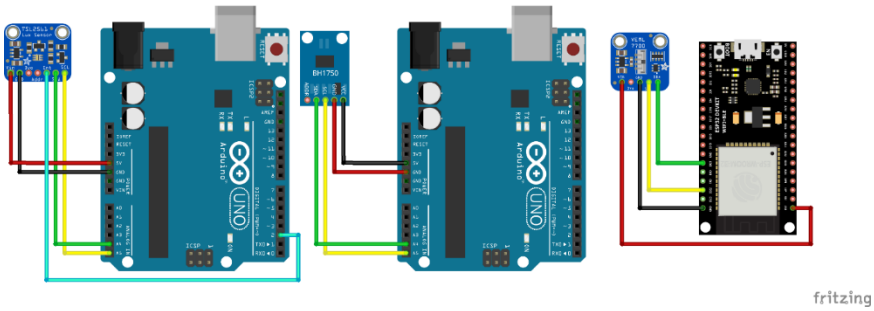


Figure 2. Wiring diagram for Arduino UNO and clone boards with I2C sensors. From left to right: TSL2561 and Arduino UNO board, BH1750 with Arduino UNO board, and VEML7700 with ESP32 board. Image is generated with Fritzing [14]

```
#include <Wire.h>
#include <BH1750.h>
BH1750 lightMeter;

void setup(){
  Serial.begin(9600);
  Wire.begin();
  lightMeter.begin();
  Serial.println(F("BH1750 Test begin"));
}

void loop() {
  float lux = lightMeter.readLightLevel();
  Serial.print("Light: ");
```

```
  Serial.print(lux);
  Serial.println(" lx");
  delay(1000);
}
```

Listing 2. Sample code for BH1750 I²C light sensor

The *AP3216* is an integrated ALS and PS module that includes a digital ambient light sensor (ALS), a proximity sensor (PS), and an IR LED in one package. This module is suitable for applications under clear or tinted glass. The proximity function is specifically aimed at near-field applications and detects an external object with a simple configurable zone controlled by registers. A 220 to 470 Ω resistor is required to connect to this module. The AP3216 has two photodiodes. One detects light in the visible range and the other in the infrared range. The photodiode voltages are amplified and digitized using an A/D converter. The visible light result is stored by the AP3216 as a 16-bit value in the ALS data register. The result from the IR measurement is stored as a 10-bit value in the IR data register. To measure distance, the LED emits infrared light. An IR photodiode detects reflected radiation in addition to ambient IR light. Based on this, the AP3216 calculates the distance as a 10-bit value and stores it in the PS data register. Data stored in the ALS data register can be converted to a lux value. PS value cannot be converted directly to distance [23].

The most important technical data for the ambient light sensor are four adjustable lux ranges and calibration functions. For proximity sensor (PS) important data are: range from 2 cm up to approx. 30 cm depending on the settings, calibration function, four gain levels, and adjustable measuring times. For the IR LED sensor important data are: selectable number of pulses per measurement, setting the LED current. Additionally, the sensor has an interrupt function: for PS and ALS or both; active-low, continuous, or single-shot measurements. The supply voltage of the sensor is 2.4 – 3.6 volts and has a power consumption of 1.7 mA. The sensor is I²C with address 0x1E (not variable). [23]

Phidgets Light Sensor 70,000 lux can measure ambient light up to 70 kilolux (roughly equivalent to direct sunlight). Each sensor is individually calibrated, and a label is affixed to the back of the plate with a calibration value that can be used in calculations to increase measurement accuracy. The sensor output is logarithmic, so it will be more accurate in low light. It can measure incandescent and fluorescent light. Sensor output type: non-radiometric. Maximum sensitivity wavelength 560 nm. The maximum response time of the sensor is 0.5 ms, the light level ranges from 3 lx to 70 klx, light to current ratio is 1.2, the current consumption is Max 500 μA, the supply voltage is 4 to 5.5 V DC, and the operating temperature ranges from -40°C to 85°C [24].

4. Sensor comparison and proposed expansion of platform

The comparison of the presented sensor is given in Table I.

Table I – Comparison of light sensors

No.	Sensor	Type	Light Level	Power Supply	Wavelength	Operating temp.	Price	Arduino
1	TSL2561	I²C	0.1 – 40,000 lx	2.7 – 3.6V	280nm – 950nm	-30 – 80°C	-6 $	yes
2	VEML7700	I²C	0 – 120 lx	3.3 – 5V	545nm	-25 – 85°C	-5 $	yes
3	Grove-Sunlight	I²C	1 – 128 klx	3.3 – 5V	280nm – 950nm	-45 – 85°C	-12 $	yes
4	APDS 9930	I²C	0,01 lx	2.2 – 3.6V	625nm	-40 – 85°C	-2 $	yes
5	BH1750	I²C	0 – 65K+ lx	3.3 – 5V	400nm – 700nm	-40 – 85°C	-5 $	yes
6	LDR	analog	N/A	2.7V – 5.5V	540nm	-30 – 70°C	-1 $	yes
7	GUVA-S12SD	analog	N/A	2.7V – 5.5V	240nm – 370nm	-30 – 85°C	-7 $	yes
8	AP3216	I²C	0 – 20661 lx	2.4V – 3.6V	595nm – 700nm	-40 – 85°C	-8 $	yes
9	CJMCU 101	analog	N/A	2.4V – 36V	650nm	0 – 70°C	-7 $	yes
10	Phidgets light	I²C	3 lx - 70 klx	4V – 5.5V	560nm	-40 – 85°C	-7 $	yes

After the overview of the presented sensors the expansion of the platform is planned with the following configuration: (1) Arduino UNO development board, (2) solar panel, (3) VEML7700 I²C, (4) SI1151 I²C, (5) BH1750 I²C, (6) DHT-22, (7) TMP 36, (8) voltage sensor, (9) UV GUVA-S12SD, (10) LDR with breakout board sensor, (A) SCL I²C bus, (B) SDA I²C bus, (C) GND bus, and (D) VCC bus. The proposed configuration is shown in Fig. 3. Fig. 3 shows a platform with three I²C sensors, although I²C bus can connect up to eight devices per bus. Only three sensors are presented in Fig. 3, because of better visibility, but in practice, the authors consider wiring from 3 to 4 I²C sensors (respectively VEML7700, SI1151, TLS2561, and BH1750). The GUVA-S12SD UV and LDR sensors are used as

analog sensors. The Arduino UNO has the limitation of up to 6 analog sensors because of the number of headers. Two headers are reserved for I^2C (A4, A5), and two other headers are used for the voltage sensor (A3), and for the TMP36 sensor (A2). TMP36 is used for measuring solar panel temperature, and the voltage sensor is used for measuring solar panel output. The platform is focused on finding the relationship between light intensity and solar panel output. The Arduino UNO also has 14 digital outputs and only one digital output is used with the DHT-22 sensor for measuring ambient temperature and humidity data. So, the next expansion of the platform will be mainly focused on adding digital sensors. The BH1750 I^2C and GUVA-S12SD UV sensor, as well as the voltage sensor accuracy, is proved in the previous version of the platform [2, 3, 4].

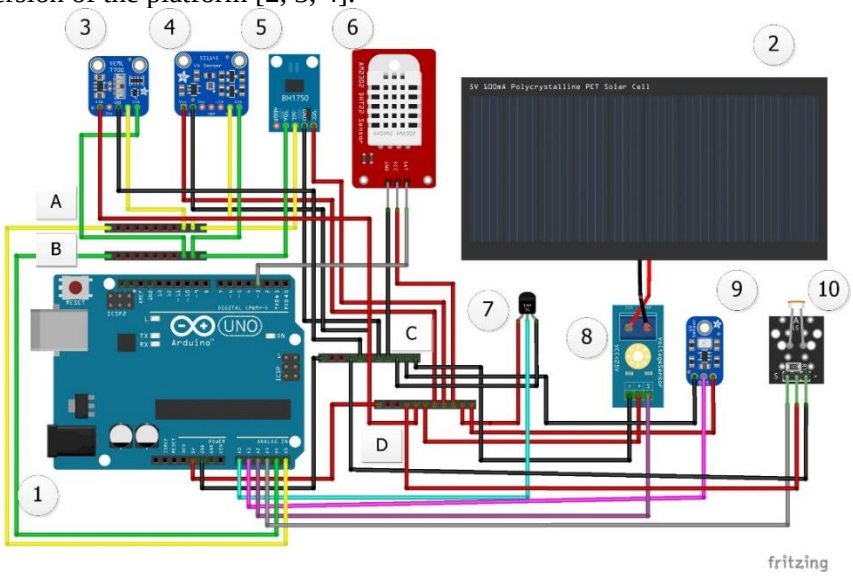


Figure 3. The proposed expansion of Arduino UNO-based solar data acquisition platform. Image is generated with Fritzing [14].

5. Conclusions

This paper is focused on using an open-hardware source platform based on Arduino and clone devices for solar radiation data acquisition. The Arduino platform can interface with various sensors for measuring light intensity. This paper gives an overview of low-cost photo sensors that can be used with such platforms. In summary, this paper presents different types of light sensors that can be used to collect solar data and to predict the voltage output of solar panels. These sensors can be analog or digital. The description of the sensors includes some important characteristics such as light intensity level, power supply, wavelength range, operating temperature, price, and capabilities of interfacing with Arduino. Wiring diagrams are shown for the majority of the presented sensors. In addition, sample codes for reading sensor outputs are presented. At the end of the paper the comparison of sensors, with the proposed expansion of the existing platform for solar radiation data acquisition is given.

Acknowledgment:

The research is conducted through the project »Creating laboratory conditions for research, development, and education in the field of the use of solar resources in the Internet of Things«, at the Technical Faculty "Mihajlo Pupin" Zrenjanin, financed by the Provincial Secretariat for Higher Education and Scientific Research, Republic of Serbia, Autonomous Province of Vojvodina, Project number 142-451-2684/2021-01/02.

References:

1. A. S. Olpenda, S. Krzysztof, B. Krzysztof, Modeling Solar Radiation in the Forest Using Remote Sensing Data: A Review of Approaches and Opportunities, *Remote Sensing*, 2018, 10, no. 5:694. URL: https://doi.org/10.3390/rs10050694
2. M. Popovic, Sensors and measurements (In Serbian), E. Sarajevo, Bosnia and Herzegovina 2004.

3. B. Radičević, Đ. Vukić, Đ. Ercegović, M. Oljača, Optical Sensors and their Application on Agricultural Machines (In Serbian), Agricultural Technique, pp. 123 – 136, 2009.
4. D. Dobrilovic, J. Pekez, E. Desnica, Lj. Radovanovic, I. Palinkas, M. Mazalica, L. Djordjević, S. Mihajlovic, Data Acquisition for Estimating Energy-Efficient Solar-Powered Sensor Node Performance for Usage in Industrial IoT, Sustainability 2023, 15, 7440. https://doi.org/10.3390/su15097440
5. D. Dobrilovic, Approach in Using Open-Source Hardware in Teaching Solar-Powered IoT Node Development, In proceedings Proceedings of the International Scientific Conference "Informatics, Mathematics, Education and their Applications" - IMEA'2022, pp 13-26, 23-25 November 2022, Pamporovo, Bulgaria.
6. D. Dobrilovic, Implementing AI on Microcontrollers in Fog and Edge Architectures, Proceedings of 4th Annual International Conference on Data Science, Machine Learning and Blockchain Technology - AICDMB 2023 (In Press), 16th-17th March, Mysuru, India,
7. I. González, J. A. Calderón, Integration of open source hardware Arduino platform in automation systems applied to Smart Grids/Micro-Grids, Sustainable Energy Technologies and Assessments, Vol. 36, 2019, 100557, URL: https://doi.org/10.1016/j.seta.2019.100557.
8. A. J. Carballo, J. Bonilla, L. Roca, M. Brenguel, New low-cost solar tracking system based on open source hardware for educational purposes, Solar Energy, Vol. 174, 2018, pp 826-836, URL: https://doi.org/10.1016/j.solener.2018.09.064.
9. J. Carpenter, Light Sensor Performance Comparisons, In Proceedings of the 15th Winona Computer Science Undergraduate Research Symposium, 2015.
10. K. P. Pachauri, P. O. Mahela, B. Khan, A. Kumar, S. Agarwal, H. H. Alhelou, J. Bai, Development of Arduino assisted data acquisition system for solar photovoltaic array characterization under partial shading conditions, Computers & Electrical Engineering, Vol. 92, 2021, 107175, URL: https://doi.org/10.1016/j.compeleceng.2021.107175.
11. M. G. Salim, H. Ismail, N. Debnath, A. Nadya, Optimal light power consumption using LDR sensor, In Proceedings of IEEE International Symposium on Robotics and Intelligent Sensors (IRIS), Langkawi, Malaysia, 2015, pp. 144-148, URL: https://doi.org/10.1109/IRIS.2015.7451601.
12. Adafruit, LDR sensor, https://www.adafruit.com/product/161, accessed 10.09.2023.
13. Adafruit. GUVA-S12SD UV sensor, https://www.adafruit.com/product/1918, accessed 10.09.2023.
14. A. Knörig, R. Wettach, J. Cohen, Fritzing: A tool for advancing electronic prototyping for designers, in: Proceedings of the 3rd International Conference on Tangible and Embedded Interaction, 2009, pp. 351–358
15. Texas Instruments, OPT 101, OPT101 Monolithic Photodiode and Single-Supply Transimpedance Amplifier, Texas Instruments Incorporated, 2023.
16. SriTu Hobby, CJMCU-101. URL: https://srituhobby.com/cjmcu-101-opt101-analog-light-sensor-with-arduino/, accessed 10.09.2023.
17. Adafruit, TLS2561, URL: https://www.adafruit.com/product/439, accessed 10.09.2023.
18. Adafruit, VEML7700, URL: https://learn.adafruit.com/adafruit-veml7700/overview, accessed 10.09.2023.
19. Seeedstudio, Grove – Sunlight Sensor, URL: https://shorturl.at/bfikr, accessed 10.09.2023.
20. AUTODESK Instructables, URL: https://www.instructables.com/Wiring-the-APDS-9930-Ambient-Light-Sense-APDS-9960/, accessed 10.09.2023.
21. Adafruit, BH1750, URL: https://learn.adafruit.com/adafruit-bh1750-ambient-light-sensor/overview, accessed 10.09.2023.
22. Github, URL: https://github.com/claws/BH1750 , Path: examples/BH1750test/BH1750test.ino
23. Wolles Elektronikkiste, URL: https://wolles-elektronikkiste.de/en/ap3216-cjmcu-3216-ambient-light-and-proximity-sensor, accessed 10.09.2023.
24. RobotShop, Phidgets Light Sensor 70,000 lux, URL: https://www.robotshop.com/products/phidgets-light-sensor-70000-lux, accessed 10.09.2023.

Smart City - Belgrade: Opportunities and Challenges

Mirjana Tomic[1], Kostadinka Stojanovic[2], Dragan Zlatkovic[3], and Nebojsa Denic[4]

[1] *Alfa BK University, Faculty of Information Technologies, Belgrade, Serbia*
[2] *Alfa BK University, Faculty of Mathematics and Computer Sciences, Belgrade, Serbia*
[3] *University "UNION-Nikola Tesla", Faculty of Information Technologies and Engineering, Belgrade, Serbia*
[4] *University of Pristina, Faculty of Sciences and Mathematics, Kosovska Mitrovica, Serbia*

tomicmirjana3@gmail.com;kostadinka.stojanovic@alfa.edu.rs;dragan.zlatkovic@fiti.edu.rs; nebojsa.denic@pr.ac.rs

Abstract:
Smart City initiatives are gaining popularity around the world to improve urban life using technology. Belgrade, the capital city of Serbia, has begun its own journey towards becoming a Smart City, with several initiatives and projects underway. Belgrade's transformation into a Smart City presents opportunities for enhancing urban sustainability, efficiency, and citizen engagement. Overcoming obstacles like funding, collaboration, and privacy concerns is crucial. By doing so, Belgrade can become a more livable city. Smart City initiatives also enable citizen participation in governance through e-governance platforms, citizen feedback systems, and social media, fostering transparency and accountability. Privacy and security considerations are vital, requiring the establishment of protective policies. Belgrade's smart parking system exemplifies technology's positive impact, offering practical solutions to improve daily life. Moreover, smart city air pollution systems leverage advanced technologies and data analytics to provide real-time solutions for enhancing air quality and reducing health risks in urban areas. This paper explores the opportunities and challenges associated with the development of a Smart City in Belgrade, with a focus on its potential to improve urban sustainability, efficiency, and citizen's engagement.

Keywords:
Citizen's engagements, development, efficiency, projects, Smart City

1. Introduction

Belgrade is the largest city and the capital of Serbia, with a population of around 1.7 million people. It is a city with a rich history, culture, and heritage, but also faces a few challenges related to urbanization, infrastructure, and environmental sustainability. The concept of a Smart City is gaining popularity around the world to address these challenges, and Belgrade is no exception. In recent years, the city has started to adopt Smart City initiatives and projects, with the aim of improving the quality of life for its citizens. The comparative method of analyzing available data was used within the case study.

2. Smart City

A smart city is a city that uses advanced technology and data analysis to improve the quality of life of its citizens, enhance sustainability, and streamline urban services. The term "smart city" refers to the integration of various technological solutions and the use of data to optimize urban systems and infrastructure [1], [2].

A smart city typically utilizes sensors, cameras, and other Internet of Things (IoT) devices to gather real-time data about various aspects of urban life, such as traffic flow, air quality, energy consumption,

and water usage through artificial intelligence (AI). This data is then analyzed and used to make more informed decisions about how to manage and improve the city [3].

In a smart city, technology is also used to improve citizen engagement and participation in civic life. For example, citizens may have access to mobile apps that allow them to report issues, provide feedback, and access information about city services and events [4].

The goal of a smart city is to create a more efficient, sustainable, and livable urban environment, while also promoting economic growth and innovation [5].

3. Opportunities of a Smart City in Belgrade

The development of a Smart City in Belgrade offers numerous opportunities for improving urban sustainability, efficiency, and citizen engagement. One of the key areas where Smart City initiatives can make a significant impact is in the area of transportation. Belgrade has a high number of vehicles on the road, which contributes to traffic congestion and air pollution. By implementing technologies such as intelligent traffic management systems, electric vehicles, and bike-sharing schemes, the city can reduce congestion and improve air quality [6].

Another opportunity presented by Smart City initiatives is the potential for energy efficiency and sustainability. Belgrade has a high demand for energy, which is met primarily by fossil fuels. The implementation of Smart Grids, renewable energy sources, and energy-efficient buildings can help the city reduce its carbon footprint and become more environmentally sustainable [7].

3.1. The Smart City initiatives

The smart city initiatives in Belgrade offer significant potential for improving energy efficiency and sustainability and city could focus its efforts in the following areas [2], [8]:

Smart buildings: One of the most significant contributors to energy consumption in cities is buildings. Implementing smart building technologies such as intelligent lighting, heating, and cooling systems, and building automation could significantly reduce energy consumption and costs.

Renewable energy sources: Belgrade has significant potential for implementing renewable energy sources such as solar and wind energy. By investing in renewable energy infrastructure, the city could reduce its reliance on fossil fuels, decrease greenhouse gas emissions, and improve energy efficiency.

Smart transportation: The implementation of smart transportation systems could significantly reduce traffic congestion and air pollution in the city. For example, intelligent traffic management systems, electric and hybrid vehicles, and smart parking solutions could improve transportation efficiency, reduce fuel consumption, and improve air quality.

Waste management: Smart waste management systems such as waste sorting and recycling could reduce the amount of waste that ends up in landfills and contribute to sustainability.

3.2. Traffic congestion and air pollution

Traffic congestion and air pollution are significant issues in Belgrade, Serbia's capital city [9]. Here are some figures related to these problems:

- According to a report by the World Health Organization (WHO) in 2018, Belgrade had an average annual concentration of particulate matter (PM2.5) in the air of 24 micrograms per cubic meter, which is twice the recommended limit by WHO. In the same year, the city was ranked as the 14th most polluted city in Europe in terms of PM2.5 levels [10].
- In 2019, a study by the Serbian Ministry of Environment and Urban Planning showed that traffic was responsible for around 80% of air pollution in Belgrade [11].
- According to a survey conducted by the City of Belgrade in 2020, traffic congestion is the most significant problem for residents, with 64% of respondents stating that it is a major issue [12].

- A study conducted by the Institute of Transportation CIP in Belgrade in 2018 estimated that the economic cost of traffic congestion in Belgrade was around 750 million euros per year.

These figures highlight the urgent need for measures to address traffic congestion and air pollution in Belgrade. The city has taken some steps towards improving air quality, such as implementing low-emission zones and promoting the use of public transport and cycling. However, more comprehensive and long-term solutions are needed to tackle these significant challenges.

4. Smart City Air Pollution Systems

Smart City Air Pollution Systems are technologies and infrastructure that are designed to monitor and control air pollution in urban areas. These systems utilize sensors, data analytics, and other advanced technologies to collect and analyze data on air quality, and to provide real-time information to citizens, local governments, and other stakeholders [13], [14].

There are various types of smart city air pollution systems, some of which are [15]:

- Air quality sensors.
- Smart traffic management.
- Electric vehicle charging infrastructure.
- Green spaces.
- Air pollution reduction technologies.

4.1. Air quality sensors

These are devices that measure air pollution levels in the environment. They can be deployed in various locations throughout the city, such as on streetlamps or buildings, and can provide real-time information on air quality to city officials and residents. Air quality sensors are a key component of smart city air pollution systems and can be used to monitor and improve air quality in urban areas [16].

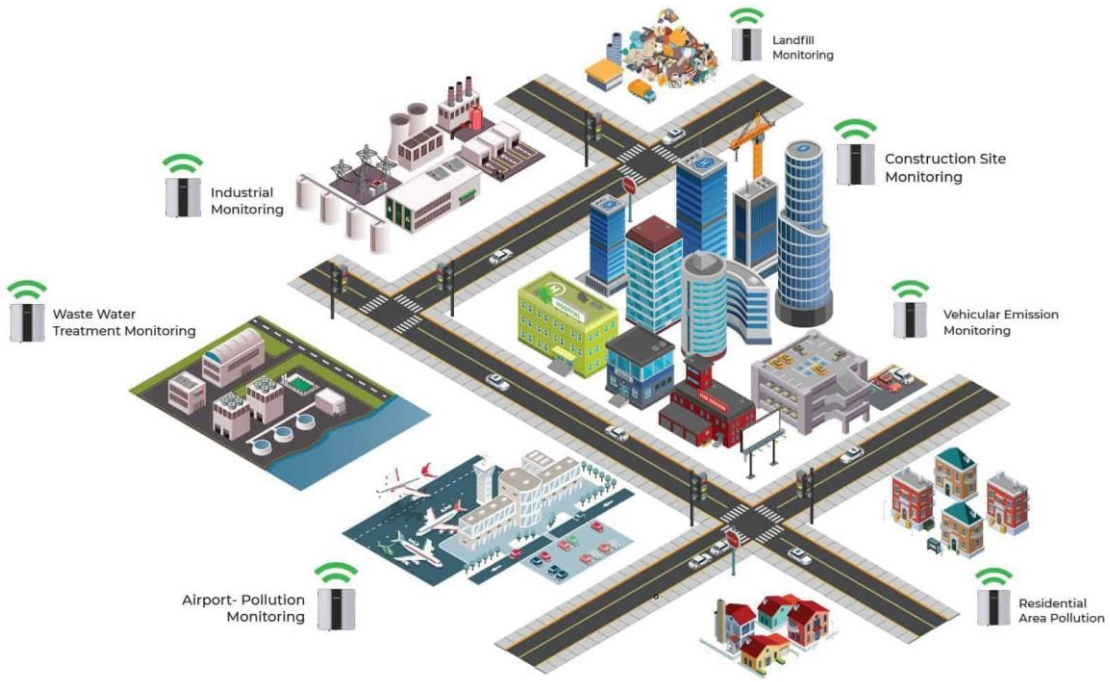


Figure 1: Air Quality Monitoring for Smart City infrastructure

There are several types of air quality sensors, including [2], [14]:

- Particulate matter (PM) sensors: These sensors measure the concentration of tiny particles in the air that can cause respiratory and cardiovascular problems. PM sensors can detect both fine (PM2.5) and coarse (PM10) particles.
- Nitrogen oxide (NOx) sensors: NOx is a common air pollutant that is emitted by vehicles and power plants. NOx sensors can measure the concentration of this pollutant in the air.
- Carbon monoxide (CO) sensors: CO is a toxic gas that is produced by the incomplete combustion of fossil fuels. CO sensors can measure the concentration of this gas in the air.
- Ozone (O_3) sensors: Ozone is a gas that is formed by the reaction of sunlight with other air pollutants. Ozone sensors can measure the concentration of this gas in the air.
- Volatile organic compound (VOC) sensors: VOCs are emitted by many sources, including vehicles, industrial processes, and consumer products. VOC sensors can detect the concentration of these compounds in the air.

According to the Environmental Agency of the Republic of Serbia (SEPA), there are currently 27 air quality monitoring stations in Belgrade, which are measuring levels of various air pollutants such as PM10, PM2.5, NO_2, SO_2, O_3, CO, and benzene. These stations are part of a national network of air quality monitoring stations in Serbia, and the data collected from them is publicly available on the agency's website.

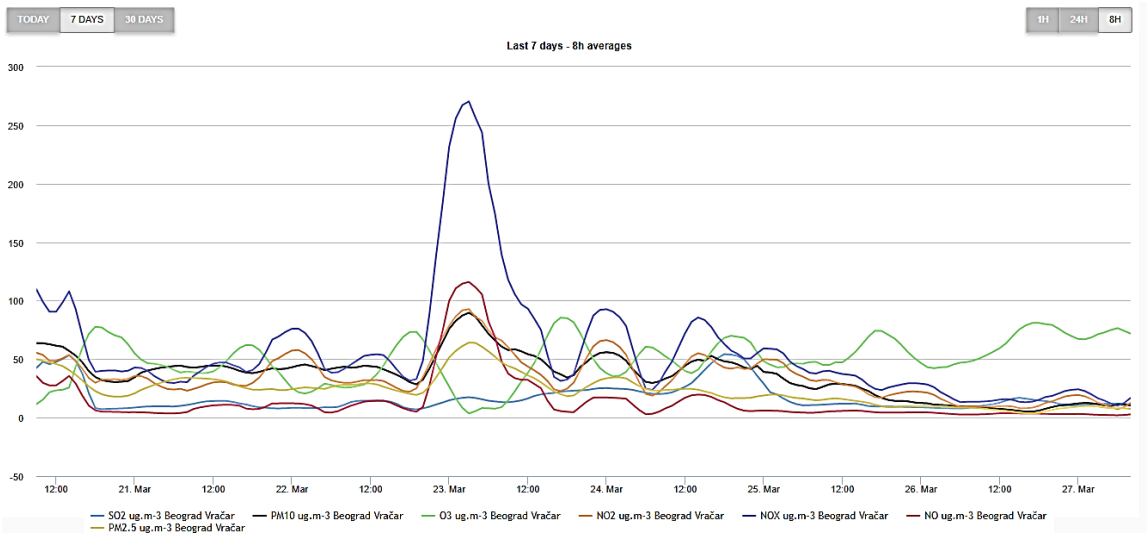


Chart 2: Data's from Belgrade "Vračar" Air Quality Monitoring Station (AQMS), 21-27 March 2023 (Source: SEPA)

4.2. Smart traffic management

Traffic is a major source of air pollution in urban areas. Smart traffic management systems use data analytics to optimize traffic flow, reducing congestion and emissions. One of good example of the smart traffic management in the city of Belgrade is the smart parking system [9].

4.2.1. Smart parking systems in Belgrade

A smart parking system is a technology-driven solution that helps drivers find available parking spots quickly and efficiently. The system uses various sensors, cameras, and other technologies to detect the presence of a vehicle and provide real-time information about parking availability to drivers [17].

With a smart parking system, drivers can easily access information about available parking spots through a mobile app or digital displays. The app provides real-time updates about available parking spaces, their location, and the cost of parking. The system can also help drivers navigate to the nearest available parking spot through GPS or other navigation tools.

Smart parking systems have several benefits, including reducing the time spent searching for parking spots, minimizing traffic congestion, and improving air quality by reducing the number of cars on the road. Additionally, these systems can help cities generate more revenue through dynamic pricing of parking fees based on demand.

One of the major issues in Belgrade is traffic congestion, and finding a parking spot in the city can be challenging. To address this problem, the city has implemented a smart parking system that enables drivers to find available parking spots quickly and efficiently.

The smart parking system is based on a network of sensors that are installed in parking lots and on-street parking spaces throughout the city of Belgrade. These sensors detect the presence of a vehicle and transmit data to a central server, which then processes the information and provides real-time information about parking availability [9].

Figure 2: Different types of parking sensors

Drivers can access this information through a mobile app, which displays available parking spots and the cost of parking. The app also provides navigation assistance to help drivers find their way to the nearest available parking spot.

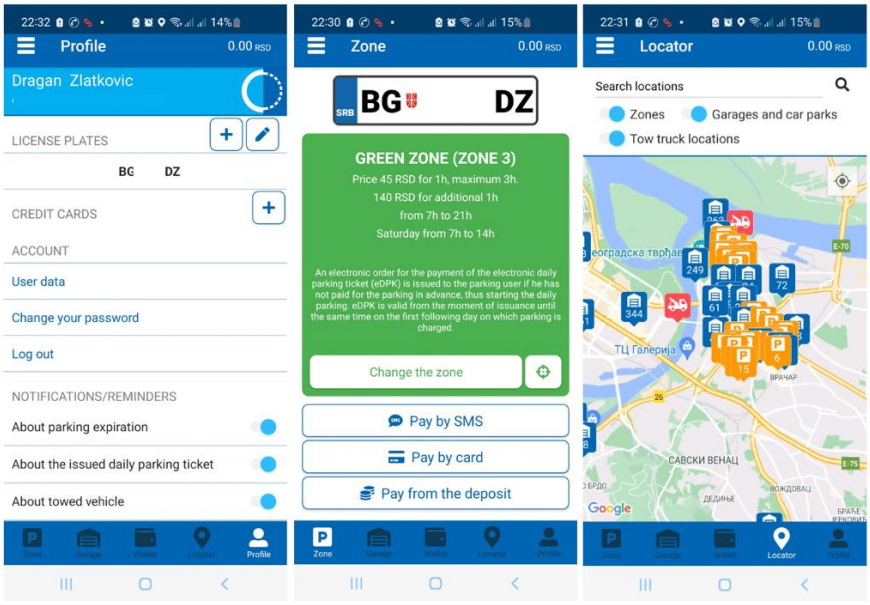


Figure 3: Belgrade Smart Parking App

The smart parking system has several benefits for drivers and the city as a whole. It reduces the time spent searching for a parking spot, which can reduce traffic congestion and air pollution. It also increases revenue for the city, as the parking fees can be set dynamically based on demand.

4.3. Electric vehicle charging infrastructure

Encouraging the adoption of electric vehicles can help to reduce air pollution in city of Belgrade. Smart charging infrastructure can be deployed throughout the city, making it easier for residents to charge their electric vehicles and reducing their reliance on gasoline-powered vehicles [18].

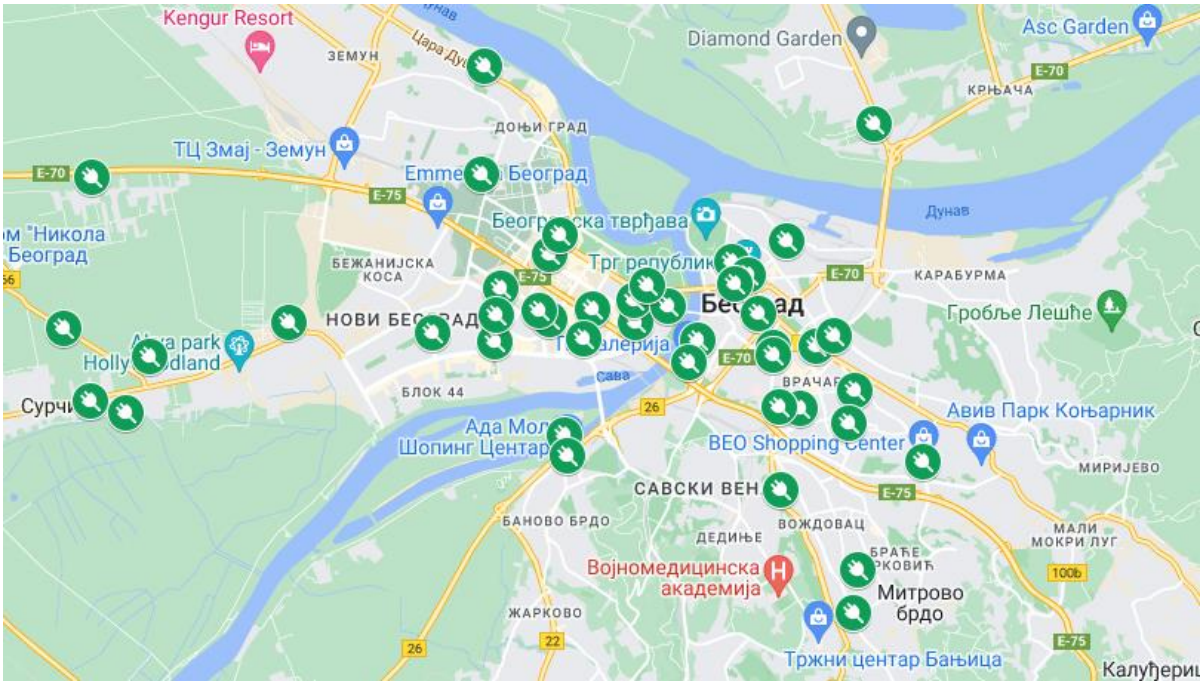


Figure 4: The network of the electric vehicle charging points in Belgrade (Source: eMobilnost)

Electric vehicle charging infrastructure refers to the network of charging stations and related equipment that enables electric vehicle owners to recharge their vehicles. This infrastructure can include a variety of charging options, such as level 1 and 2 charging stations for home or workplace charging, and level 3 DC fast charging stations for public use [19].

4.4. Green spaces

Green spaces such as parks and trees can help to absorb pollutants and improve air quality. Belgrade smart city planners can incorporate more green spaces into urban areas, utilizing data analytics to determine the best locations for these spaces [20].

Smart technologies can also be used to enhance the functionality and sustainability of green spaces in smart cities. For example, sensors can be installed to monitor soil moisture levels, temperature, and air quality, which can help optimize irrigation and maintenance schedules. Furthermore, digital tools such as mobile apps can help residents locate and access green spaces, report maintenance issues, and participate in community events.

4.5. Air pollution reduction technologies

There are various technologies that can be used to reduce air pollution, such as scrubbers on industrial smokestacks or filters on HVAC systems[1]. Smart city planners can identify areas with high levels of pollution and deploy these technologies to reduce emissions. There are several technologies available to reduce air pollution, from fuel-switching to exhaust gas treatment, low-emission vehicles, air purification systems, and alternative energy sources [21].

5. Citizen's Engagement in Smart City Development

Citizen engagement is critical for the success of smart city development. Engaging citizens in the process of developing smart cities can lead to more inclusive and responsive solutions that meet the needs and expectations of the community. Citizen's engagement is crucial for the success of smart city development. Creating open channels of communication, participatory planning, co-creation, education, and outreach, and monitoring and evaluation are some ways to encourage citizen engagement in smart city development [22].

6. Challenges

Belgrade, like many other cities around the world, faces several challenges in implementing smart city initiatives. As per our point of view and research some of these challenges include:

Funding: One of the biggest challenges is securing the necessary funding for smart city projects. Implementing smart city technologies requires significant investment in infrastructure, data analytics, and IoT devices.

Data Management: Gathering and managing large amounts of data from multiple sources can be a complex task, especially when different departments and organizations are involved. Ensuring data privacy and security is also a major challenge.

Citizen Engagement: Engaging citizens in smart city initiatives can be a challenge, as some may not be familiar with the technology or may not have access to it. It's important to ensure that all citizens have equal access to the benefits of smart city technology.

Infrastructure: Upgrading existing infrastructure to support smart city technologies can be a challenge. In Belgrade, for example, the city's aging infrastructure may require significant upgrades to support new technologies.

Coordination: Smart city initiatives often involve multiple stakeholders, including government agencies, private companies, and community groups. Ensuring coordination between these stakeholders can be a challenge.

Cultural Factors: Cultural factors such as resistance to change and bureaucratic structures may also hinder the implementation of smart city initiatives.

Smart City initiatives in Belgrade require a multidisciplinary approach and the involvement of different actors to be successful. So, addressing these challenges will require strong leadership, collaboration, and a commitment to innovation and sustainable development.

7. Conclusions

The development of Belgrade as a Smart City offers numerous opportunities for improving urban sustainability, efficiency, and citizen engagement. However, there are also several challenges that need

[1] HVAC stands for Heating, Ventilation, and Air Conditioning. HVAC systems are used to control and regulate the temperature, humidity, and air quality of indoor spaces.

to be addressed, including funding, collaboration, and privacy concerns. By addressing these challenges and building on the opportunities presented by Smart City initiatives, Belgrade can become a more sustainable, efficient, and livable city for its citizens.

Smart City initiatives also have the potential to improve citizen engagement and participation in urban governance. By implementing technologies such as e-governance platforms, smart citizen feedback systems, and social media, the city can engage citizens in decision-making processes and improve the transparency and accountability of local government [22], [23].

Privacy and security concerns are also important considerations in the development of Smart City initiatives. The collection and use of data by Smart City technologies raises questions about privacy and security, and the city will need to develop policies and regulations to protect the rights of its citizens.

The smart parking system in Belgrade is an excellent example of how technology can be used to improve urban life. It shows how smart city initiatives can provide practical solutions to everyday problems and improve the quality of life for citizens [9].

The smart city air pollution systems aim to improve air quality and reduce the negative health impacts of air pollution in urban areas. By utilizing advanced technologies and data analytics, these systems can provide real-time information and solutions to help address this critical issue [23].

In general, the smart city initiatives in Belgrade offer significant opportunities for improving energy efficiency and sustainability. By investing in smart technologies and systems, the city could reduce its carbon footprint, increase energy efficiency, and improve the quality of life for its citizens [2].

References:

[1] O. Gassmann, J. Böhm and M. Palmié, Smart Cities: Introducing Digital Innovation to Cities, Bingley: Emerald Publishing Limited, 2019.

[2] D. Zlatkovic and N. Denic, Smart cities: from urban development to digital infrastructure and cyber security, *Ecologica*, vol. 99, pp. 443-451, 2020.

[3] D. Munera, P. Tobon, J. Aguirre and G. N. Gomez, IoT-based air quality monitoring systems for smart cities: A systematic mapping study, *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 4, p. 3470~3482, 2021.

[4] J. Kisin, A. Okanović i S. Raletić-Jotanović, Pametan grad - budućnost održivog urbanog razvoja, *Ecologica*, t. 92, pp. 799-804, 2018.

[5] J. Woetzel, J. Remes, B. Boland, K. Lv, S. Sinha, G. Strube, J. Means, J. Law, A. Cadena and V. Taan, Smart Cities: Digital solutions for a more livable future, San Francisco: McKinsey Global Institute, 2018.

[6] D. Ilic, I. Milosevic and T. Ilic-Kosanovic, Application of Unmanned Aircraft Systems for smart city transformation: Case study Belgrade, *Technological Forecasting and Social Change*, vol. 176, p. 121487, 2022.

[7] M. Milosevic, D. Milosevic, D. Stevic and A. Stanojevic, Smart City: Modeling Key Indicators in Serbia Using IT2FS, *Sustainability*, vol. 11, no. 13, p. 3536, 2019.

[8] P. Neirotti, A. De Marco, C. A. Cagliano, G. Mangano and F. Scorrano, Current trends in Smart City initiatives: Some stylised facts, *Cities*, vol. 38, no. 1, pp. 25-36, 2014.

[9] D. Zlatkovic, V. Mikic, N. Denic and M. Ilic, Implementation of the Smart Parking System inthe city of Belgrade as part of Smart-City Subsystem, in *7th Conference with International Participation - Knowledge Management and Informatics*, Vrnjacka Banja, 2021.

[10] J. M. Friedrich, Air Pollution Is Greatest Environmental Threat to Health, *JAMA*, vol. 319, no. 11, p. 1085, 2018.

[11] N. Veljković and D. Lekić, Životna sredina u Srbiji 2004-2019, Ministartvo životne sredine - Agencija za zaštitu živote sredine, Beograd, 2019.

[12] Pametan Beograd, 25 May 2020. [Online]. Available: https://pametanbeograd.rs/sr/form/anketa1-saobracaj-. [Accessed 27 March 2023].

[13] L. Butler, T. Yigitcanlar and A. Paz, Smart Urban Mobility Innovations: A Comprehensive Review and Evaluation, *IEEE Access*, vol. 8, pp. 196034-196049, 2020.

[14] B. Ostojic, M. Vukovic and M. Bogdanovic, Concept of smart cities development strategy in the Western Balkans, *Ecologica*, vol. 105, pp. 88-98, 2022.

[15] R. Saborido and E. Alba, Software systems from smart city vendors, *Cities*, vol. 101, p. 102690, 2020.

[16] O. Popoola, D. Carruthers, C. Lad, B. Bright, M. Mead, M. Stettler, J. Saffell and R. Jones, Use of networks of low cost air quality sensors to quantify air quality in urban settings, *Atmospheric Environment*, vol. 194, pp. 58-70, 2018.

[17] J. Yang, J. Portilla and T. Riesgo, Smart parking service based on Wireless Sensor Networks, in *38th Annual Conference on IEEE Industrial Electronics Society*, Montreal, 2012.

[18] P. Morrissey, P. Weldon and M. O'Mahony, Future standard and fast charging infrastructure planning: An analysis of electric vehicle charging behaviour, *Energy Policy*, vol. 89, pp. 257-270, 2016.

[19] D. Yang, N. Sarma, M. Hylnd and R. Jayakrishnan, Dynamic modeling and real-time management of a system of EV fast-charging stations, *Transportation Research Part C: Emerging Technologies*, vol. 128, p. 103186, 2021.

[20] M. Kondo, J. Fluehr, T. McKeon and C. Branas, Urban Green Space and Its Impact on Human Health, *Int. J. Environ. Res. Public Health*, vol. 15, no. 3, p. 445, 2018.

[21] D. Vashi, P. Vashi, D. Shah and K. Kurumi, Air pollution and its control measures, *Advances in Applied Science Research*, vol. 6, no. 5, pp. 118-121, 2015.

[22] N. Goodman, A. Zwick, Z. Spicer and N. Carslen, Public engagement in smart city development: Lessons from communities in Canada's Smart City Challenge, *The Canadian Geographer*, vol. 64, no. 3, pp. 416-432, 2020.

[23] N. Bojković, T. Živojinović and N. Zornić, Air pollution in the beginning of the COVID-19 pandemic: analysis of emissions in Belgrade and Sarajevo, *Ecologica*, vol. 102, pp. 301-308, 2021.

Challenges of Knowledge Management in Industry 4.0 – Preliminary Literature Review

Jelena Slavić[1] , Zeljko Stojanov[1]

[1]*University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*

jelena.slavic@tfzr.rs; zeljko.stojanov@uns.ac.rs

Abstract:
Today, organizations are facing significant changes brought about by Industrial Revolution 4.0. With the development of industrial revolutions, knowledge management, one of the critical success factors, developed simultaneously. This paper aims to investigate and present the challenges of knowledge management in Industry 4.0. The research methodology was implemented through a preliminary literature review, where twelve primary studies were identified and research challenges were highlighted. The importance of the research is to summarize some of the critical problems faced by various industries worldwide. Knowledge management in Industry 4.0 ensures the use of vast amounts of data, effective sharing of knowledge, retention of knowledge within the organization and contribution to improving decision-making and innovation.

Keywords:
knowledge management, Industry 4.0, research challenges, literature review

1. Introduction

Industrial revolutions have introduced significant changes in business and manufacturing, as well as to human society. Manufacturing and business organizations should accept or adapt to them in order to be competitive in the ever-changing market. The period of industrial revolution 4.0 is the period of complete digitization and introducing completely new technologies. Development of Industry 4.0 (I4.0) is based on the four key components: cyberphysical systems (CPS), Internet of Things, Internet of Service and Smart factory. Among these four components, Skobolev et al. [1] lists six technologies used in I4.0: Industrial internet of things, additive production, Big Data, artificial intelligence, collaborative robots and virtual reality. I4.0 requires the combination of knowledge about manufacturing technologies, total quality and supply chain management, which demands effective knowledge management [2].

Industry 4.0 poses new challenges for future smart factories driven by four disruptions [3]:
1. increased volume of data, computing power and connectivity,
2. the emergence of analytical skills and business intelligence,
3. new forms of human-machine interaction,
4. improvements in the transfer of digital instructions to the physical world.

Industry 4.0 is the connection between digital technology, people and other physical systems and the integration of the digital and physical worlds through cyber-physical systems and the Internet of Things [4]. Industry 4.0 promotes new socio-technical infrastructures by transforming different aspects of a workplace such as health management and work organization, lifelong learning and career path models, team structures and knowledge management [5].

Knowledge management (KM) refers to the development of methods that promote the flow of knowledge between individuals, as well as identifying, processing and the use of this knowledge. KM also has as its primary interest the study of the contribution of information technology (IT) as a

mechanism to stimulate the creativity of individuals to develop new value for business [6]. Three key factors enabling knowledge management have been identified [7]: organizational structure, organizational culture and technology.

Knowledge management includes activities involving using, sharing and collecting knowledge within the organization, assuming that organizational learning influences knowledge management in manufacturing firms [8].

Technological changes have greatly affected knowledge management. Industry 4.0 has changed the way knowledge is developed in companies and constantly requires new managerial skills to facilitate learning [6]. According to the Abubakar et al. [7] knowledge process involves 4 factors: creation, retrieval/storage, transfer and application. Organizations must know the best ways of sharing knowledge in the organization. Some of the most successful ways of sharing knowledge are using modern technologies, artificial intelligence, information technologies and programs with large databases. In every organization, it is important to have experts from various fields needed for business and encouraging knowledge exchange activities. The technologies enabling the emergence of Industry 4.0 can simplify the exchange of information and knowledge between people at work [9]. People are the most important factor from the viewpoint of knowledge management in the Industry 4.0 paradigm and this new technological approach 4.0 requires changes in talent management practices and the nature of work skill-sets [10].

The aim of this paper is to identify the challenges related to knowledge management and technologies used in Industry 4.0. Based on various research conducted in the last few years in the countries of the world, primary studies were singled out and processed through a preliminary literature review.

2. Related work

This section will present some of the recently published literature reviews in the field of knowledge management in Industry 4.0, which is important for summarizing, categorizing, and challenging existing knowledge in the field of research [11]. Due to the recognized importance of literature reviews, several guidelines for conducting them exist [12,13,14].

Piccarozzi et al. [15] conducted systematic literature review on the topic of Industry 4.0 in management literature, aimed at discovering the gaps in literature and outlining future avenues of research. The most important topics discovered in the analyzed literature relate to production methods, business model, strategy, impact and consequences, and human resources. However, the authors did not identify knowledge management as an important issue in analyzed literature.

Through a systematic literature review Alkhazaleh et al. [16] inquired the most important factors affecting the success of technology transfer in I4.0 and reviewed existing models for technology transfer targeting I4.0. Based on the analyzed literature the authors proposed a conceptual framework of technology transfer for I4.0.

Ribeiro et al. [2] conducted a literature review to inquire how knowledge management supports Industry 4.0 implementation. By analyzing 27 empirical studies and 14 review papers, the authors identified three broader themes: technology (infrastructure), KM and learning (importance of both hard and soft skills, and enabler factors for KM), and worker engagement (communication and cultural aspects).

Dimensions regarding the relation between organizational learning and Industry 4.0 are identified through systematic literature review conducted by Belinski et al. [17]. The authors identified nine dimensions: management, Industry 4.0, general industry, technology, sustainability, application, interaction between industry and the academia, education and training and competency and skills. The

authors indicate the tight relation between knowledge management and organizational learning in Industry 4.0.

3. Literature review method

This work was done through a preliminary literature review, which was guided by the recommendations for conducting literature reviews in management [18,11] . The proposed process for systematic literature review was simplified for this study and contains the following steps, which are shown in Figure 1:

1. Setting the research question,
2. Defining search criteria and defining keywords,
3. Searching for studies,
4. Classification and selection of primary studies,
5. Analysis of selected primary studies.

Based on the proposed objective in the introduction section, the proposed reseearch question (RQ) is:

RQ: What are the main challenges for knowledge management in Industry 4.0?

For searching adequate literature, and based on the proposed research question, the following search string, with two keywords was formed and used:

"knowledge management" **AND** "Industry 4.0"

The literature search was done in Google Scholar, Science Direct, and Research Gate by using proposed search string.

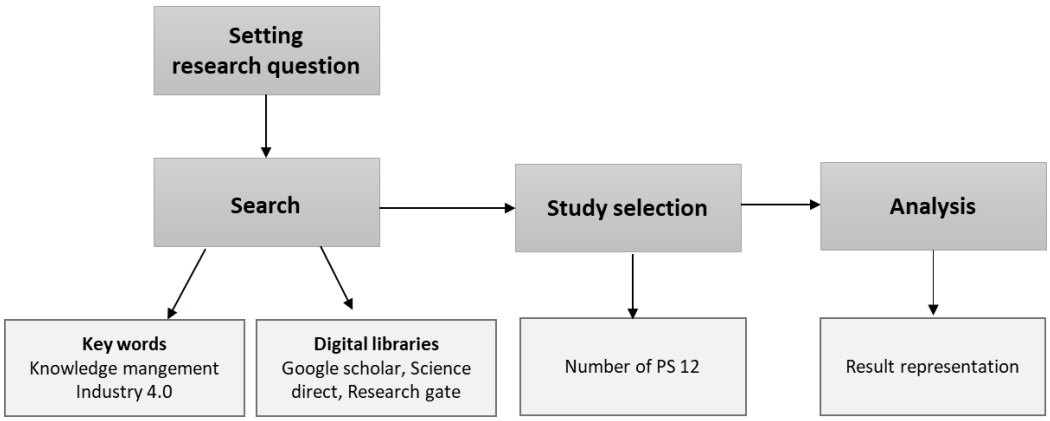


Figure 1. Method of preliminary systematic literature review

In order to classify the works, specific criteria were set that were followed during the search of primary studies. A period of time has been determined for the works taken into account to obtain the newest and the most actual research and the freshest results. In addition, they are firstly selected based on the title, abstract and keywords. For detailed checking of obtained studies, inclusion and exclusion criteria are proposed.

Inclusion critera:

I1. Paper published in refereed journal and conference.
I2. Paper is published between 2010 and 2023.
I3. Paper is available electronically.
I4. Paper directly address KM and I4.0 issues.

Exclusion criteria:

E1. Paper is not written in English.
E2. Paper has less than three pages.
E3. Paper is literature review.

Where this was insufficient to determine inclusion, we provisionally included publications for the classification phase. Final inclusion or exclusion was done based on reading the full text of the study.

After conducting the proposed steps for literature review, 12 primary studies (PS) were selected for detailed analysis (see Table 1).

Table 1:
List of primary studies

Primary Study	Reference
PS1	Lista, A. P., & Tortorella, G. L. (2022). Integration of Industry 4.0 technologies and Knowledge Management Systems for Operational Performance improvement. IFAC-PapersOnLine, 55(10), 2042-2047.
PS2	Khedr W. M, Gohar N. M. (2023). The Role of Knowledge Management in Adopting Industry 4 Technology: The Mediating Role of Market Orientation. 1-30.
PS3	Li, D., Fast-Berglund, Å., & Paulin, D. (2019). Current and future Industry 4.0 capabilities for information and knowledge sharing: Case of two Swedish SMEs. The International Journal of Advanced Manufacturing Technology, 105, 3951-3963.
PS4	Salvadorinho, J., & Teixeira, L. (2021). Organizational knowledge in the I4. 0 using BPMN: a case study. Procedia Computer Science, 181, 981-988.
PS5	Tortorella, G. L., Vergara, A. M. C., Garza-Reyes, J. A., & Sawhney, R. (2020). Organizational learning paths based upon industry 4.0 adoption: An empirical study with Brazilian manufacturers. International Journal of Production Economics, 219, 284-294.
PS6	Ghouri, A. M., & Mani, V. (2019). Role of real-time information-sharing through SaaS: An industry 4.0 perspective. International Journal of Information Management, 49, 301-315.
PS7	Eslami, M. H., Achtenhagen, L., Bertsch, C. T., & Lehmann, A. (2023). Knowledge-sharing across supply chain actors in adopting Industry 4.0 technologies: An exploratory case study within the automotive industry. Technological Forecasting and Social Change, 186, 122118.
PS8	Shafiei Nikabadi, M. (2014). A framework for technology-based factors for knowledge management in supply chain of auto industry. Vine, 44(3), 375-393.
PS9	Mao, H., Liu, S., Zhang, J., & Deng, Z. (2016). Information technology resource, knowledge management capability, and competitive advantage: The moderating role of resource commitment. International Journal of Information Management, 36(6), 1062-1074.
PS10	Brizolla, R. K., Patias, T. Z., & DORION, E. C. H. (2019). The understanding and the implementation of Industry 4.0: an exploratory Study of a Brazilian metal-mechanic SME. RAUnP-ISSN 1984-4204-Digital Object Identifier (DOI): http://dx. doi. org/10.21714/raunp., 11(2), 5-20.
PS11	Cimini, C., Boffelli, A., Lagorio, A., Kalchschmidt, M., & Pinto, R. (2020). How do industry 4.0 technologies influence organisational change?

	An empirical analysis of Italian SMEs. Journal of Manufacturing Technology Management, 32(3), 695-721.
PS12	Librita Arifiani, S. K., Dyah Budiastuti, M. M., & Wibowo Kosasih, E. (2019). The effect of disruption technology, and the future knowledge management toward service innovation for telecommunication industry 4.0 in Indonesia. Int. J. Eng. Adv. Technol, 8, 247-257.

4. Findings

The summary findings of the literature review are presented in a Table 2, which includes labels and descriptions of the identified challenges, as well as a list of studies in which the challenges appear.

Table 2:
Challenges for KM in I4.0

Label	Challenge title (description)	Primary studies
CH1	Impact of I4.0 technologies	PS1, PS5, PS8
CH2	KM processes influence the association between I4.0 technologies and operational performance improvement	PS1
CH3	Affect of KM activities on industry 4.0	PS2
CH4	KM in an unstable environment that characterizes the I4.0	PS4
CH5	Knowledge sharing in I4.0 environment	PS6, PS7, PS3
CH6	IT resources in KMC	PS9
CH7	Industry 4.0 concept and pillars	PS10
CH8	Industry 4.0 threats and opportunities	PS10
CH9	Adopting Industry 4.0 (I4.0) technologies	PS11
CH10	Impact of KM and big data on service innovation and competitive advantage	PS12

CH1 was considered in studies PS1, PS5 and PS8. In PS1, the focus is on the impact of I4.0 technologies adoption on KM practice regarding the 6 KM process, as PS5 researches the impact of I4.0 technologies adoption on organization learning development. Technology factors for knowledge management in the automotive industry for supply chain was the main aim of PS8 in CH1. CH2 in PS1 was based on the KM process's influence on the association between I4.0 technologies and operational performance improvement.

In PS2, the main CH3 is the affect of KM activities as knowledge creation, acquisition, storage, sharing, dissemination and application on Industry 4.0 in the manufacturing and service industry. CH4 show that BPM tool is essential for organizations and knowledge management in the environment of Industry 4.0.

In Industry 4.0, one of the main challenges is knowledge sharing. CH5 as CH1 includes the most primary studies PS6, PS7 and PS3. PS7 showed the crucial role played by knowledge-sharing in facilitating the adoption of I4.0 technologies. CH6 indicates the impact that IT resources have on knowledge management capability.

CH7 and CH8 deals with the challenges, opportunities, pillars and threats that Industry 4.0 has in small and medium enterprises in the metallurgic industry in Brazil. CH9 in PS11 shows the organizational implications of adopting Industry 4.0 technologies. CH10 was considered in PS12 that research impact of knowledge management and big data on service innovation and competitive advantage.

5. Conclusion

This paper points to the challenges of using KM practice in industries operating in the 4.0 environment. It can be concluded that one of the biggest challenges for industries operating around the world is the impact of I4.0 technologies and the sharing of knowledge in the I4.0 environment. Through the analysis of primary studies, it can be confirmed that knowledge management has a crucial impact on Industry 4.0.

During the implementation of this research, certain limitations were observed. The search was not conducted in the databases of all leading publishers of literature, which may affect the quality of the findings. The number of included studies is another possible limitation. Including a more significant number of primary studies can indicate a significantly larger number of challenges that industries in the 4.0 era face. In order to conduct better research in the future, it is necessary to include many more databases for literature search. Another future goal is to research specific aspects of using KM in I4.0 factories in Serbia through a field study.

References:

[1] P. O. Skobelev, S. Y. Borovik, On the way from Industry 4.0 to Industry 5.0: From digital manufacturing to digital society. Industry 4.0, 2(6) (2017) 307-311.
[2] V. B. Ribeiro, D. Nakano, Jr. Muniz, R. B. Oliveira, Knowledge management and Industry 4.0: a critical analysis and future agenda. Gestão & Produção, 29 e5222 (2022). doi: 10.1590/1806-9649-2022v29e5222.
[3] A. Wortmann, B. Combemale, O. Barais, A systematic mapping study on modeling for Industry 4.0, Project-Team DiverSE, 2017
[4] S. Khin, D. M. H. Kee, Factors influencing Industry 4.0 adoption. Journal of Manufacturing Technology Management, 33(3) (2022) 448-467.
[5] X. Xu, Y. Lu, B. Vogel-Heuser, L. Wang, Industry 4.0 and Industry 5.0—Inception, conception and perception. Journal of Manufacturing Systems, 61 (2021) 530-535.
[6] R. V. D. Gonzalez, M. F. Martins, Knowledge Management Process: a theoretical-conceptual research. Gestão & Produção, 24 (2017) 248-265.
[7] A. M. Abubakar, H. Elrehail, M. A. Alatailat, A. Elçi, Knowledge management, decision-making style and organizational performance. Journal of Innovation & Knowledge, 4(2) (2019) 104-114.
[8] A. Noruzy, V. M. Dalfard, B. Azhdari, S. Nazari-Shirkouhi, A. Rezazadeh, Relations between transformational leadership, organizational learning, knowledge management, organizational innovation, and organizational performance: an empirical investigation of manufacturing firms. The International Journal of Advanced Manufacturing Technology, 64 (2013) 1073-1085.
[9] D. Li, Å. Fast-Berglund, D. Paulin, Current and future Industry 4.0 capabilities for information and knowledge sharing: Case of two Swedish SMEs. The International Journal of Advanced Manufacturing Technology, 105 (2019) 3951-3963.
[10] Z. Whysall, M. Owtram, S. Brittain, The new talent management challenges of Industry 4.0. Journal of management development, 38(2) (2019) 118-129.
[11] C. Fisch, J. Block, Six tips for your (systematic) literature review in business and management research. Management Review Quarterly, 68 (2018) 103-106. doi: 10.1007/s11301-018-0142-x.
[12] J. Webster, R. T. Watson, Analyzing the past to prepare for the future: Writing a literature review. MIS quarterly, 26(2) (2002) xiii-xxiii.
[13] M. J. Grant, A. Booth, A typology of reviews: an analysis of 14 review types and associated methodologies. Health information & libraries journal, 26(2) (2009) 91-108.
[14] A. Booth, A. Sutton, D. Papaioannou, Systematic approaches to a successful literature review. SAGE Publications. London, UK. (2016)

[15] M. Piccarozzi, B. Aquilani, C. Gatti, Industry 4.0 in management studies: A systematic literature review. Sustainability, 10(10) (2018) 3821.
[16] R. Alkhazaleh, K. Mykoniatis, A. Alahmer, The success of technology transfer in the industry 4.0 era: A systematic literature review. Journal of Open Innovation: Technology, Market, and Complexity, 8(4) (2022) 202.
[17] R. Belinski, A. M. Peixe, G. F. Frederico, J. A. Garza-Reyes, Organizational learning and Industry 4.0: findings from a systematic literature review and research agenda. Benchmarking: An International Journal, 27(8) (2020) 2435-2457. doi: 10.1108/BIJ-04-2020-0158.

ChatGPT and AI for Learning – Opportunities and Challenges

Mimoza Anastoska-Jankulovska[1]

[1] *International Slavic University "St. Nicholas", Sveti Nikole-Bitola, North Macedonia*

jankmj2@yahoo.com

Abstract:
The goal of this paper is to assess the impact of Artificial Intelligence (AI) in learning. To achieve this goal. qualitative research through literature review is conducted. Initially, technology was used in education in the form of computers and computer related technologies, later transitioning to web-based and online intelligent education systems, and lately with web-based chatbots, that function with or without educators. To use these systems and platforms efficiently, educators and students need to be trained in prompt engineering – asking questions/queries or giving requests to generative AI. After receiving answers from the AI, educators and students need to be able to verify them in order to use them in the process of learning. The research resulted in short guidelines for creating good prompts, which will be explained more in depth later in this paper. Additionally, this paper identifies advantages that education can benefit from and disadvantages that should be considered while using AI. A lot of AI tools and platforms are listed in this paper that can be used for educational purposes. Being a very new field, there are still a lot of topics about AI in the educational process that should be studied and researched in the future.

Keywords:
Generative Artificial Intelligence, Machine Learning, ChatGPT, Education, Learning, Prompt engineering

1. Introduction

Artificial intelligence (AI) uses Machine Learning (ML). ML makes use of statistical models to develop predictions. More specifically, ML uses algorithms that take empirical or historical data in, analyze it, and generate outputs based on that analysis. However, machine learning exists for a long time already. For example, the tool of Autocorrect in text processors uses a big number of data to predict what we intended to write and to correct it into that assumption. When we search online, Google search suggests finalization of our queries through the use of ML. We receive suggested posts in social media, YouTube, Spotify again with the help of AI and ML.

AI, according to britannica.com, is the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. "AI is a loose umbrella term that refers to a collection of methods, capabilities, and limitations—many of which are often not explicitly articulated by researchers, education technology companies, or other AI developers." [1] ChatGPT has become the most widely known AI. It was released in November 2022 by OpenAI Foundation and is claimed to be a free tool for everyone who will register. GPT stands for Generative Pre-trained Transformer. **Generative** in linguistics involves applying a finite set of rules in order to produce the well-formed items of a language. **Pre-trained** means that it was developed (trained) by using a big amount of data and, generally, is not using newer data. The data used to train ChatGPT included everything that was online on the Internet until 2021. A large language model (LLM) is constructed with artificial neural networks of large size. They are used for AI to be trained. **Transformer** is a deep learning model that is weighting the significance of each part of the input data, and thus creating an output. ChatGPT is a chatbot, this means a robot that communicates with users through chats. ChatGPT is online but is not able to browse through the Internet and has limited knowledge of the world and events after 2021.

Generative AI goes beyond performing specific tasks based on predefined rules and patterns and strives to create entirely new data that resembles human-created content. This paper deals with generative AI.

Generative AI provides reliable and accurate answers to many queries. But it can also encourage prejudices from the texts it analyses and, in some cases, can provide false or misleading information, even supported by non-existent sources. Sometimes, it will admit its limitations, for example when asked to make predictions, to summarize texts or to make personal reflections, but many authors had experiences in it giving wrong answers. It will be dangerous if we completely rely on artificial intelligence and assume that it knows everything. And that is the most common case, especially among students and users who do not have much experience with technology.

"Too often we only ask how a new technology will change education. A more interesting question is: How will education shape our reception and steer the integration of new technology – both technology that is here today and technology that remains on the horizon? Our education systems can define a trajectory and establish norms for how we understand world-changing technology – and, by extension, how we allow it to influence us and our world." [2]

2. Advantages while using ChatGPT and other AI tools

ChatGPT and other AI tools can be very useful for a lot of tasks. Their implementation opens a lot of new opportunities in research and education. Some of the advantages that can be used are:

ChatGPT can create content very quickly

ChatGPT is developed and trained on a large database and is designed to generate responses as an answer to a question or a query. Users can ask a question such as: "What are Newton's three laws?" or "What are the downsides of global warming" and ChatGPT will generate a written, concise, and generally accurate answer. It can also offer its own analysis and it can do all this in different languages.

Moreover, ChatGPT can generate code that can be used in software programs and can code in various programming languages.

Since ChatGPT is not directly connected to the Internet and has limited knowledge of the world and events after 2021, it will not be able to generate specific answers to questions that require more recent information. For example, asking the system to answer, "What was the average temperature in Bitola in August 2023?" will elicit a response that since it cannot browse the Internet, ChatGPT can only provide general information about the weather in Bitola.

ChatGPT 3.5 is currently free

Since ChatGPT was developed and distributed by OpenAI, the announcement is that there will always be a free-to-use service. Users only need to be registered in the system of the respective AI. Another service from the OpenAI Foundation is an AI-enabled DALL-E image generator that requires a fee. There is a version ChatGPT 4.0, which one must pay to use.

ChatGPT learns from its interactions with users

This can be done through "tuning", a process in which ChatGPT acquires a new database of topic-specific conversations or the area in which the conversation takes place. This allows ChatGPT to learn more about the specific language and content so that it can generate more relevant responses in the future.

One of its main advantages is that it can be easily adapted for specific tasks or domains, allowing it to generate relevant responses to a particular conversation or user need. It can also handle a wide variety of conversation styles and can generate responses that are appropriate for different types of conversations, including casual, formal, and technical.

It should be noted that ChatGPT cannot learn in the same way as a human, but it can improve its performance to generate more appropriate responses by processing new data from user conversations.

ChatGPT is only one of the various AI tools available

There are a growing number of artificial intelligences similar to ChatGPT that can assist in the writing process, be it individual or collaborative writing. There are artificial intelligences that can create images, audio, or video materials according to given instructions. Some of these artificial intelligences are free and some require a monthly fee to be used.

3. Disadvantages of using ChatGPT and other AI tools

Having presented all the strengths, we should also be aware of the disadvantages of ChatGPT and other AI tools. Some of them are:

Cannot browse online looking for more recent information and data

ChatGPT is a standalone AI chatbot that is trained at the beginning and does not have access to the latest information and data which is widely available online.

Dependence on large amounts of data

ChatGPT, like other language models, needs large amounts of data to learn patterns and generate responses. This can be an advantage in some cases. On the other hand, in areas where limited data is available, ChatGPT may not perform well.

ChatGPT does not always provide correct answers

There is no guarantee that the information provided by ChatGPT will be accurate or complete, just like the results obtained from a search on Google or any other search engine. This can lead to students receiving inaccurate or misleading information, which can slow down their learning and understanding.

If generative AI does not have the requested information, instead of just giving a response something like "Error" or "Please try again", ChatGPT actually makes up something. Because ChatGPT sometimes provides misleading information, some websites have banned responses generated by ChatGPT (e.g., StackOverflow). On top of this, some countries have completely banned the use of ChatGPT.

Sometimes, if the same query/prompt is repeated to ChatGPT, it can give different answer. This is not always desirable. Users should be equipped with knowledge and possibilities to check the answers generated by ChatGPT.

Reproduce biases and stereotypes that exist

ChatGPT is trained using a huge amount of data from the Internet. However, there is a possibility that this data contained prejudices and stereotypes that are widespread in society. Because of this, ChatGPT may generate responses that are offensive, discriminatory, or harmful to certain groups of people. For example, if a student asks a question on ChatGPT about a specific group of people, the model can generate an answer that is based on stereotypes or prejudices, due to pre-existing Internet stereotypes and prejudices about that particular group.

Lack of empathy and human touch

ChatGPT is unable to understand or respond to emotions the same way humans do. It cannot always generate appropriate responses in emotionally complicated situations. In the previous example about generated answer with stereotypes or prejudices, ChatGTP cannot understand its own mistake and is not able to correct it.

ChatGPT and other AI can generate responses that are appropriate for different types of conversations. However, all AI systems, if it is not specifically asked from them, cannot personalize their responses to individual users in the same way as humans do.

When all students receive the same information and feedback based on model results, it can limit their ability to develop their own unique perspectives and voices, and to benefit from personalized instruction or guidance from human educators.

Lack of common sense

Language models like ChatGPT do not have the ability to understand or use common sense knowledge in the same way that humans have. This can lead to generated results that are not appropriate or relevant to the conversation. For example, on the question "What would happen if the Earth stop spinning?", ChatGPT will give some changes like losing day/night and so on. It finishes with the statement that humans can easily adapt to these changes and if Earth stops spinning, it will not cause any major problems for humans and humanity. To any human who is reading this, the answer simply does not seem reasonable.

4. ChatGPT and AI in Learning

"There is no single or fixed definition of AI, but there is common agreement that machines based on AI are potentially capable of imitating or even exceeding human cognitive capacities, including sensing, language interaction, reasoning and analysis, problem solving, and even creativity." [3] It is obvious that there can be numerous benefits in including AI in learning. The benefits can be quicker access to more information, adaptable materials and curriculum, personalized answers, help in administration of education, better inclusion of vulnerable groups, better access to different opportunities, etc. "AI in education has also eliminated some barriers to access to learning opportunities, such as national and international borders, enabling global access to learning through online and web-based platforms" [4].

Generative AI can help teachers and researchers generate useful text and other outputs to support their work. However, it is not always a straight-forward process. "... the community should focus ... on support the role of teachers, ..." [5] It may take multiple repetitions of the request before the desired result is achieved. Students, being less experienced than teachers, may unknowingly and without critical reflection accept the answer provided by generative AI even though it may be superficial, inaccurate or even harmful.

That is why it is very important to know how:

1. To ask adequate prompts that will guide AI to generate desired output. This is the reason why a completely new field has emerged – Prompt Engineering.
2. To reflect critically on the received output generated by AI, and to research other resources in order to verify received generated output.

The education should strive to prepare teachers and students to be able to perform the above two steps while working with generative AI.

4.1. Prompt engineering for learning

Users can converse with generative AI by asking questions, queries, or prompts. Although all three terms can be found in the literature and are equally used in this paper, the term "prompt" starts to emerge as the most used in these conversations. A completely new field has been developed – prompt engineering – which helps in preparing prompts in the way so that generative AI generates the answer that is closest to the user's idea/need and that is the most accurate one.

"Prompt-engineering refers to the processes and techniques for composing input to produce Generative AI output that more closely resembles the user's desired intent." [6] It is essential to be able to formulate clear, ethical, and responsible prompts when interacting with generative AI systems to ensure productive and responsible AI usage.

In this short period in conversation with generative AI, the following guidelines appeared to be very helpful in formulating the prompt to AI:

1. State the purpose and focus of your chat

Good example: "Prepare a text about black holes for 15 years old students of Physics"

Bad example: "Prepare text about black holes in 2 pages"

2. Be Specific as much as possible

"It's important to use specific and clear instructions when asking for information, as the model is not able to infer the context or meaning behind vague or ambiguous requests." [7]

Good examples: "Suggest to me a recipe with eggs, tortillas, ham and cheese".
"Create a logo for a start-up working on education about AI. The start-up creators are 3 people in their 20s. The target group for education are people older than 50 years."

Bad example: "How to bake a cake?"

3. Provide context and examples if possible

Good example: "I want to travel to Macedonia. I have 6 days and want to see at least 3 cities. I don't want to spend a lot of time traveling. Suggest an itinerary to me. For example, I liked when last year I visited Poland and went to Krakow, Katowice and Warsaw."

Bad example: "I want to travel to Macedonia. I have 6 days. Suggest an itinerary to me."

4. State your desired outcome

Good example: "Write 5 multiple choice questions for a quiz for 15 years old students of Physics on the topic of black holes."

Bad example: "Prepare questions for a quiz on black holes."

Generally, the following elements are considered as parts of a good prompt:

- Instructions
- Purpose
- Context
- Examples
- Desired characteristic of the output
- Clearness and specificity

4.2. Different AI tools and platforms for learning

There are a growing number of LLMs that are similar to ChatGPT and can support humans in their work. Some of them are focused on writing, others on creating music, visuals, audio files, etc. New AI tools and platforms are emerging very quickly. It is good to know at least some of them in order to be able to use them in own work. The list below, adapted from [6] and own research, presents different tools and platforms that are operational in September 2023. Most of these are free to use (within certain limits), while some are open-source.

AI tools that help in **generating text** are:

- **Ecree** (https://www.ecree.com/about) – provides widespread access to fast, easy and personalized writing feedback that can help every student and every teacher.
- **TooWrite** (https://www.tswrevolution.com/) – a scientific writing tool for researchers.
- **Writefull** (https://www.writefull.com/) – Writefull's AI helps you write, paraphrase, edit texts.
- **CoAuthor** (https://coauthor.stanford.edu/) – a human-AI collaborative writing dataset that captures interaction between 63 writers and four instances of GPT-3 in English.
- **Alpaca** (https://crfm.stanford.edu/2023/03/13/alpaca.html) – a fine-tuned version of Meta's Llama, from Stanford University, which aims to address LLMs' false information, social stereotypes and toxic language.
- **Bard** (https://bard.google.com) – a LLM from Google, based on its LaMDA and PaLM 2 systems, that has access to the internet in real time, which means it can provide up-to-date information.

- **Chatsonic** (https://writesonic.com/chat) – a revolutionary AI like Chat GPT - Chatsonic (now with GPT-4 capabilities), the conversational AI NLP processor that addresses the limitations of ChatGPT.
- **Hugging Chat** (https://huggingface.co) – made by HuggingFace, who emphasized ethics and transparency throughout its development, training and deployment. In addition, all data used to train their models are open source.
- **YouChat** (https://you.com) - A LLM that incorporates real-time search capabilities to provide additional context and insights in order to generate more accurate and reliable results.

AI tools that are helping with **colaborative writing** are:

- **CoWrite** (https://cowrite.com/en/) - GPT-4 powered writing platform, easy and accurate and has been proven to increase the efficiency of customer's writing with 45%.
- **ParagraphAI** (https://paragraphai.com/) - Boost your productivity and professionalism with this free AI writing assistant.
- **Compose AI** (https://www.compose.ai/) - Compose AI is a Chrome extension that cuts your writing time by 40% with AI-powered autocompletion & text generation.

AI tools that help with **generating images** from text prompts are:

- **Craiyon** (https://www.craiyon.com) – (former DALL•E mini) Creates AI Art with our free AI image generator.
- **DALL•E 2** (https://openai.com/product/dall-e-2) - OpenAI's image Generative AI tool.
- **DreamStudio** (https://dream.ai/create) - Stable Diffusion's image Generative AI tool.
- **Fotor** (https://www.fotor.com/) - Incorporates Generative AI in a range of image-editing tools.
- **Midjourney** (https://www.midjourney.com) - An independent image Generative AI tool.
- **NightCafe** (https://creator.nightcafe.studio) - Interface to Stable Diffusion and DALL•E 2.
- **Photosonic** (https://writesonic.com/photosonic-ai-art-generator) - WriteSonic's AI art generator.

Examples of **video generating AI** include the following:

- **Elai** (https://elai.io) - Converts presentations, websites and text into videos.
- **GliaCloud** (https://www.gliacloud.com) - Generates videos from news content, social media posts, live sporting events and statistical data.
- **Pictory** (https://pictory.ai) - Automatically creates short videos from long-form content.
- **Runway** (https://runwayml.com) - Offers a range of video (and imaging) generation and editing tools.

Music generating AI are:

- **Aiva** (https://www.aiva.ai) - Automatically creates personalized soundtracks.
- **Boomy** (https://boomy.com) - Creates original songs in seconds, even if you've never made music before.
- **Soundraw** (https://soundraw.io) - Creates beats with the power of AI
- **Voicemod** (https://www.voicemod.net/text-to-song) - Generates songs from any text, and require no music composition knowledge.

Some other AI tools with **miscellaneous characteristics** are:

- **Knewton** (https://www.knewton.com/) - Tailors educational content to the needs of individual student, allowing for a more personalized learning experience that helps students achieve mastery of learning outcomes
- **ChatPDF** (https://www.chatpdf.com) - Summarizes and answers questions about submitted PDF documents.
- **Elicit: The AI Research Assistant** (https://elicit.org) - Aims to automate parts of researchers' workflows, identifying relevant papers and summarizing key information.

- **Perplexity** (https://www.perplexity.ai) - Provides a 'knowledge hub' for people seeking quick answers tailored to their needs.
- **WebChatGPT** (https://tools.zmo.ai/webchatgpt) - Gives ChatGPT internet access to enable more accurate and up-to-date conversations.
- **TeamSmart AI** (https://www.teamsmart.ai) - Provides a 'team of virtual assistants.
- **Wiseone** (https://wiseone.io) - Get reliable information everywhere, explore the web while reading and receive instant answers to your most complex questions.

4.3. Other measures to be employed in education

"... it's important to understand the capabilities and limitations of the technology (AI and ChatGPT), as well as the ethical considerations that come with its use." [7]

One of the biggest worries in academia is that ChatGPT will be used for cheating and plagiarism. To avoid this, there are various steps that educators and educational institutions can take. Some of the actions can be:

- Include artificial intelligence technology in academic settings as part of the curriculum to prepare students better in responsible use of technology and AI.
- Educate students and educators in ways to reflect critically on the received output generated by AI, and to research other resources in order to validate received output.
- Train educators to use ChatGPT in ways that support learning and academic achievement instead of a replacement for traditional forms of teaching, learning and assessment.
- Develop guidelines and policies for the use of generative AI (including ChatGPT) in academic work and ensure that students and educators are aware of them and follow these guidelines.
- Monitor the use of generative AI in academic settings and take appropriate action if it is used for cheating or other unethical purposes.

Education research is another field where AI can have significant input and employment. "... (AI) can significantly reduce the time and effort required for data analysis, enabling researchers to focus on higher-level tasks, such as interpreting the results and formulating actionable insights because of their emergent abilities." [8]

As with any other change, we need to be prepared to be able to use efficiently the opportunities that arise. "Used ethically and with due consideration of the need to build individual and institutional capacity, ChatGPT could support HEIs to provide students with a more personalized and relevant learning experience, make administrative processes more efficient, and advance research and community engagement." [9]

5. Discussion and Conclusions

The release of numerous AI tools and platforms has put the whole world in shock. There are a lot of generative AI tools emerging every day which are not giving education a chance to test and validate their possibilities. Instead, the educational institutions and educators should focus on fostering a comprehensive understanding of when, by whom, and in which way this novel technology should and should not be employed. AI prompts to reassess the education practices, how they are done, when and why.

"While the current state of generative AI technology represented by ChatGPT is impressive but flawed, it is only a preview of what is to come. It is important for engineering educators to understand the implications of this technology and study how to adapt the engineering education ecosystem to ensure that the next generation of engineers can take advantage of the benefits offered by generative AI while minimizing any negative consequences." [10]

By employing generative AI, and more specifically ChatGPT, in learning and education, the systems will be preparing students to grow into productive citizens of tomorrow. "The use of simulation, virtual

reality and other aspects of AI in learning was shown to prepare students for futuristic trends with the gradual gravitation towards keeping pace with the application of AI in industry." [3] AI is here to stay, and the world needs to learn how to use it for own benefit and prosperity.

Between the first release of ChatGPT in November 2022 and today, a short period has elapsed. However, there are a number of research and studies conducted. Educators and students are getting new experiences with every passing hour. "... the most appropriate ways to deploy AI for particular (assessment) activities and in specific contexts remains an area of debate" [1] and future research. In addition, future research should focus on what results have been accomplished by using curriculum and materials adapted with the help of generative AI, lessons learnt from prompt engineering, what kind of regulation about using AI are in place, and how others should be adapted.

References:

[1] N. Gillani at all, Unpacking the "Black Box" of AI in Education, Educational Technology & Society, 26(1), 99-111., 2023
[2] S. Giannini, Reflections on generative AI and the future of education, 2023
[3] UNESCO and World Commission on the Ethics of Scientific Knowledge and Technology (COMEST), Preliminary Study on the Ethics of Artificial Intelligence, 2019
[4] L. Chen, P. Chen and Z. Lin, Artificial Intelligence in Education: A Review, in IEEE Access, vol. 8, pp. 75264-75278, (2020), doi: 10.1109/ACCESS.2020.2988510
[5] S. Bulathwela at all, Could AI Democratise Education? Socio-Technical Imaginaries of an EdTech Revolution, Conference on Neural Information Processing Systems, 2021
[6] UNESCO, Guidance for generative AI in education and research, 2023
[7] S. Atlas, ChatGPT for Higher Education and Professional Development: A Guide to Conversational AI, University of Rhode Island, College of Business, 2023
[8] A. Katz at all, The Utility of Large Language Models and Generative AI for Education Research, Department of Engineering Education, USA, 2023
[9] UNESCO, ChatGPT and Artificial Intelligence in higher education: Quick start guide, 2023
[10] J. Qadir, Engineering Education in the Era of ChatGPT: Promise and Pitfalls of Generative AI for Education, IEEE Global Engineering Education Conference (EDUCON), Kuwait, Kuwait, (2023), pp. 1-9, doi: 10.1109/EDUCON54358.2023.10125121.

AI and Tracking Data Exchanges on Maps

Marko Blažić[1,2], Dubravka Sladić[1], Višnja Ognjenović[2], Ivana Berković[2] and Katarina Vignjević[2]

[1] *University of Novi Sad, Faculty of Technical Science, Novi Sad*
[2] *University of Novi Sad, Technical Faculty "Mihailo Pupin", Zrenjanin*

markoblazic93@mail.com;dudab@uns.ac.rs;ivana.berkovic62@gmail.com;visnjaognjenovic@gmail.com;katarina.vignjevic@tfzr.rs

Abstract:
In the year 2023, artificial intelligence (AI) reached a significant milestone as it became a mainstream phenomenon. AI has been in development for a decade, and numerous companies have adopted AI technologies across various sectors. Notably, user-friendly AI models like ChatGPT gained popularity, marking AI's enduring presence. The competition between industry giants Microsoft and Google in the AI arena further emphasized its importance. The influence of AI extended to the field of cartography, where creating accurate and ethical maps with minimal misinformation was paramount. AI tools offer a range of capabilities, including map style transfer, retrieval, generalization, and design critique. However, the opaque nature of AI raised concerns about the level of trust that could be placed in machine-generated maps. One noteworthy AI model for generating maps was DALL-E 2, developed by OpenAI. DALL-E 2 relied on prompts to generate diverse and realistic images, allowing users to select the mapped region, map location on the image, and map type. Google Maps also harnessed AI, using it to predict traffic patterns, collect real-time data, and collaborate with DeepMind for advanced road analysis. AI's impact extended to indoor mapping, retail maps, and airport maps, utilizing various technologies such as proximity, trilateration, fingerprinting, and motion positioning. AI-powered robots played a role in indoor mapping, addressing challenges related to localization through metric or topological approaches.

Keywords:
AI, Google Maps, Indoor maps, in-store maps, AI-generated maps, Airport maps

1. Introduction

The year 2023 can be considered as the year AI went mainstream. This concept has been present for a decade and many companies have used this concept in their business in many varieties. Rise of the consumer-friendly AI, such as ChatGPT, and the "battle" between two leaders such as Microsoft and Google, are proof that AI is here to stay in years to come [1].

Cartographers have recognized the significance of making cards that are trustworthy and ethical, cards that have truthful information about region with as little as possible misinformation [2]. With the rapid development in AI, use of AI in map development presents opportunities and concerns [3]. For example, there are many tools that cartographers can use in design decision on things such as:

- map style transfer [4],
- map retrieval [5],
- map generalization [6],
- map design critique [7]

Despite all this quality characteristic, the AI and machine learning as a concept is considered as "black-box", so the question that is asked by cartographers is "how much should be trusted in machine-generated maps?" [8].

DALL-E 2 is one of the possibilities for generating maps using AI. It is an advanced model developed by OpenAI that depends on prompt for generating images. This model is trained on pairs of images, allowing us to generate realistic and diverse images based on those prompts. DALL-E allows us to choose mapped region, the location of the map on the image and map type [9].

- Mapped region – based on the possibility that maps can vary based on scale that they are presented, maps can be generated in 3 scale: state, country, and continent.
- Map type – this parameter includes six categories: general map, heat map, political map, physical map, reference map and choropleth map.

The first two parameters are required, the third parameter is optional, and it presents the location of the map in the image. That could be table, desk, field rather than map be entire image.

2. Google Maps and AI

There are couple of things that should be pointed out about Google Maps and AI in combination [10]:

- Google Maps is using Machine Learning to suggest roads with low trafficking - To predict traffic patterns in the near future, Google Maps analyzes historical data of traffic on the same route over time. For example, if we take North California and freeway 280, the data tells us that around 6-7am vehicles are traveling around 110km/h but only 50 km/h in the afternoon. Then maps are using this data in combination with traffic conditions to give us a prediction [11].
- Google collects location from all users in area to give live traffic details – Google traffic collects all data from users who turned on location on, in google maps app. Google collects that data and shares with other users on that specific location through app [12].
- Google Maps uses historic traffic patterns of the road to present probably traffic condition at the same time on the same road- To get this information google maps requires two types of data: data from history about same route, at the same time and on the same day. The second data comes from real-time data sent by sensors and smartphones reporting how fast cars are moving [13].
- Google Maps and DeepMind together will use advanced Machine learning for analyzing roads, quality, speed limits and closures. With the newest information Google Maps will give a better estimated time before the trip is even started [14].

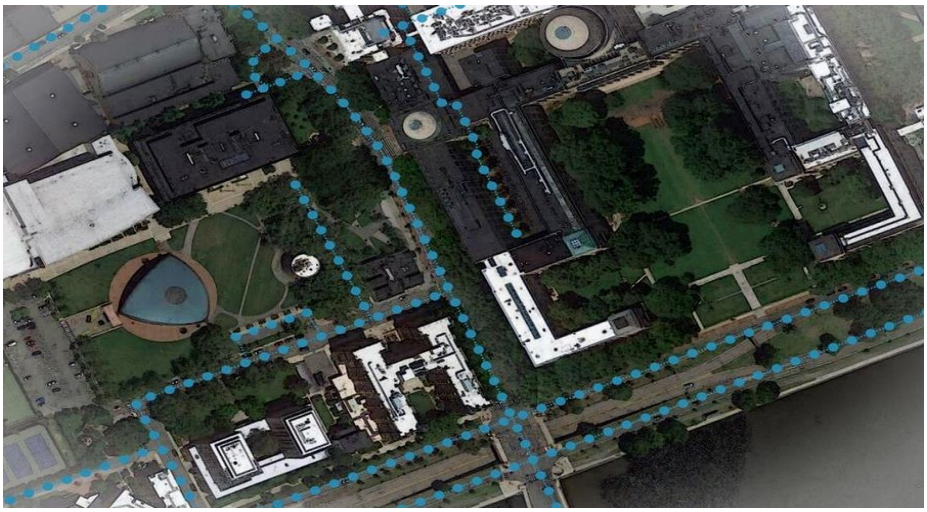


Figure 1: Google map powered by AI [15]

Dijkstra algorithm is commonly used for finding the shortest path between two nodes in a communication network [16]. He is the most popular algorithm and widely used in real-time applications, because of his complexity. Additionally, he is used for optimal path planning and for identifying collision-free paths [17]. Modification of this algorithm is used in calculations for planning shortest path for vehicles [18].

3. Indoor Maps and AI

Indoor mapping would mean applying numerous technologies to determine devices and their location indoors. While indoor mapping has some similarities with GPS, it does not rely on satellite tracking [19].

Indoor mapping includes [19]:

- Proximity positioning – represents distance between receiver and device.
- Trilateration positioning – It is used to determine the position of an object with several reference points. Signal strength is the parameter for distance between two objects.
- Fingerprinting positioning – represents pinpointing objects across venues, and when a new object becomes available on the network, their location is mapped according to specific coordinates already collected.
- Motion positioning - Motion positioning can be determined by smartphones. There are three components needed:
    1. Accelerometers – measuring acceleration and they are equipped with axis-based sensors.
    2. Gyroscope – detecting orientation of a device. For example, rotating smartphone from portrait to landscape requires gyroscope.
    3. Magnetometers – determining position of smartphone in physical space.

Figure 2: Indoor map powered by AI [20]

To build an indoor map, we can also use robots powered by AI. If we want for robot to build a map of its environment, robots must know where it is. Since the motion of the robot is not precise, robots must solve localization problems. These problems can be solved by one of two major paradigms [21]:

- Metric – In this representation robot's environment is defined by single global coordinate system, in which all mapping and navigation takes place. Each cell of the map presents the same amount space in the real world. This grid of cells is pretty convenient at representing the structure of the world [22].
- Topological – In this representation robot's environment is description of environment-based graph, where nodes are connects to places or landmarks, and arcs connects to action that connect neighbor places [23]. In short terms, topological maps are quality representation of robot environment, in which is presented as places connected with other places and connection between them.

4. In-Store Maps and AI

An important fact to point out is that in-store maps are not physical maps, they are digital maps that visualize operations, objects and people. Both customers and managers have benefits from this type of maps, because they bind marketing experience, events and product together [24].

In-store maps point customers to the desired product, and how to get to them. One research done by Google/Ipsos shows that 63% percent of people would rather buy product that is presented online through web sites and mobile apps. In-store mapping is not only for blending online and offline insights to create a 360-degree experience [24].

The retail mapping software are given based on the tags that pinpoint the location of the user. Users just need to use an indoor store app, to be navigated to the desired product with the shortest route possible. This platform can also provide information about discounts, promotion etc. For example, an app can warn users about products that have been sold out, or products on sale nearby. One research has shown that purchase done this way, last minute offer, was done in 78% [25].

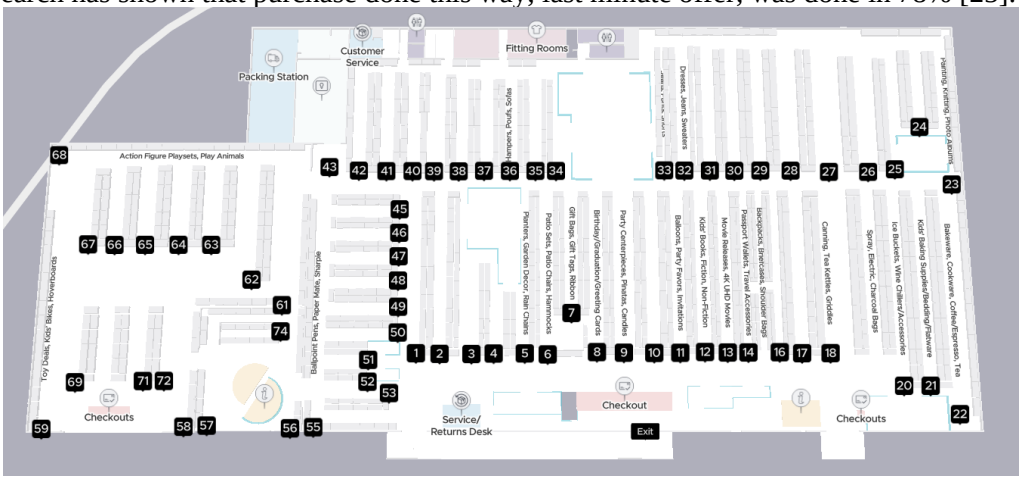


Figure 3: Example of in-store mapping [26]

5. Airport Maps and AI

One of the possibilities for AI are airports. If it has been taken into consideration their complexity and the fact that they are very crowded, it is necessary to develop airport maps in order to make traffic as fluent as possible. The solution that comes in handy is AI based Airport Maps.

For starters, airports are composed of a couple specific features such as: aprons, terminals, taxiways and runways that can be easily seen from OSM (open street maps). Additional features such as helipads or hangars can be showed as rectangles, polygons or lines. In order to generate maps correctly, airports are modelled as complex of these objects [27]. There are couple of algorithms that can generate these features: taxiway collapse and typification, runway collapse, apron and terminal amalgamation and algorithm change whether the taxiways/runways are lines or polygons. Version of airport with this guideline is presented for OSM airports into ScaleMaster [28].

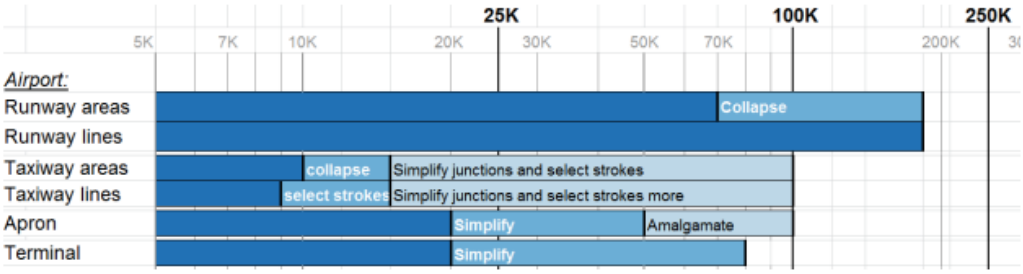


Figure 4: ScaleMaster modelling of airports using guidelines [28]

It will be mentioned algorithm that can help with generalization while scale level decrease.

Runway Collapse is a simple yet efficient method for simplifying runway representations. It involves finding the longest straight segment within a runway polygon that matches its orientation. This method is less intricate than certain other algorithms and is adept at generating valuable results for runway mapping in GIS or mapping applications [29].

6. Tracking data exchange on Maps

Based on time series data, it is possible to display time series maps [30]. Figure 5 shows only one map based on time series where data from different time series are marked with a different color [31].

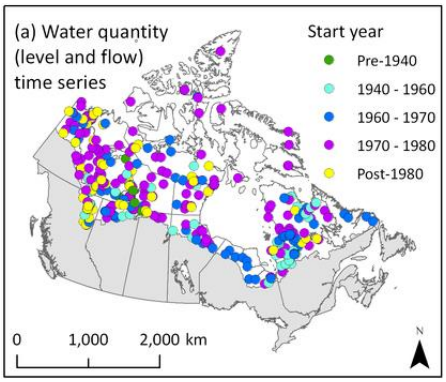


Figure 5: Tracking data exchange on maps [31]

There is also another way of displaying time series by showing an individual map for each time series. While the first way is cluttered because the elements of one series are over the elements of another series, there is a series of maps corresponding to the thesis time series to follow.

Based on AI algorithms, it is possible to monitor only larger oscillations among time series data. The authors of this paper came up with the idea that the maps could only show data changes and that to the extent of the data difference of two time series. One color could show the difference in increase and another color the difference in decrease. It was noted that this idea has not been used and that it would be useful to test it.

7. Conclusions

In conclusion, the year 2023 marks a significant milestone for AI, with its integration into mainstream applications such as mapping and navigation systems. Advancements in AI technology have enabled the creation of customizable maps through tools like DALL-E 2, providing options for mapped regions, locations, and map types. However, there remains a challenge in trusting machine-generated maps due to the inherent "black-box" nature of AI algorithms, raising questions among cartographers.

In the context of Google Maps, AI is harnessed to predict traffic patterns, analyze historical data, and offer real-time updates, enhancing user experience and route planning. Collaboration between Google Maps and DeepMind further refines predictions by analyzing roads, speed limits, and closures.

Indoor mapping, powered by AI, employs various techniques such as proximity positioning, trilateration, fingerprinting, and motion positioning, providing precise location data indoors. Robots, equipped with AI, assist in mapping indoor environments, solving localization challenges through metric or topological representations.

Moreover, AI-driven in-store maps bridge online and offline experiences, aiding customers and managers alike. These digital maps guide users to products, offer discounts, and enhance the overall

shopping experience. In complex environments like airports, AI-generated maps streamline traffic flow, considering the intricate features of aprons, terminals, taxiways, and runways.

Furthermore, AI algorithms enable the visualization of time series data on maps, offering insights into data changes over time. Novel approaches are being explored, such as displaying differences in data increase and decrease using distinct colors, showing promise for future applications.

In essence, the integration of AI into mapping and navigation systems has revolutionized how we perceive, interact with, and benefit from spatial data, shaping a future where intelligent mapping solutions continue to evolve and enhance our daily lives.

References:

[1] Pointr - https://www.pointr.tech/blog/ai-maps-how-is-ai-being-used-in-mapping
[2] Mark Monmonier. How to lie with maps. University of Chicago Press, 2018.
[3] Gengchen Mai, Weiming Huang, Jin Sun, Suhang Song, Deepak Mishra, Ninghao Liu, Song Gao, Tianming Liu, Gao Cong, Yingjie Hu, Chris Cundy, Ziyuan Li, Rui Zhu, and Ni Lao. On the opportunities and challenges of foundation models for geospatial artificial intelligence, 2023. arXiv:2304.06798
[4] Yuhao Kang, Song Gao, and Robert E Roth. Transferring multiscale map styles using generative adversarial networks. International Journal of Cartography, 5(2-3):115–141, 2019.
[5] Anthony C Robinson and Xi Zhu. Visualizing viral cartography with mapreverse. GI_Forum, 10(1):91–97, 2022. D. Harel, First-Order Dynamic Logic, volume 68 of Lecture Notes in Computer Science, Springer-Verlag, New York, NY, 1979. doi:10.1007/3-540-09237-4.
[6] Azelle Courtial, Guillaume Touya, and Xiang Zhang. Deriving map images of generalised mountain roads with generative adversarial networks. International Journal of Geographical Information Science, 37(3):499–528, 2023.
[7] Taisheng Chen, Menglin Chen, A-Xing Zhu, and Weixing Jiang. A learning-based approach to automatically evaluate the quality of sequential color schemes for maps. Cartography and Geographic Information Science, 48(5):377–392, 2021
[8] Yuhao Kang, Song Gao, and Robert Roth. A review and synthesis of recent geoai research for cartography: Methods, applications, and ethics. In AutoCarto 2022, 2022
[9] Kang Y, Zhang Q, Roth R. The Ethics of AI-Generated Maps: A Study of DALLE 2 and Implications for Cartography. 2023. Accessed September 19, 2023. https://ezproxy.nb.rs:2076/login.aspx?direct=true&db=edsarx&AN=edsarx.2304.10743&site=eds-live
[10] Postoast - https://www.postoast.com/how-google-maps-work-and-its-algorithm/
[11] Google The Keyword - https://blog.google/products/maps/google-maps-101-how-ai-helps-predict-traffic-and-determine-routes/
[12] Science ABC - https://www.scienceabc.com/innovation/how-does-google-maps-know-about-traffic-conditions.html#:~:text=Google%20Traffic%20works%20by%20crowdsourcing,geographic%20location%20with%20the%20app.
[13] Howstuffworks - https://electronics.howstuffworks.com/how-does-google-maps-predict-traffic.htm#:~:text=Google%20Maps%20bases%20its%20traffic,how%20fast%20cars%20are%20moving
[14] TechBallon - https://www.techballoon.com/news/google-maps-uses-ai-to-improve-travel-time-accuracy-for-users
[15] MIT News - https://news.mit.edu/2020/artificial-intelligence-digital-maps-0123
[16] Candra, A.; Budiman, M.A.; Hartanto, K. Dijkstra's and A-Star in Finding the Shortest Path: A Tutorial. In Proceedings of the 2020 International Conference on Data Science, Artificial Intelligence, and Business Analytics (DATABIA), IEEE, Medan, Indonesia, 16–17 July 2020; pp. 28–32

[17] Luo, M.; Hou, X.; Yang, J. Surface Optimal Path Planning Using an Extended Dijkstra Algorithm. IEEE Access 2020, 8, 147827–147838.
[18] Lee, D.C. Proof of a Modified Dijkstra's Algorithm for Computing Shortest Bundle Delay in Networks with Deterministically Time-Varying Links. IEEE Commun. Lett. 2006, 10, 734–736.
[19] Brilliant.org - https://brilliant.org/wiki/dijkstras-short-path-finder/#citation-4
[20] AIhouse - https://www.aihouse.com/product/design
[21] Zavlangas, P.G., Tzafestas, S.G. (2002). Integration of Topological and Metric Maps for Indoor Mobile Robot Path Planning and Navigation. In: Vlahavas, I.P., Spyropoulos, C.D. (eds) Methods and Applications of Artificial Intelligence. SETN 2002. Lecture Notes in Computer Science (), vol 2308. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-46014-4_12
[22] Moravec, H.P., Elfes, A.: High resolution maps from wide angle sonar. Proc. of the IEEE International Conference on Robotics and Automation, (1985) pp. 116-121
[23] Dudek, D., Jenkin, M., Milios, E., Wilkes, D.: Robotic exploration as graph construction. IEEE Trans. on Robotics and Automation, 7(6) (1991) 859-865
[24] Mapsted - https://mapsted.com/blog/what-is-in-store-navigation
[25] SmallBusinessBonFire - https://smallbusinessbonfire.com/indoor-store-mapping-features-and-benefits-for-retailers/
[26] https://www.pointr.tech/grocery-store-map?hsCtaTracking=e8b86003-11c0-4b81-89be-fb93ff50b8ec%7C2f41bcb6-82cb-43f6-bb18-5a543b840e8e
[27] Touya, Guillaume & Girres, Jean-François. (2014). Generalising Unusual Map Themes from OpenStreetMap.
[28] Brewer, C. A. and B. Buttenfield (2007). Framing Guidelines for Multi-Scale Map Design Using Databases at Multiple Resolutions. Cartography and Geographic Information Science 34 (1), 3-15.
[29] Duchêne, C., S. Bard, X. Barillot, A. Ruas, J. Trevisan, and F. Holzapfel (2003). Quantitative and qualitative description of building orientation. In fifth workshop on progress in automated map generalisation, Paris, France
[30] geoapify - https://www.geoapify.com/visualizing-time-series-data-on-the-map
[31] https://www.researchgate.net/figure/Summary-of-data-sources-from-northern-monitoring-programs -located-within-the_fig1_236631422 and the full text is at https://www.researchgate.net/publication/236631422_Workshop_Report_Developing_the_Canadian_Arctic_Freshwater_Biomonitoring_Network

Machine Learning Algorithms for Heart Disease Prognosis using IoMT Devices

Anita Petreska[1], Blagoj Ristevski[1] Daniela Slavkovska[2,] Saso Nikolovski[1], Pero Spirov[2], Nikola Rendevski[1], Snezhana Savoska[1]

[1] *Faculty of Information and Communication Technologies, Bitola, University "St. Kliment Ohridski"- Bitola, Republic of Macedonia*
[2] *UTMS Skopje, Faculty of Computer Science, Republic of North Macedonia,*

petreska.anita@uklo.edu.mk, blagoj.ristevski@uklo.edu.mk, dslavkovska@yahoo.com, saso.nikolovski@ uklo.edu.mk, spirov.pero@gmail.com, nikola.rendevski@uklo.edu.mk, snezana.savoska@uklo.edu.mk

Abstract:
Artificial intelligence (AI) and Machine learning (ML) algorithms have improved the capabilities of Internet of Medical Things (IoMT) systems in the direction of better quality of life and efficient healthcare. The paper analyzes various IoMT technologies and the classification, architecture, and communication of IoMT. The paper presents the application of machine learning algorithms for the detection of anomalies crucial in the process of prognosis of heart diseases. Cardiovascular disease is the primary global cause of death, prompting increased interest in leveraging artificial intelligence to analyze data obtained from wearable devices. The methodology for predicting heart disease risk using IoMT data includes data collection, pre-processing, and application of machine learning algorithms. A comparative evaluation of five machine learning models: logistic regression, Support Vector Machine (SVM), Decision Tree, Random Forest, and k-nearest neighbors (KNN) was conducted. The purpose of this paper is to emphasize the role of ML in the field of cardiology and the critical importance of data quality, as well as the selection of an appropriate algorithm in order to improve cardiovascular risk assessment. The introduction of ML in the prognosis of heart disease is a significant step towards the realization of predictive, preventive, and personalized health care and the reduction of health care costs.

Keywords:
Internet of Medical Things, machine learning, artificial intelligence, medical Big Data, exploratory data analysis

1. Introduction

The Internet of Medical Things are devices that continuously collect and transmit health data in real-time, enabling early disease detection and personalized patient care. Heart disease is a global health concern and its accurate prediction is of paramount importance. This paper explores the application of machine learning algorithms to cardiac disease prognosis using data collected with the IoMT.

The second part looks at the role of IoMT in the healthcare sector, the architecture, communication, and management of IoMT devices, and defines the technologies used and their security.

Section 3 presents the motivation for introducing ML in healthcare with special reference to commonly used ML algorithms for detecting patterns and anomalies crucial in the process of heart disease prognosis modeling.

In Section 4, five machine learning algorithms are analyzed to detect patterns and anomalies that are key to detecting and diagnosing heart disease risk. The methodology used, data collection, data pre-processing, exploratory data analysis [9], and metrics for evaluation and comparison of results are presented. Section 5 discusses the limitations faced by ML algorithms used to detect heart disease. Section 6 discusses the conclusions and future challenges of ML algorithms for heart disease prediction.

2. The role of IoMT in healthcare, architectural framework, technologies, and security

The IoMT is a network of medical devices, sensors, and software applications connected over the Internet. This network collects and delivers real-time data related to patient health, providing healthcare professionals with invaluable real-time insights into patient conditions enabling personalized healthcare interventions [1]. IoMT encompasses different categories of devices: wearable, implantable, ingestible, stationary, diagnostic, therapeutic, assistive, smart, industrial, emergency, telemedicine, and prosthetic [2]. IoMT devices facilitate continuous, remote patient monitoring, and dynamic data-driven insights, enabling healthcare professionals to detect subtle changes in a patient's condition and adjust interventions accordingly. A core feature of the transformative potential of IoMT is the integration of multimodal data, providing healthcare providers with a comprehensive perspective on patient health and thereby facilitating more effective diagnostic and treatment strategies.

The IoMT architecture consists of a synergy of sensors and the IoMT Gateway, which work together to enable the collection, transmission, analysis, and storage of vital health data. Wireless communication serves as the key, with technologies such as Bluetooth, Wi-Fi, ZigBee, and cellular networks offering mobility and flexibility, while near-field communication (NFC) and radio frequency identification (RFID) facilitate contactless communication [3]. Standardized communication protocols such as Health Level 7 (HL7) and Digital Imaging and Communications in Medicine (DICOM) ensure seamless interoperability and data exchange between IoMT devices and Electronic Health Records (EHR). IoMT management is segmented across different layers, each with a distinct function.

In the field of technology and security [4], the IoMT uses a comprehensive range of tools and measures: cloud computing, blockchain technology, big data analytics, interoperability standards, edge computing, artificial intelligence, machine learning, blockchain technology, and virtual and augmented reality. , medical device integration (MDI) and security measures underpin the IoMT landscape [5], [6]. Cloud computing assumes a key role, offering scalable storage and computing resources [7]. Blockchain technology protects the security and privacy of data, creating a decentralized database [8]. Artificial intelligence identifies patterns and improves patient care. Virtual and augmented reality technologies contribute to professional training, remote consultation with patients, and monitoring of treatment progress.

Security measures [9] and [10] are imperative due to the sensitivity of medical data. Encryption, authentication, access control, physical security, vulnerability assessments, and timely remediation are necessary to protect patient data and system integrity [11], [12].

3. Motivation for using ML to detect heart disease

The motivation for using ML in cardiac disease prognosis within the IoMT is multifaceted. First, heart disease is a leading cause of morbidity and mortality worldwide, requiring early diagnosis and intervention. ML algorithms offer the potential to analyze vast datasets, spanning diverse clinical and physiological attributes to uncover complex patterns and dependencies that may not be apparent through conventional methods.

Heart disease encompasses a range of conditions affecting the heart and blood vessels that can lead to a variety of health problems, including narrowing of the blood vessels, chest pain, stroke, and heart attack. The main causes of heart disease are diabetes, obesity, unhealthy diet, increased weight, excessive alcohol use, and physical inactivity. Early prediction of heart disease is paramount for patients and healthcare providers. Early identification allows health professionals to implement preventive measures, effective diagnosis and treatment, and patients valuable insights into their health. Machine learning plays a key role in identifying and predicting heart disease. Machine learning algorithms applied to relevant medical data serve as powerful tools for identifying patients at risk of heart disease before symptoms become apparent.

Numerous AI algorithms and models use the data generated by wearable sensor devices. In the domain of establishing diagnostic and predictive models using data from wearable devices, classical

machine learning and deep learning techniques are of paramount importance. The development of appropriate algorithms and models tailored to specific categories of heart disease remains imperative.

By using machine learning algorithms, healthcare providers and patients can work together to detect heart disease in its early stages, enabling timely interventions and informed decisions.

Deep learning is a powerful tool in the field of heart disease prediction that uses multilayer neural network architectures to automatically learn and extract complex patterns and features from complex medical data. Convolutional Neural Networks (CNNs) are used to automatically detect anomalies and irregularities in ECG signals, helping in the early detection of cardiac conditions such as arrhythmias [13]. Convolutional neural networks (CNNs) are applied to medical imaging data, including heart scans (echocardiograms, MRIs, CT scans). They help in the automatic interpretation of the images, helping in the diagnosis of structural problems of the heart. Recurrent neural networks (RNNs) and their specialized variant [14], long-short-term memory (LSTM) networks excel at modeling temporal sequences [15]. They can capture patterns and dependencies in time series data, making them invaluable in predicting heart attacks and strokes.

4. Machine learning algorithms for prognosis of heart diseases

Machine learning uses different algorithms and models to predict heart disease. These algorithms analyze medical data, such as patient records, medical images, and diagnostic tests, to generate predictive insights. Some of the most common machine learning algorithms for heart disease prediction include logistic regression, random forests, support vector machines, neural networks, gradient boosting, and K-nearest neighbors, among others. These algorithms take into account multiple factors, including a patient's medical history, genetics, lifestyle, and environmental factors to estimate the risk of developing heart disease.

Efficient machine learning models have been developed for the prognosis of heart disease. Five machine learning algorithms were used in our study: logistic regression, SVM (Support Vector Machines), Decision Tree, Random Forest, and KNN [16].

Logistic regression is used to predict the probability of heart disease based on a combination of clinical and physiological factors, as input characteristics and the binary outcome of the presence (1) or absence (0) of heart disease.

Support Vector Machine is a powerful machine learning algorithm that recognizes complex patterns in medical data, particularly when classifying individuals into different risk categories for heart disease. Support vector machine is also efficient for high-dimensional data and non-linear relationships between them, making it excellent for capturing complex dependencies in medical data.

The Decision Tree algorithm builds on a tree-like hierarchical structure that efficiently evaluates data, aiding early detection and personalized management of heart disease.

Random Forest combines the predictions of many decision trees, providing a robust and accurate way to estimate heart disease risk. Random Forest excels in dealing with complex and high-dimensional medical data by being able to capture complex relationships in this data, which cannot be done with simple methods. Random Forest not only increases the prediction accuracy and improves the generalization of the model, but also can reduce overfitting.

The K-Nearest Neighbors (KNN) algorithm is an intuitive method for classifying individuals into risk categories based on their similarity to other patients. KNN although simple and flexible can detect complex relationships in data, making it a good choice for heart disease risk prediction.

4.1. Methodology

The purpose of this study is to determine whether a patient will experience a heart attack based on the data collected by the IoMT device. For this purpose, machine learning methods have been proposed that would be used by doctors in order to diagnose heart diseases more easily.

The methodology includes a series of procedures that take place in 3 phases. In the first phase, data from IoMT devices is collected and prepared for processing. In the second stage, preprocessing is performed, which includes analysis of missing values, cleaning, and standardization. The third stage

involves applying a classifier to build an appropriate machine learning model. The proposed model uses 5 machine learning algorithms: logistic regression, SVM, Decision Tree, Random Forest and KNN.

The process of heart disease risk prediction is shown in Figure 1. Data are loaded from Kaggle database [17], cleaned, significant features are extracted. The database is divided into training and testing sets with the ratio (80% and 20%). The data from the five considered machine learning models are trained and tested and finally the results are evaluated and compared in order to evaluate the efficiency of the proposed methodology by evaluating the models and determining the accuracy, cross-validation result, ROC_AUC result.

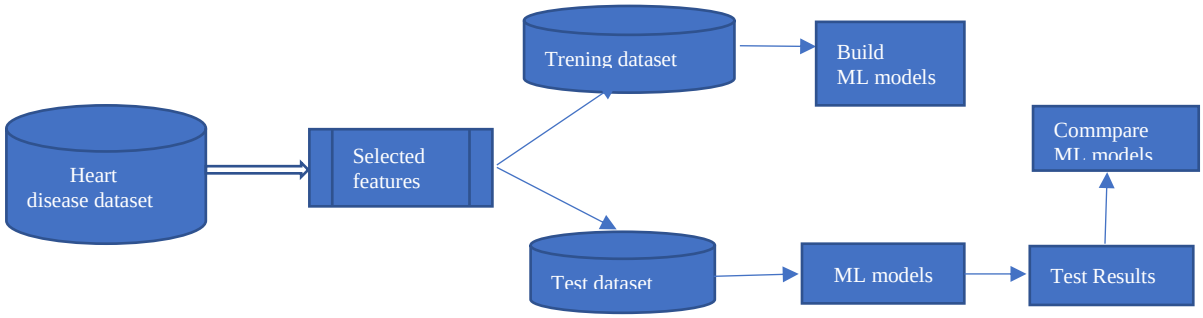


Figure 1: Flowchart of ML algorithms for heart disease prediction

4.2. Data collection

In our study, a database downloaded from Kaggle is used, which consists of 303 queues containing 14 attributes of which 13 are input attributes and one target attribute indicating the presence or absence of heart disease. All 14 attributes shown in Table 1 contain demographic data and clinical parameters that cover a different aspect of an individual's health profile.

Table 1:
Heart prediction dataset attributes and information

Feature	Characteristic representation	specific	Range/Value
age	Patient's age in years	[continuous variable]	29-77
sex	0 = female ; 1 = male	[categorical variable]	0,1
cp	Chest pain type 0 = typical angina 1 = atypical angina 2 = non-anginal pain 3 = asymptomatic	[categorical variable]	0,1,2,3
trestbps	Resting blood pressure (mm hg)	[continuous variable]	94-200
chol	Cholesterol (mg/dl)	[continuous variable]	126-564
fbs	1 = True, 0 = False	[categorical variable]	0,1
restecg	Resting electrocardiographic results ~ 0 = Normal, 1 = ST-T wave normality, 2 = Left ventricular hypertrophy	[categorical variable]	0,1,2
thalach	Maximum heart rate achieved	[continuous variable]	71-202
exng	Exercise induced angina ~ 1 = Yes, 0 = No	[categorical variable]	0,1
oldpeak	Previous peak	[continuous variable]	0-6,2
slp	Slope of exercise ST segment 0 = unslope 1 = flat 2 = downslope	[categorical variable]	0,1,2
Ca	No. of major vessels [0-2] colored by fluoroscopy	[continuous variable]	0,1,2
thal	Defect type 1 = fixed defect; 2 = normal; 3 = reversable defect	[categorical variable]	1,2,3
target	Has heart disease or not, 0 = no 1 = yes	[target variable]	0,1

In the data preprocessing process, categorical variables are converted into numerical variables, where each category is represented by a binary (0 or 1) indicator variable.

Identifying and addressing missing values is critical. Depending on the extent of missing data, rows with missing data may be discarded or imputation methods may be used that include filling in missing values with the mean, and median, or using more advanced techniques such as interpolation.

It is essential to detect and address outliers appropriately. This may include visualizations (eg, box plots, scatter plots) and statistical methods to identify data points that deviate significantly from the norm.

Scaling is crucial when numerical features have different scales. Standardization involves transforming the data to have a mean of 0 and a standard deviation of 1, while normalization scales the data to a specified range (for example, between 0 and 1).

The listed steps contribute to the preparation of the data for further analysis or machine learning modeling. It is worth noting that the specific techniques and methods used may vary based on the nature of the data, the problem at hand, and the requirements of the chosen analysis or model [18].

Chart 1 and chart 2 present the distribution of numerical characteristics and categorical characteristics respectively.

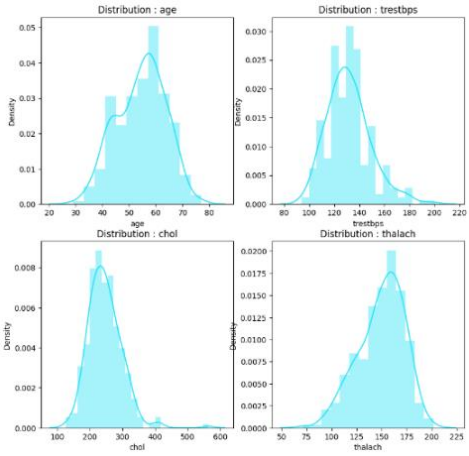


Chart 1: Distribution of numerical features

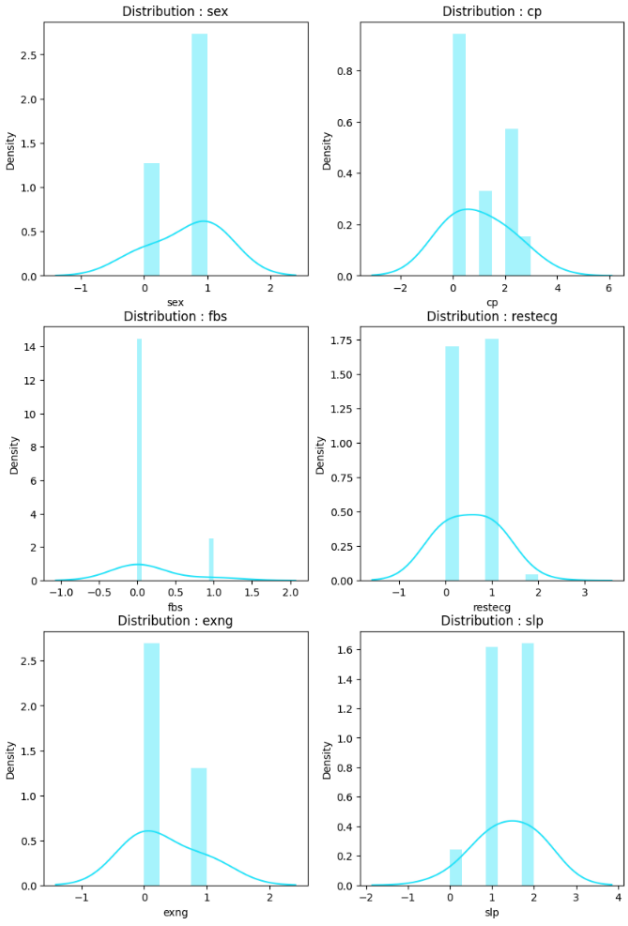


Chart 2: Distribution of categorical features

Table 2 presents the basic statistical parameters of the database.

Table 2:
Basic statistical parameters

	count	mean	std	min	25%	50%	75%	max
age	303.00	54.37	9.08	29.00	47.50	55.00	61.00	77.00
sex	303.00	0.68	0.47	0.00	0.00	1.00	1.00	1.00
cp	303.00	0.97	1.03	0.00	0.00	1.00	2.00	3.00
trestbps	303.00	131.62	17.54	94.00	120.00	130.00	140.00	200.00
chol	303.00	246.26	51.83	126.00	211.00	240.00	274.50	564.00
fbs	303.00	0.15	0.36	0.00	0.00	0.00	0.00	1.00
restecg	303.00	0.53	0.53	0.00	0.00	1.00	1.00	2.00
thalach	303.00	149.65	22.91	71.00	133.50	153.00	166.00	202.00
exng	303.00	0.33	0.47	0.00	0.00	0.00	1.00	1.00
oldpeak	303.00	1.04	1.16	0.00	0.00	0.80	1.60	6.20
slp	303.00	1.40	0.62	0.00	1.00	1.00	2.00	2.00
ca	303.00	0.73	1.02	0.00	0.00	0.00	1.00	4.00
thal	303.00	2.31	0.61	0.00	2.00	2.00	3.00	3.00
target	303.00	0.54	0.50	0.00	0.00	1.00	1.00	1.00

4.3. Exploratory data analysis

Exploratory Data Analysis (EDA) is a fundamental step in predicting heart disease using machine learning and aims to explore and understand essential insights and patterns in a database. EDA is performed with a systematic approach, includes visual and statistical examinations, provides a deep understanding of the data, and creates a basis for the development of accurate and relevant models [19]. Visual displays such as histograms, scatterplots, and ROC curves provide a comprehensive understanding of data distributions, feature importance, and model performance metrics. Data visualization is an indispensable tool in the field of predicting heart disease using machine learning models, helping to understand the complex relationships between characteristics such as age, cholesterol levels, and blood pressure and their impact on heart disease risk. Metrics such as ROC curves, confusion matrices, and precise regression curves provide insight into the accuracy and ability of the model to distinguish patients with and without heart disease [20]. Figure 2 shows the mean values of all attributes for cases with and without heart disease risk.

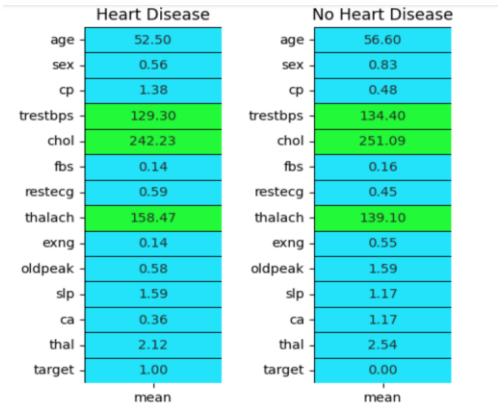
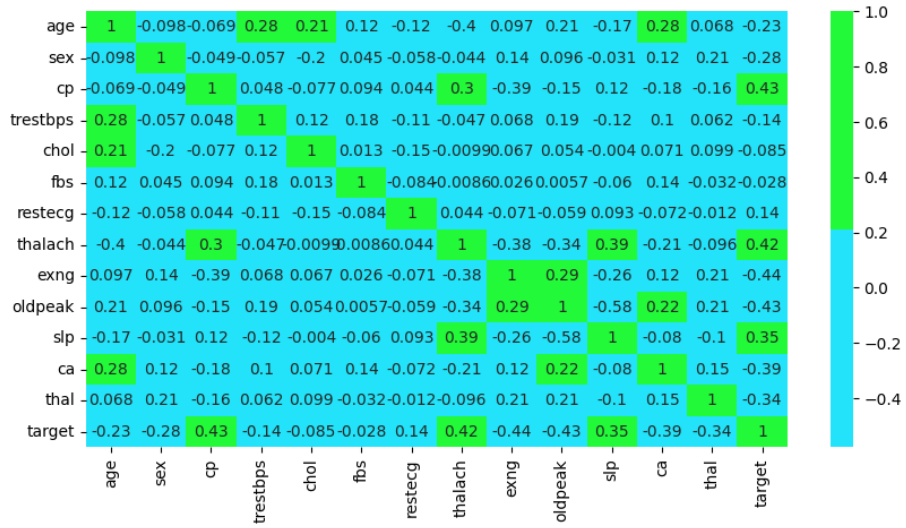


Figure 2: Mean values of all characteristics there is and there is no risk of heart disease

Correlation was performed to reveal potential relationships between variables in order to highlight interactions and dependencies in the dataset. This analysis provides valuable insights into the potential influence between variables as well as their joint influence on heart disease risk. From the correlation matrix, insight is obtained into which characteristics are positively or negatively correlated with each other and with the target variable. According to Table 3, the target variable is a dependent variable that

is negatively correlated with some of the factors such as thali, and positively correlated with some of the parameters such as that, ca, etc. Visualization of the correlation matrix provides insight into which characteristics are positively or negatively correlated with the target variable.

Table 3: Correlation matrix



4.4. Evaluation metrics

Evaluation metrics are crucial for evaluating the performance of machine learning models [21]. Accuracy is a metric that measures the percentage of correctly classified cases out of all cases and represents the ability to correctly identify true positives while minimizing false positives, which is critical to patient safety.

Cross-validation divides the dataset into multiple subsets so that training and testing are performed multiple times, allowing each subset of data to be part of both the training and testing set. Cross-validation helps to identify potential problems such as overfitting and underfitting [22].

Receiver Operating Characteristics – AreaUnder Curve (ROC-AUC) is used to quantify the ability of the predictive model to distinguish between positive and negative cases. In binary classification tasks. The ROC-AUC metric helps assess the ability of a model to maintain a balance between true positives and false positives at different classification thresholds [23].

The confusion matrix provides a detailed overview of the classification results of the model, categorizing the predictions into four basic categories: true positives (correctly predicted cases of heart disease), true negatives (correctly predicted cases without heart disease), false positives (incorrectly predicted cases of heart disease), and false negatives (incorrectly predicted cases without heart disease [24].

4.5. Comparison of results

Table 4 shows the Evaluation Scores of the five considered ML algorithms by evaluating their performance in terms of accuracy, Cross Validation Score, and ROC_AUC Score in order to build a model for predicting heart diseases.

In terms of accuracy, SVM achieved the highest accuracy of 90.16%, indicating that it correctly predicted the heart disease status for the majority of cases. Random forest and logistic regression also performed well, with 86.89% accuracy. KNN and Decision Tree scored lower for accuracy, with KNN at 85.25% and Decision Tree at 77.05%.

In terms of the Cross Validation Score metric, Random Forest received the highest cross-validation score of 89.67%, indicating its consistent performance across different data subsets. SVM had a slightly

lower but still strong cross-validation score of 87.96%. Logistic regression also performed well in cross-validation, with a score of 88.52%. KNN and Decision Tree had lower cross-validation scores of 84.98% and 79.82%, respectively.

Table 4: Accuracy, Cross Validation Score, and ROC AUC Score for different ML models

model	accuracy	Cross Validation Score	ROC_AUC Score
Logistic regression	86.89%	88.52%	87.34%
SVM	90.16%	87.96%	90.62%
Decision Tree	77.05%	79.82%	76.99%
Randoom Forest	86.89%	89.67%	87.34%
KNN	85.25%	84.98%	85.61%

Analysis of the ROC-AUC results showed that SVM achieved the highest ROC-AUC score of 90.62%, indicating its excellent ability to discriminate between patients with and without heart disease. Random forest and logistic regression share the same ROC-AUC score of 87.34%, indicating their comparable discrimination power. KNN and Decision Tree had slightly lower ROC-AUC scores of 85.61% and 76.99%, respectively, suggesting that they may not perform as well in distinguishing between positive and negative cases.

Of the algorithms analyzed, SVM stands out as the top-performing model, boasting the highest accuracy and ROC-AUC score, making it a strong choice for heart disease prediction. Random Forest and Logistic Regression also provide robust performance across all metrics. KNN and Decision Tree, while achieving reasonably good results, have slightly lower scores in accuracy, cross-validation, and ROC-AUC, indicating slightly lower predictive power in this particular dataset.

The choice of the best model ultimately depends on the specific application requirements and the relative importance of the specified metrics in the context of heart disease prediction. The methodology must be constantly improved and developed, so that other parameters such as echocardiographic data and medical recordings may be included in some future considered models.

5. Limitations of machine learning models for heart disease prognosis

Data quality and quantity play a key role in the performance of ML algorithms for heart disease prognosis. IoMT devices generate huge amounts of data, which often leads to problems with data imbalance, noise, and missing values. The need for comprehensive, standardized data collection protocols is becoming apparent, as the absence of such practices can hinder the accuracy and generalizability of predictive models. The interpretability and transparency of ML algorithms in health applications, especially for the prognosis of critical heart diseases, despite their high prediction accuracy, cannot often provide meaningful insights into the decision-making process, which raises ethical and regulatory concerns, given the fact that doctors and patients demand transparent explanations for the recommendations made by these algorithms. The issue of bias and fairness in ML models for heart disease prognosis is a critical limitation.

Biases in data and algorithmic decisions may disproportionately affect certain demographic groups, potentially leading to unfair results. Providing robust, universally applicable models requires careful consideration of these factors.

Limitations imposed by the hardware and software capabilities of the IoMT devices themselves can hamper the effectiveness of ML algorithms. Connectivity issues, security issues, and device compatibility challenges must be addressed to facilitate seamless integration of these devices with ML algorithms.

Future research should focus on developing more robust and personalized models to improve cardiovascular risk assessment.

Machine learning algorithms integrated with IoMT devices hold great promise for advancing heart disease prediction, enabling early detection, personalized care, remote monitoring, and more efficient allocation of healthcare resources. However, addressing data privacy, data quality, regulatory

compliance, and model interpretability are essential to unlocking the full potential of this technology to strengthen patient trust and data security.

6. Conclusions and challenges of machine learning models for heart disease prognosis

IoMT devices improve patient outcomes by enabling early detection of health problems, continuous monitoring, and personalized interventions. IoMT has the potential to reduce healthcare costs by minimizing hospital readmissions, optimizing resource utilization, and streamlining workflows and patient care in rural settings. It empowers individuals to actively participate in their own health management and promotes patient engagement.

The data processing process applies advanced analytics, machine learning algorithms, and artificial intelligence techniques to extract meaningful insights from the data. Choosing an appropriate machine learning algorithm depends on the specific task and the characteristics of the data. These insights help healthcare professionals make the right and personalized decisions. IoMT facilitates remote patient monitoring, allowing healthcare providers to monitor patients' health conditions and intervene in real-time when needed. IoMT enables virtual consultations, remote diagnostics, and telehealth services, and supports chronic disease monitoring, medication management, and preventive care.

The future of machine learning models in the prediction of numerous diseases brings enormous challenges in the direction of applying advanced algorithms such as deep learning, neural networks, hybrid model research, and enabling real-time monitoring. The continued evolution of machine learning has the potential to revolutionize early diagnosis, intervention, and patient care in healthcare.

Data is securely transmitted to cloud-based platforms or local servers, where it can be stored, analyzed, and processed, however, the widespread adoption of IoMT raises security and privacy concerns given that it is sensitive patient data. Interoperability issues between different IoMT devices and systems pose a challenge for seamless integration and data exchange. Healthcare organizations need to ensure their IoMT devices comply with regulatory requirements such as HIPAA and GDPR [25] to avoid legal and financial penalties.

References

[1] Mansour, Romany Fouad, et al. "Artificial intelligence and internet of things enabled disease diagnosis model for smart healthcare systems." *IEEE Access* 9 (2021): 45137-45146.

[2] Dwivedi, Ruby, Divya Mehrotra, and Shaleen Chandra. "Potential of Internet of Medical Things (IoMT) applications in building a smart healthcare system: A systematic review." *Journal of oral biology and craniofacial research* 12.2 (2022): 302-318.

[3] Huang, Chenxi, et al. "Internet of Medical Things: A Systematic Review." *Neurocomputing* (2023): 126719.

[4] Kamalov, Firuz, et al. "Internet of medical things privacy and security: Challenges, solutions, and future trends from a new perspective." *Sustainability* 15.4 (2023): 3317.

[5] Dang, Van Anh, et al. "Intelligent Healthcare: Integration of Emerging Technologies and Internet of Things for Humanity." *Sensors* 23.9 (2023): 4200.

[6] Ameen, Aya Hamid, Mazin Abed Mohammed, and Ahmed Noori Rashid. "Dimensions of artificial intelligence techniques, blockchain, and cyber security in the Internet of medical things: Opportunities, challenges, and future directions." *Journal of Intelligent Systems* 32.1 (2023): 20220267.

[7] Ashfaq, Z., Rafay, A., Mumtaz, R., Zaidi, S.M.H., Saleem, H., Zaidi, S.A.R., Mumtaz, S. and Haque, A., 2022. A review of enabling technologies for the Internet of Medical Things (IoMT) Ecosystem. Ain Shams Engineering Journal, 13(4), p.101660

[8] Elhachmi, J. A. M. A. L., and A. B. D. E. L. L. A. T. I. F. Kobbane. "BLOCKCHAIN-BASED SECURITY MECHANISMS FOR INTERNET OF MEDICAL THINGS (IOMT)." *International Journal of Computer Networks and Communications* (2022): 115-136.

[9] Kamruzzaman, M. M., Ibrahim Alrashdi, and Ali Alqazzaz. "New opportunities, challenges, and applications of edge-AI for connected healthcare in the internet of medical things for smart cities." *Journal of Healthcare Engineering* 2022 (2022).
[10] Savoska, Snezana, et al. "Integration of heterogeneous medical and biological data with electronic personal health records." Medical Informatics, Biometry and Epidemiology (MIBE) 19 (2023): 1-11.
[11] Savoska, Snezana, Blagoj Ristevski, and Vladimir Trajkovik. "Personal Health Record Data-Driven Integration of Heterogeneous Data." *Data-Driven Approach for Bio-medical and Healthcare*. Singapore: Springer Nature Singapore, 2022. 1-21.
[12] Gopalan, Subiksha Srinivasa, Ali Raza, and Wesam Almobaideen. "IoT security in healthcare using AI: A survey." *2020 International Conference on Communications, Signal Processing, and their Applications (ICCSPA)*. IEEE, 2021.
[13] Bao, Xinqi, et al. "Time-frequency distributions of heart sound signals: A comparative study using convolutional neural networks." *Biomedical Engineering Advances* 5 (2023): 100093.
[14] Chitra, S., and V. Jayalakshmi. "Prediction of heart disease and chronic kidney disease based on internet of things using RNN algorithm." *Proceedings of Data Analytics and Management: ICDAM 2021, Volume 1*. Springer Singapore, 2022.
[15] Zhang, Wanruo, et al. "Motion prediction of beating heart using spatio-temporal LSTM." *IEEE Signal Processing Letters* 29 (2022): 787-791.
[16] Uddin, Shahadat, et al. "Comparing different supervised machine learning algorithms for disease prediction." *BMC medical informatics and decision making* 19.1 (2019): 1-16.
[17] https://www.kaggle.com/datasets/rashikrahmanpritom/heart-attack-analysis-prediction-dataset
[18] Md, Abdul Quadir, et al. "Enhanced Preprocessing Approach Using Ensemble Machine Learning Algorithms for Detecting Liver Disease." *Biomedicines* 11.2 (2023): 581.
[19] Kandpal, Jyoti, Geetanjali Baluti, and Praful Ranjan. "Integration of Artificial Intelligence with Opto-VLSI for the Biomedical and Healthcare Industries." *Opto-VLSI Devices and Circuits for Biomedical and Healthcare Applications*. CRC Press 161-175.
[20] Senthilkumar, S. A., et al. "Big data in healthcare management: a review of the literature." *American Journal of Theoretical and Applied Business* 4.2 (2018): 57-69.
[21] Hicks, Steven A., et al. "On evaluation metrics for medical applications of artificial intelligence." *Scientific Reports* 12.1 (2022): 5979.
[22] Xiong, Zheng, et al. "Evaluating explorative prediction power of machine learning algorithms for materials discovery using k-fold forward cross-validation." *Computational Materials Science* 171 (2020): 109203.
[23] Habenicht, Richard, et al. "Machine Learning Based Linking of Patient Reported Outcome Measures to WHO International Classification of Functioning, Disability, and Health Activity/Participation Categories." *Journal of Clinical Medicine* 12.17 (2023): 5609.
[24] Vujović, Ž. "Classification model evaluation metrics." *International Journal of Advanced Computer Science and Applications* 12.6 (2021): 599-606.
[25] Shuaib, Mohammed, et al. "Compliance with HIPAA and GDPR in a blockchain-based electronic health record." Materials Today: Proceedings (2021).

Comparative Analysis of ML Algorithms for Breast Cancer Detection

Daniela Slavkovska[1], Blagoj Ristevski,[2] Anita Petreska[2]

[1]*UTMS, Skopje Faculty of Computer Science, Republic of North Macedonia*
[2]*Faculty of Information and Communication Technologies - Bitola, University "St. Kliment Ohridski"- Bitola, Republic of North Macedonia*

dslavkovska@yahoo.com, blagoj.ristevski@uklo.edu.mk, petreska.anita@uklo.edu.mk

Abstract:
Artificial intelligence and machine learning algorithms with their advanced predictive and diagnostic techniques help healthcare providers make the right decisions in the process of disease prevention and early diagnosis. Breast cancer as a disease of modern dynamic living is gaining momentum. Its early detection is crucial to increase the chances of survival through better treatment options.
Exploratory data analysis (EDA) as a key step in data analysis involves systematic examination and visualization of data to discover patterns, outliers and dependencies, enabling hypothesis generation and making correct decisions. As powerful algorithms applied for classification, Logistic regression is used, K-nearest neighbors (KNN), Naïve Bayes, Decision Tree, Random Forest, XGBoost Support, Deep Learning and NeNetwork. In our study, a comparative analysis of the most commonly used Machine Learning (ML) algorithms has been done by evaluating various metrics such as accuracy, F-measure, confusion matrix and specificity.
Limitations of machine learning algorithms often include issues such as overhead, high computational requirements, and data quality. Challenges may arise from the need for large labeled datasets, algorithmic bias, and concerns about interpretability.
Future work in machine learning should focus on developing more robust models that can generalize well to diverse data, improving the interpretability of complex models.
Keywords:
Machine Learning, Breast Cancer, Artificial intelligence, Medical big data, Healthcare, Data mining, Exploratory Data Analysis

1. Introduction

In recent years, healthcare has been facing numerous challenges, both from the increased accuracy of diagnosis and therapy and the need for an efficient way of managing large amounts of health data[1]. The ability to extract useful knowledge hidden in large amounts of data and act on that knowledge is becoming increasingly challenging. Cancer, especially breast cancer [2][7], is one of the most dangerous diseases in the world, claiming the lives of countless women every year. Machine learning algorithms [8] play a major role in the early detection and diagnosis of cancer. Section 2 reviews the role of machine learning and Data Mining in healthcare [6][15], and Section 3 presents the most commonly used ML algorithms in healthcare [3]. Part 4 presents the application of machine learning algorithms in order to discover patterns and anomalies that are crucial in the detection and diagnosis of breast carcinoma, the used methodology, data collection, data preprocessing, Exploratory Data Analysis [9], evaluation metrics and comparison are presented of the results. In the fifth part, the limitations and challenges faced by ML algorithms are presented, and in the last, sixth part, the conclusions reached and future challenges of the ML algorithms are presented.

2. The role of machine learning in healthcare

Artificial intelligence [5][20] has a great impact on medical research that has direct application in real medical applications. Given the exponential growth of data, traditional manual data analysis is not enough. Methods for efficient computational analysis such as technologies developed in the field of data mining are necessary to discover insights from Big Data. Data Mining (DM) enables the discovery of knowledge in databases that includes understanding the domain, preprocessing the data, extracting regularities hidden in the data in order to create patterns or models. Numerous Big Data applications have been developed with innovative methods for improving the quality of health care in the direction of prevention or early detection of diseases.

Machine learning as one of the key components of artificial intelligence is used to train algorithms and models that can analyze data, extract patterns and make predictions based on experience. Machine learning leverages AI's ability to learn from data and improve its performance with experience, making AI more intelligent and adaptable to different tasks and environments.

3. Most commonly used algorithms of machine learning in healthcare

Various machine learning algorithms are used in the healthcare industry covering various aspects of medicine and healthcare.

Logistic regression [8] is a simple and interpretable algorithm that is often used for binary classification tasks, such as breast cancer prediction. It models the probability that a given case belongs to a certain class. Characteristics related to patient demographics, genetic factors, and medical history can be used as input variables.

Support vector machine (SVM) [6][16]is a powerful algorithm for both binary and multiclass classification that uses a hyperplane that maximizes the margin between different classes. SVM can effectively handle high-dimensional spaces, making it suitable for genetic and medical data.

Decision trees [6][12] are used for classification by recursively partitioning the data into subsets based on the most informative features. They are interpreted and can provide insight into which features are most important in predicting breast cancer. Decision trees may require pruning to prevent overgrowth.

Random forests [7][8] are an ensemble method that builds multiple decision trees and combines their predictions. They reduce the risk of overload associated with single-decision trees. Random forests can handle high-dimensional data and capture complex interactions between features.

KNN [6][8] is a non-parametric algorithm that classifies data points based on the majority class among their k-nearest neighbors. It can be used to predict breast cancer by looking at similar cases of patients. The choice of the 'k' parameter is important and should be determined through cross-validation.

The Naive Bayes [8][17] classifier works by calculating the probability of breast cancer occurrence based on characteristics such as age, family history, tumor characteristics, and medical test results. Although it simplifies the assumption of independence of features, it can still be effective in aiding breast cancer diagnosis and risk assessment.

XG Boost [8][18] is known for its speed, efficiency, and high prediction performance. The ability to handle a wide range of input features allows it to be applied to breast cancer detection by training a model containing numerous features extracted from breast cancer-related medical images, patient data, or genetic information.

Neural networks [17][19][21], especially deep learning models, can automatically learn complex patterns and relationships in data, making them suitable for tasks involving complex visual data. They have outstanding performance in image recognition tasks, including medical image analysis. Table 1 lists the advantages and disadvantages of the most commonly used ML algorithms.

3.1. Advantages and disadvantages of machine learning algorithms

An advantage of logistic regression is its simplicity and interpretability, which makes it suitable for solving problems related to linear dependence and basic modeling, and a disadvantage is its limited ability to model complex, non-linear data patterns.

Support Vector Machine is efficient for complex, high-dimensional data and can achieve high accuracy with proper kernel selection, but can be computationally expensive for large datasets.

The advantages of decision trees include their simplicity and interpretability, while their disadvantages include susceptibility to overfitting and instability due to small variations in the data.

The advantage of Case Forest is high accuracy, robustness to overload, and suitability for different data types and large data sets, and the disadvantage is reduced interpretability compared to a single decision tree and can be computationally intensive.

The K-Nearest Neighbors algorithm is simple and effective to implement for locally structured data without making assumptions about the data distribution and can be effective for classification tasks, but it can be computationally expensive, especially with large datasets and many rely on the choice of the "k" parameter, which can affect its performance making it less suitable for high-dimensional databases.

The advantage of XGBoost lies in its excellent power as an ensemble learning method, efficiency, and adaptability for complex problems, known for its high predictive accuracy and efficiency due to gradient boosting, but it can be prone to overfitting if not properly tuned and can to require more computational resources compared to simpler algorithms.

An advantage of a neural network is its ability to model complex relationships in data to deal with complex tasks with large data sets, including image and text analysis, and the disadvantage of a neural network is that it requires a large amount of data and computational resources, lacks transparency, which can hinder their interpretability, especially in critical applications where understanding the model's decision-making process is crucial.

3.2. Comparison of Machine Learning Algorithms in terms of Accuracy, Interpretability and Adaptability

The accuracy of machine learning algorithms can vary depending on the dataset and parameters, but in general, Random Forest, XGBoost, and Neural Networks tend to achieve high accuracy due to their ability to capture complex patterns, logistic regression and decision trees can tend to work well for simpler problems, in more complex scenarios they have lower accuracy. SVM and neural networks are very adaptable to different data types and complexity, while decision trees and random forests may require additional techniques to deal with overload, KNN is sensitive to the volume and dimensionality of the data. Algorithm selection should consider trade-offs between these aspects based on the specific problem and dataset.

In terms of interpreting the results of decision tree and logistic regression is relatively simple, they are generally more interpretable because they provide explicit rules and coefficients that connect the input characteristics to the predictions. Neural networks, SVM, Random Forest, KNN, and XGBoost are less interpretable due to their complex, non-linear, and ensemble-based nature, which allows meaningful insights to be extracted.

In terms of adaptability SVMs and neural networks are highly adaptable to different data types and complexity, making them suitable for a wide range of tasks. Decision trees and random forests are scalable but may require techniques to resolve overload, while KNNs can be sensitive to data volume and dimensions, affecting their adaptability to certain scenarios.

4. Most commonly used algorithms of machine learning in healthcare

In this chapter, a study is made based on medical data that may come from hospitals, research institutions, or publicly available data sets in order to predict the risk of breast cancer. In the study, machine learning methods are proposed for the purpose of breast cancer diagnosis.

4.1. Methodology

Methodology [4] for cancer assessment includes 10 stages shown in Table 1.

Table 1:
Methodology for breast cancer assessment

Phase	description
First stage	Data collection and pre-processing
Second phase	Data cleaning and pre -processing
Third phase	Normalize or standardize
Fourth phase	Feature selection and engineering
The fifth stage	Split the database
The sixth stage	Model selection and training
Seventh stage	Model evaluation
Eighth stage	Interpretation and visualization of the model
Ninth stage	Deployment and testing
The tenth stage	Ethical considerations and compliance

The first phase involves gathering relevant medical data, including patient demographics, genetic information, medical history, and diagnostic characteristics of breast cancer. The data cleaning is performed in order to deal with missing values, analysis of outliers, and data quality. To ensure uniform scaling, it is necessary to perform normalization and standardization of numerical characteristics and coding of categorical variables. Then divide the database into training, validation, and test sets for model evaluation. Unnecessary features are removed in order to reduce dimensionality. Univariate or multivariate analysis techniques are used in order to examine the correlation between variables.

The next step is the selection of appropriate machine learning algorithms. The selected models are trained based on the training data. In order to prevent overfitting of the model, the performance of the model is evaluated on the validation set. Models are evaluated using appropriate metrics for binary classification, such as accuracy, precision, recall, F1-score area under the ROC curve (AUC-ROC), confusion matrix, and cross-validation. It is good practice to visualize the results of the model to gain insight and interpretability. It requires continuous monitoring of the model and implementing mechanisms to update it when new data becomes available. We need to ensure that the use of patient data complies with ethical and legal regulations, such as HIPAA (in the United States) or GDPR (in Europe). The final phase involves documenting the entire methodology, including data sources, preprocessing steps, model selection, and evaluation results.

4.2. Data collection

In our study, a database downloaded from Kaggle **Error! Reference source not found.** was used, consisting of 33 rows containing 33 attributes of which 32 are input attributes and one target attribute that indicates the risk of breast cancer. Attributes contain demographic data and clinical parameters that encompass different aspects of an individual's health profile.

4.3. Data pre-processing

In the process of data preprocessing, categorical variables are converted into /indicator variables, missing values are checked if there are any, which is of crucial importance to either be rejected or filled with a mean or interpolated value. Table 2 presents the basic statistical parameters of the database

Table 2:
Basic statistical parameters

	count	mean	std	min	25%	50%	75%	max
Radius mean	569.000000	14.127292	3.524049	6.981000	11.700000	13.370000	15.780000	28.110000
Texture mean	569.000000	19.289649	4.301036	9.710000	16.170000	18.840000	21.800000	39.280000
Perimeter mean	569.000000	91.969033	24.298981	43.790000	75.170000	86.240000	104.100000	188.500000
Area mean	569.000000	654.889104	351.914129	143.500000	420.300000	551.100000	782.700000	2501.000000
Smoothness mean	569.000000	0.096360	0.014064	0.052630	0.086370	0.095870	0.105300	0.163400
Compactness mean	569.000000	0.104341	0.052813	0.019380	0.064920	0.092630	0.130400	0.345400
Concavity mean	569.000000	0.088799	0.079720	0.000000	0.029560	0.061540	0.130700	0.426800
Concave points mean	569.000000	0.048919	0.038803	0.000000	0.020310	0.033500	0.074000	0.201200
Symmetry mean	569.000000	0.181162	0.027414	0.106000	0.161900	0.179200	0.195700	0.304000
Fractal dimension mean	569.000000	0.062798	0.007060	0.049960	0.057700	0.061540	0.066120	0.097440
Radius se	569.000000	0.405172	0.277313	0.111500	0.232400	0.324200	0.478900	2.873000
Texture se	569.000000	1.216853	0.551648	0.360200	0.833900	1.108000	1.474000	4.885000
Perimeter se	569.000000	2.866059	2.021855	0.757000	1.606000	2.287000	3.357000	21.980000
Area se	569.000000	40.337079	45.491006	6.802000	17.850000	24.530000	45.190000	542.200000
Smoothness se	569.000000	0.007041	0.003003	0.001713	0.005169	0.006380	0.008146	0.031130
Compactness se	569.000000	0.025478	0.017908	0.002252	0.013080	0.020450	0.032450	0.135400
Concavity se	569.000000	0.031894	0.030186	0.000000	0.015090	0.025890	0.042050	0.396000
Concave points se	569.000000	0.011796	0.006170	0.000000	0.007638	0.010930	0.014710	0.052790
Symmetry se	569.000000	0.020542	0.008266	0.007882	0.015160	0.018730	0.023480	0.078950
Fractal dimension se	569.000000	0.003795	0.002646	0.000895	0.002248	0.003187	0.004558	0.029840
Radius worst	569.000000	16.269190	4.833242	7.930000	13.010000	14.970000	18.790000	36.040000
Texture worst	569.000000	25.677223	6.146258	12.020000	21.080000	25.410000	29.720000	49.540000
Perimeter worst	569.000000	107.261213	33.602542	50.410000	84.110000	97.660000	125.400000	251.200000
Area worst	569.000000	880.583128	569.356993	185.200000	515.300000	686.500000	1084.000000	4254.000000
Smoothness worst	569.000000	0.132369	0.022832	0.071170	0.116600	0.131300	0.146000	0.222600
Compactness worst	569.000000	0.254265	0.157336	0.027290	0.147200	0.211900	0.339100	1.058000
Concavity worst	569.000000	0.272188	0.208624	0.000000	0.114500	0.226700	0.382900	1.252000
Concave points worst	569.000000	0.114606	0.065732	0.000000	0.064930	0.099930	0.161400	0.291000
Symmetry worst	569.000000	0.290076	0.061867	0.156500	0.250400	0.282200	0.317900	0.663800
Fractal dimension worst	569.000000	0.083946	0.018061	0.055040	0.071460	0.080040	0.092080	0.207500

4.4. Exploratory data analysis

The EDA process [13] begins with summarizing the database, providing basic statistics and identifying data types, figuring out the size and structure of the database. A missing values analysis follows, which allows us to assess the extent of missing data and decide on appropriate techniques to be applied in order to ensure the completeness of the modeling data. Descriptive statistics and data distribution visualizations reveal the statistical properties of numerical features, making it easier to identify outliers and understand the distribution of features.

Correlation analysis reveals the relationships between features and allows us to measure multicollinearity between them. Particular attention is paid to the distribution of the target variable, understanding the prevalence of breast cancer cases and its impact on the class imbalance of the model.

During EDA, data visualization plays a key role in discovering relationships, identifying outliers, and gaining insight into potential feature engineering opportunities. The insights gained from the comprehensive EDA, feature selection, and model development steps indicate that the breast cancer prediction ML model is built on a solid understanding of the complexity of the database. Ethical considerations when handling sensitive medical data need to be addressed to address issues of privacy and bias, contributing to responsible AI applications in healthcare. Figure 1 shows a flowchart of ML algorithms for breast cancer prediction.

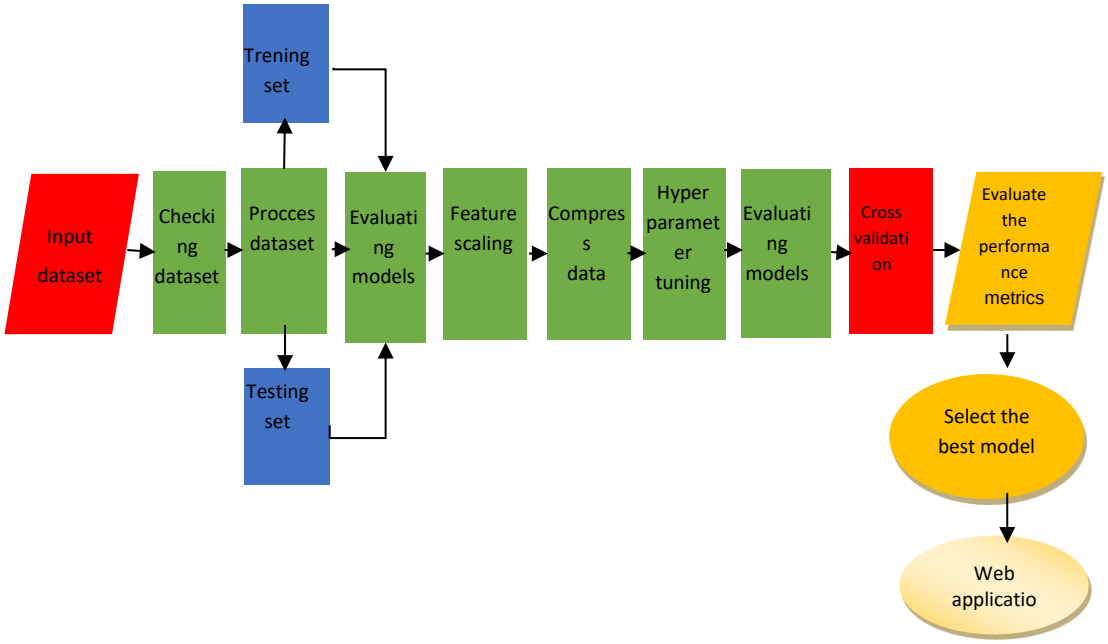


Figure 1: Flowchart of ML algorithms for breast cancer prediction

A correlation was made in order to reveal the potential relationships between the variables so that interactions and dependencies could be seen. This analysis provides valuable insights into the potential influence between variables as well as their joint influence on breast cancer risk. Figure 2 shows a correlation matrix.

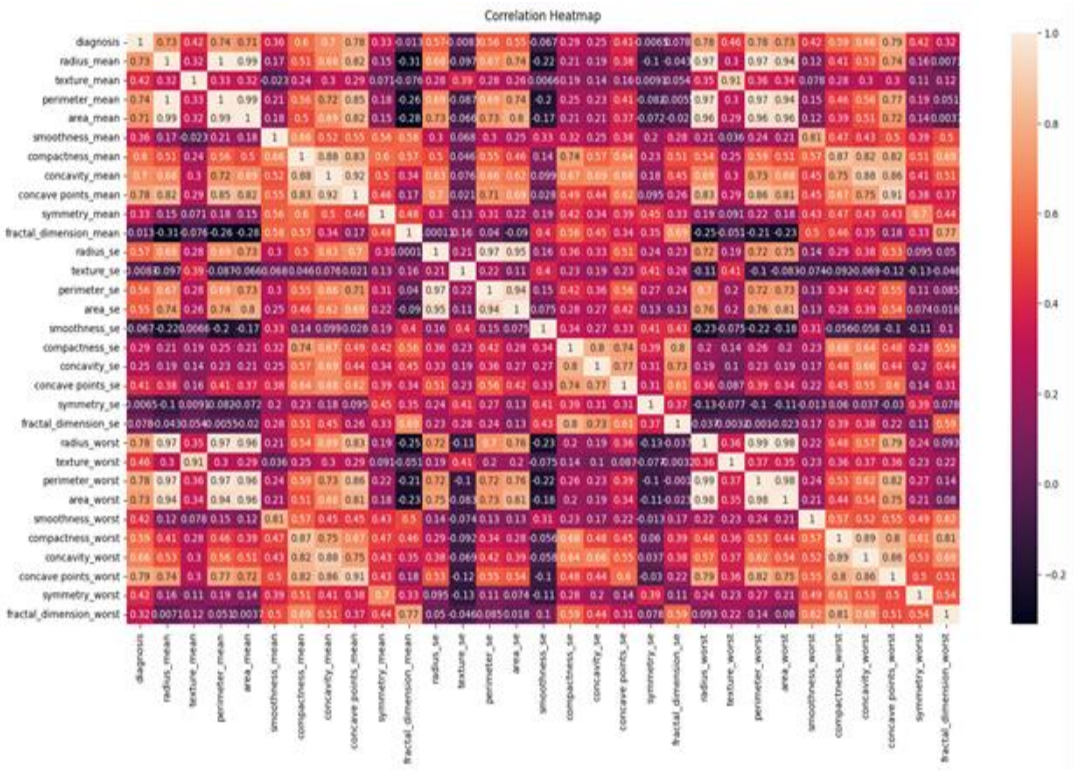


Figure 2: Correlation matrix

4.5. Evaluation Metrics

Using appropriate metrics [14] to evaluate machine learning models aims to determine the best algorithm for a specific task according to the objectives of the problem. The choice of evaluation metrics can vary depending on the machine learning model chosen. An overview of the most commonly used metrics for evaluating classification and regression tasks is provided:

Accuracy: Measures how many of the total number of instances in the test set were correctly classified as true. Accuracy = (TP + TN) / (TP + TN + FP + FN), where TP (True Positives) is the number of true positives, TN (True Negatives) is the number of true negatives, FP (False Positives) is the number of false positives, and FN (False Negatives) is the number of false negatives.

Precision: Measures how many instances classified as positive result in true positives. Accuracy = TP / (TP + FP).

Sensitivity (Sensitivity, Recall): Measures how many of the true positive instances were identified as positive. Recall = TP / (TP + FN).

F1-Score: A joint measure of precision and recall that measures a harmonic average between them. F1-score = 2 * (Precision * Recall) / (Precision + Recall).

Area under the receiver operating characteristic curve (AUC-ROC): Measures the performance of the model for different decision threshold levels. In other words, it is the probability that the model will give an earlier score to a randomly selected positive instance than to a randomly selected negative instance.

Area under the precision-recall curve (AUC-PR): Measures model performance for different decision threshold levels, but focuses on precision and recall. In other words, it is the probability that the model will give an earlier score to a randomly selected positive instance than to a randomly selected negative instance.

Confusion Matrix: A table showing the number of true and false positive and negative classifications of the model. It can be used to calculate all of the above metrics.

The choice of evaluation metric(s) should align with the specific goals of your machine learning project. It is advisable to consider multiple metrics to get a comprehensive overview of model performance.

4.6. Comparison of results

Table 3 shows the Evaluation Scores of the considered ML algorithms by evaluating their performance in terms of accuracy, Cross Validation Score, and ROC_AUC Score in order to build a model for breast cancer risk prediction.

Table 3:
Accuracy, f1 score, precision, recall, balanced accuracy in different ML models

ML model	accuracy	f1 score	precision	recall	balanced accuracy
K-Nears Neighbors	0.973684	0.963855	1.000000	0.930233	0.965116
Naïve Bayes	0.973684	0.963855	1.000000	0.930233	0.965116
Logistic Regression	0.964912	0.952381	0.975610	0.930233	0.958074
Random Forest	0.964912	0.952381	0.975610	0.930233	0.958074
XG Boost	0.964912	0.952381	0.975610	0.930233	0.958074
Neural Network	0.956140	0.941176	0.952381	0.930233	0.951032
Decision Tree	0.929825	0.909091	0.888889	0.930233	0.929905

K-Nears Neighbors and Naïve Bayes: Both models achieved the highest accuracy (97.37%) and F1 score (0.9639) among the models. They also have perfect precision (1.000) and high recall (0.9302), indicating a good balance between correctly classifying positive examples and minimizing false positives. Their balanced precision is also the highest at 0.9651, indicating good performance in both classes.

Logistic Regression, Random Forest, and XG Boost: These three models have similar performance indicators, with an accuracy of 96.49%, an F1 score of 0.9524, and a precision of 0.9756.

Their recall is also 0.9302, indicating that they are efficient at finding positive examples. The balanced accuracy is slightly lower than K-Nearest Neighbors and Naïve Bayes at 0.9581, but it is still a strong performer.

The Neural Network model has slightly lower accuracy (95.61%) and F1 score (0.9412) compared to previous models. Its accuracy and recoil are similar to previous models, both at 0.9524 and 0.9302, respectively. The balanced precision is 0.9510, which is still a good performance.

The Decision Tree model has the lowest accuracy (92.98%) and F1 score (0.9091) among the models.

Its precision (0.8889) is also lower, indicating a higher rate of false positives. However, its recall is still 0.9302, suggesting that it is effective in finding positive examples. Balanced accuracy (0.9299) is the lowest among the models, indicating slightly lower performance of both classes.

In summary, K-Nearest Neighbors and Naïve Bayes stand out as the best-performing models for this classification task, followed by logistic regression, random forest, and XG Boost. The neural network performs slightly below these models but is still a strong contender. The Decision Tree model lags behind in terms of performance. Choosing the best model should consider factors beyond these metrics, such as model complexity, interpretability, and computational requirements.

In this study, we attempted to predict breast cancer risk using machine learning methods and medical data from Kaggle. This study emphasized the importance of data quality and the application of statistical and machine analyses to identify important correlations in medical data. Evaluation of the models with

different metrics provided information about their effectiveness and reliability. It should be kept in mind that this study is only one step towards improving the diagnosis of breast cancer and that continuous research and innovation is needed.

5. Limitations and challenges

Machine learning algorithms in healthcare face numerous limitations and challenges, related to the quality and availability of healthcare data, potential biases in data collection leading to biased models and health disparities, and the interpretability of models. Ethical concerns and privacy regulations present significant obstacles in managing patient data while ensuring confidentiality.

Imbalanced data, temporal aspects of health data, resource constraints, regulatory compliance, and the need to continuously adapt the model to evolving health environments are significant obstacles. The integration of machine learning into existing clinical workflows can be complex and may require changes to established practices, ensuring model validation and accountability is critical in healthcare.

Machine learning (ML) algorithms in the field of healthcare have shown significant progress, but still struggle with significant limitations and face various challenges, such as data quality, data privacy, interoperability, ethical issues, clinical validation, generalizability, model explanation, resource constraints, regulatory hurdles, lack of standardization, human oversight, barriers to adoption, cost-benefit analysis.

Addressing these challenges requires a collaborative, multidisciplinary approach involving healthcare professionals, data scientists, ethicists, policymakers, and various stakeholders to develop robust, ethical, and clinically valuable machine learning applications in the healthcare sector.

6. Conclusions and future work

Machine learning (ML) and artificial intelligence (AI) are revolutionizing healthcare by automating diagnostic processes, predicting disease risks, and personalizing treatment plans based on patient data. These technologies improve medical image analysis, support healthcare professionals in decision-making, and streamline healthcare data management. Although they offer significant benefits, their ethical use and privacy of patient data remain important considerations in their implementation.

Machine learning has made great strides in the healthcare field, offering great potential to improve patient care, diagnostics, and treatment, ensuring quality and standardization of healthcare data, which are critical to training accurate and reliable ML models. This implies a dedicated effort in data collection, cleaning, and harmonization to obtain the desired results.

Ethical dilemmas and privacy are significant aspects, requiring continued research into privacy-preserving techniques and ethical frameworks aimed at protecting patient data while enabling meaningful analysis.

Interpretability and explain ability of ML models are of critical importance, especially in the healthcare decision-making process, continuous research on interpretable AI methodologies and model explanation is needed. Mitigating bias in health data and algorithms to promote fairness, the development of unbiased models becomes necessary to ensure equitable health care outcomes.

Validation studies are necessary to confirm the safety and efficacy of ML models in clinical settings. Collaboration between researchers, clinicians, and regulatory bodies is essential.

ML models must adapt to new data and protocols in healthcare. Research devoted to adaptive and lifelong learning approaches becomes imperative.

Integrating ML into clinical workflows is a huge challenge, but it is still essential that research should concentrate on developing user interfaces and workflows that healthcare professionals can easily adopt.

Effective allocation of resources in resource-constrained healthcare facilities requires careful consideration.

In the future, priority should be given to patient-focused AI applications that enable individuals to make informed decisions about their health.

Interdisciplinary collaboration among data scientists, healthcare professionals, ethicists, policymakers, and other stakeholders remains necessary for the successful development, validation, and deployment of ML algorithms in healthcare.

The future calls for continued advances in AI and ML in the healthcare sector, with an increased focus on personalized medicine, early disease detection and improving patient outcomes. Addressing these challenges will be critical to realizing the full potential of ML in healthcare.

References:

[1] Kirola, Madhu, et al. "A comprehensive review study on: optimized data mining, machine learning and deep learning techniques for breast cancer prediction in big data context." Biomedical and Pharmacology Journal 15.1 (2022): 13-25.

[2] Thagaard, Jeppe, et al. "Pitfalls in machine learning-based assessment of tumor-infiltrating lymphocytes in breast cancer: a report of the international immuno-oncology biomarker working group." The Journal of Pathology (2023).

[3] Ozaydin B, Berner ES, Cimino JJ. Appropriate use of machine learning in healthcare. Intelligence-Based Medicine. 2021 Jan 1; 5: 100041.

[4] Gou, Qiheng, et al. "Systematic evaluation of tumor microenvironment and construction of a machine learning model to predict prognosis and immunotherapy efficacy in triple-negative breast cancer based on data mining and sequencing validation." Frontiers in Pharmacology 13 (2022): 995555.

[5] Clark, Phoebe, Jayne Kim, and Yindalon Aphinyanaphongs. "Marketing and US Food and Drug Administration Clearance of Artificial Intelligence and Machine Learning Enabled Software in and as Medical Devices: A Systematic Review." JAMA Network Open 6.7 (2023): e2321792-e2321792.

[6] Yue, Wenbin, et al. "Machine learning with applications in breast cancer diagnosis and prognosis." Designs 2.2 (2018): 13.

[7] Amrane, Meriem, et al. "Breast cancer classification using machine learning." 2018 Electric Electronics, computer science, biomedical engineering meeting (EBBT). IEEE, 2018.

[8] Asri, Hiba, et al. "Using machine learning algorithms for breast cancer risk prediction and diagnosis." Procedia Computer Science 83 (2016): 1064-1069.

[9] Shalini, M., and S. Radhika. "Machine learning techniques for prediction from various breast cancer datasets." 2020 Sixth international conference on biosignals, images, and instrumentation (ICBSII). IEEE, 2020.

[10] Christobel, Angeline, and Y. Sivaprakasam. "An empirical comparison of data mining classification methods." International Journal of Computer Information Systems 3.2 (2011): 24-28.

[11] Baughan, Natalie, Lindsay Douglas, and Maryellen L. Giger. "Past, present, and future of machine learning and artificial intelligence for breast cancer screening." Journal of Breast Imaging 4.5 (2022): 451-459.

[12] Bari Antor, Morshedul, et al. "A comparative analysis of machine learning algorithms to predict Alzheimer's disease." Journal of Healthcare Engineering 2021 (2021).

[13] Indrakumari, R., T. Poongodi, and Soumya Ranjan Jena. "Heart disease prediction using exploratory data analysis." Procedia Computer Science 173 (2020): 130-139.

[14] Li, Yixuan, and Zixuan Chen. "Performance evaluation of machine learning methods for breast cancer prediction." Appl Comput Math 7.4 (2018): 212-216.

[15] Ahmad, L. Ghasem, et al. "Using three machine learning techniques for predicting breast cancer recurrence." J Health Med Inform 4.124 (2013): 3.

[16] Sivakami, K., and Nadar Saraswathi. "Mining big data: breast cancer prediction using DT-SVM hybrid model." International Journal of Scientific Engineering and Applied Science (IJSEAS) 1.5 (2015): 418-429.

[17] Men, Kuo, et al. "Fully automatic and robust segmentation of the clinical target volume for radiotherapy of breast cancer using big data and deep learning." Physica Medica 50 (2018): 13-19.

[18] Rane, Nikita, et al. "Breast cancer classification and prediction using machine learning." International Journal of Engineering Research and Technology 9.2 (2020): 576-580.

[19] Mahesh, T. R., et al. "Performance analysis of xgboost ensemble methods for survivability with the classification of breast cancer." Journal of Sensors 2022 (2022): 1-8.
[20] Saoud, Hajar, et al. "Using feature selection techniques to improve the accuracy of breast cancer classification." Innovations in Smart Cities Applications Edition 2: The Proceedings of the Third International Conference on Smart City Applications. Springer International Publishing, 2019.
[21] Abdulaziz, Meliboyev. "ACCURATE BREAST CANCER CLASSIFICATION BY USING ARTIFICIAL INTELLIGENCE ALGORITHMS." Journal of Universal Science Research 1.5 (2023): 1579-1586.
[22] Baraniya, S., et al. "Breast Cancer Classification and Recurrence Prediction Using Artificial Neural Networks and Machine Learning Techniques." 2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT). IEEE, 2023.
[23] Petrovski, G., Savoska, S., Ristevski, B., Bocevska, A., Jolevski, I., & Blazheska-Tabakovska, N. (2022). Visual Data Analysis for EU Public Sector Data using the app MyDataApp.
[24] Ristevski, Blagoj, Snezana Savoska, and Natasha Blazheska-Tabakovska. "Opportunities for big data analytics in healthcare information systems development for decision support." (2020): 38-50.
[25] https://www.kaggle.com/datasets/nancyalaswad90/breast-cancer-dataset?select=data.csv

Future Challenges for Object Detection and Image Recognition Techniques

Buen Bajrami[1], Kostandina Veljanovska[2] and Zoran Kotevski[3]

[123] *Faculty of Information and Communication Technologies, University St. Kliment Ohridski – Bitola, 1 Maj bb., 7000 Bitola, North Macedonia*

bajrami.buen@uklo.edu.mk, kostandina.veljanovska@uklo.edu.mk, zoran.kotevski@uklo.edu.mk

Abstract:
Object detection and object recognition are the two main techniques that enable the identification of objects in images and videos. These techniques utilize Deep Learning and Machine Learning by simulating human-like image recognition. There are various methods developed recently, which have made significant progress in object detection or image recognition, and therefore, through this work, we present the latest advances in this field and their operation through Machine Learning and Deep Learning. We discuss the challenges that appear in object detection and the necessary improvements that must be made in order to increase the efficiency of this technology. Furthermore, the future challenges for object recognition has been addressed as well, by providing an analysis on the possibilities of continuous development and facilitation of these techniques in various domains.

Keywords:
Object Detection, Image Recognition, Deep Learning, Machine Learning, Convolutional Neural Networks

1. Introduction

Object detection and image recognition techniques have recently had substantial development in a fair proportion to the advancement of technology in the world. Many Machine Learning and Computer Graphics researchers have intensified their work in improving and advancing these techniques. Object detection has found high application in video games, in automotive industry, in medical devices, in traffic safety etc. On the other hand, face recognition is increasingly becoming an authentication paradigm, finding application in a large number of mobile applications that require face recognition, as well as in biometric applications. These vast application possibilities present an urgent requirement to increase security, due to the rapid development of digitalization. In addition to these two techniques, we will also focus on some other paradigms, to analyze their implementation progress, as well as to see the shortcomings and sensitive points where we estimate that there is still a need for improvements in the future.

The rest of this paper is organized as follows. In section 2we first make a review of the literature to give a right orientation to our analysis. We emphasize the importance of this technology and its development so far. Next, in section 3 and section 4 we provide separate analyzes for both technologies, i.e. object detection techniques and image recognition techniques and the challenges and issues in their further development. Finally, in Section 5, we give a recommendation about what issues developers should deal with more specifically, in order to make greater advancement of this field.

2. State-of-the-art

Object detection and image recognition techniques are the two highly used object recognition techniques. In order to achieve a complete understanding of the image, it must be made clear that we have several other tasks to deal with in addition to classifying the different images, i.e. to accurately

assess the concepts and locations of the objects contained in each image. This task is referred to as object detection [1]. After the realization of this first phase, the next phase consists of several other subtasks such as face detection [2], pedestrian detection [3] and skeleton detection [4]. In order to be able to recognize objects better and have have stronger semantics, we need to extract representative features, which is achieved through Scale-Invariant Feature Transform (SIFT) [5], Histogram of Oriented Gradients (HOG) [6] and Haar-Like that considers adjacent rectangular regions at a specific location in a detection window of the image then sums up the pixel intensities in each region and finaly calculates the difference between these sums [7]. HOG is a feature descriptor that is used in computer vision and processing of images for object detection and recognition. FiSIFT as a technique for image matching can identify and match features in images that are invariant to scaling, rotation, and affine distortion and it is widely used in computer vision, image matching, object recognition, and also, 3D reconstruction. We can say that Deep Neural Networks (DNNs), or the more representative Convolutional Neural Networks (CNNs), operate in a completely different manner compared to traditional approaches. CNNs have deeper architectures with the ability to learn more complex features than shallow ones. Also, expressiveness and powerful training algorithms enable us to obtain features of objects without having to draw them by hand [8].

3. Object Detection

When we talk about object detection, we mean the computer's ability to detect objects within a given frame, similar to how humans are able to detect objects within a given image. Object recognition [9] is widely used in machine vision industry for inspection, registration and manipulation tasks. We will cover some of the object detection tasks, for example, objects under different angles, lighting and variations within the classroom and challenges in object detection [10]. We estimate that object detection is divided into two periods. Object detection Based on traditional methods until 2014, and after 2014 we have object detection based on deep learning. We will address the biggest developments, applications of these technologies as well as the challenges these technologies are facing.

A. Object Detection in Computer Vision

Computer vision is extremely important in improving object detection. These algorithms have found high application in medicine, such as in detection of tumors, in detection of potential anomalies of children yet to be born etc. Other implementations of object detection include security systems with face detection, traffic management with car license plates recognition and implementation in automotive industry where high precision detectors have been developed to detect objects that may have a direct or indirect impact on the moving vehicle. Of course, object detection is used in a very large number of everyday-life fields as well, such as ensuring that people on the "No Fly" list do not pass through security gates at airports, animal monitoring in agricultural farms and zoos, detection of roads, pedestrians and traffic lights in autonomous vehicles, scanning and verification of faces against passports at airports, detection of health abnormalities etc.

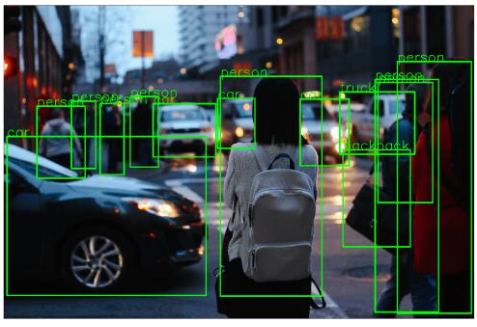


Figure 1: R-CNN Object Detection

In recent years, object detection algorithms have been further enhanced by deep learning, increasing processing speed, as well as providing more accurate results. Some of the most popular deep learning

architectures in object detection are known as: Fast R-CNN, Yolo, SSD, R-CNN. The Fast Region-based Convolutional Neural Network method (Fast R-CNN) works as a CNN (usually pre-trained on the ImageNet classification task) with its final pooling layer replaced by a region-of-interest pooling (ROI pooling) layer and its final fully connected layer (FC) (that operates on a flattened input where each input is connected to all neurons) is replaced by two branches: a (K + 1) category softmax layer branch and a category-specific bounding box regression branch. YOLO (You Only Look Once) deep learning architecture as a popular object detection algorithm has made a revolution in the field of computer vision with its characteristics as fast, efficient and easily applicable in real-time object detection tasks. Single Shot Detectors architecture (SSD) is also a popular and efficient in object detection. SSD uses a single CNN in order to predict bounding boxes and class labels for objects in an image, and that makes it faster and more efficient than other methods. For the last architecture, regions with convolutional neural networks (R-CNN) it is a technique where objects are detected in an image by combining rectangular region proposals with convolutional neural network features.

B. Object Detection Algorithms and Detectors

Below is a list of some of the most important developments in object detection technology:

a) The detectors of Viola and Jones [11], where in 2001 they achieved for the first time in real time the detection of human faces without any limitations. This VJ detector was hundreds of times faster than the algorithms of that time.
b) HOG detector, N. Dalal and B. Triggs developed this detector during 2005 where a significant improvement of scale invariant feature transformation was achieved [12] and shaping contexts of his time [13].

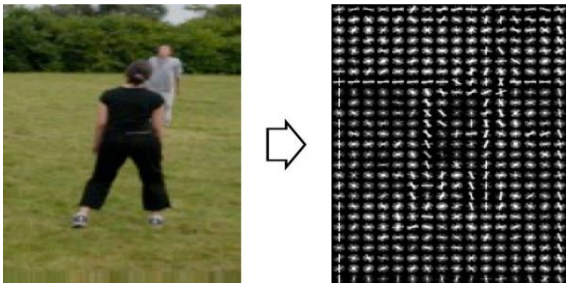


Figure 2: Input Image & Visualization of the HOG features

c) R-CNN, makes the extraction of a set of proposed objects through a selective search. Then all of these propositions are converted into still images and fed into a CNN model, so that features can be extracted from them [14].
d) SPPNet model is an improvement over CNN models, as they require input of fixed size, while SPPNet does not depend on fixed dimensions of input images. It also has an object detection speed about 20 times faster than R-CNN.
e) Fast R-CNN, is a new model developed in 2015. Which is an improvement over R-CNN and SPPNet models, where it has an object detection speed of about 200 times greater than R-CNN. This model allows to simultaneously handle a detector and a regressor within the same link.
f) FPN algorithm, developed during 2017 and proposed by T.-Y.Lin [15], has increased accuracy in locating objects in contrast to the above models.

C. Object Detection Challenges

Throughout all these developments, Object detection has also had challenges, some of which have been improved while others still need to be worked on in the future. Below we list some of the key challenges for which we believe more can be done.

a) Disadvantages of RCNN are excessive feature calculations on a large number of overlapping propositions (over 2000 boxes from one image) lead to an extremely slow detection speed (14 seconds per image with GPU)
b) SPPNet still has some shortcomings, such as the training is still performed in multi-stage and it only adjusts its fully connected layers, while simply ignoring all previous layers.
c) Deformation. In many cases object detection detectors are trained for solid objects and have difficulty detecting fluid or very flexible objects.
d) Occlusion. Object detection detectors also encounter difficulties when asked to find an object, where part of it is covered by another object. For humans this is not a problem.
e) Illumination conditions. In different illuminations the detectors assign different features to the same object and as a result several different objects are detected.

4. Image Recognition Techniques

Image recognition is a machine vision, which has the ability to recognize objects, images, people, etc., knowing how to distinguish and compare the same features in different images. Image recognition requires many processes, which require high processing speed, then accuracy in recognition, as well as small tolerance in the obtained results. Below we have made an analysis about image recognition through deep learning and machine learning. We have also shown the most frequent applications of this technology and finally the challenges that this technology faces.

It is known that image recognition has found a very large application in practice, where today it is used by a large part of the industry, security systems, smart phones, etc. Some of the applications of this technology include:

a) Medical diagnosis, i.e. in health, image recognition is used quite a lot, especially in software that is integrated into devices through which many recordings are made, such as X-Ray devices, etc., where through advanced algorithms it is possible to recognize abnormalities in patients.

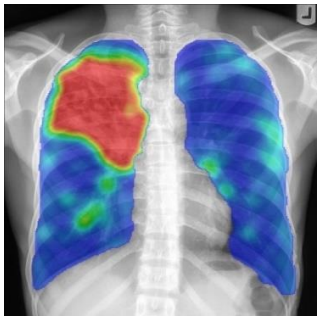


Figure 3: Detection anomalies in X-Ray with CNN

b) People identification. In these cases the security institutions use special equipment that enables people to be recognized through the collected photographs and videos.
c) Fingerprint recognition is one of the simplest implementations based on small features such as fingerprints-impressions.
d) Face recognition [16] is a technology that is mostly used in smart phones, in access control, and recently in cars.
e) Visual search is a search method that allows people to search for something via input image, as Google Lens tool makes searches.
f) Iris scan recognition. After DNA, iris is the second most unique organ in the human body and contains more information about the human than fingerprint impressions [17]. Therefore, recently we have technology developments where iris recognition is used as authentication

instead of fingerprint or face recognition. This technology has started to be used also in ATM devices.

Figure 4: ATM authentication with Iris recognition

A. Image Recognition with Machine Learning & Deep Learning

Deep learning [18] as a new field, comes from machine learning which aims to build a neural network, which is capable of analyzing data and learning similarly to humans, through advanced algorithms. Deep learning has boosted image recognition in a rapid manner. Mainly, traditional models have applied color for image recognition, as well as features of image shape and structure [19]. One of the main concepts in image recognition through deep learning is Convolutional Neural Networks (CNN). This concept consists of several smaller layers of neurons that all contain parts of the image, and after image processing, all parts are superimposed in one place. And so the layer below repeats this process, learning more about the composition of the image. CNN [20] brought a revolution in the field of computer vision by increasing the accuracy of image classification, but also many times improved scene classification, object detection, semantic segmentation of biological images and face detection, text recognition and human body recognition in natural images. The main practical success of CNN is the face recognition and autonomous driving of cars. Also, we have Multi-Layer Perceptron (MLP) with a feed forward learning algorithms. MLP is a frequent choice because of its simplicity and its capability in supervised pattern matching. It has been successfully applied to many pattern classification problems [21, 22]. Linear discriminant analysis (LDA) is a powerful method for face recognition. This model gives effective representation that linearly transforms the original data space into a low-dimensional feature space. A subspace analysis method for face recognition called kernel discrimination locality preserving projections (MMDLPP) was proposed in [21] based on LDA analysis.

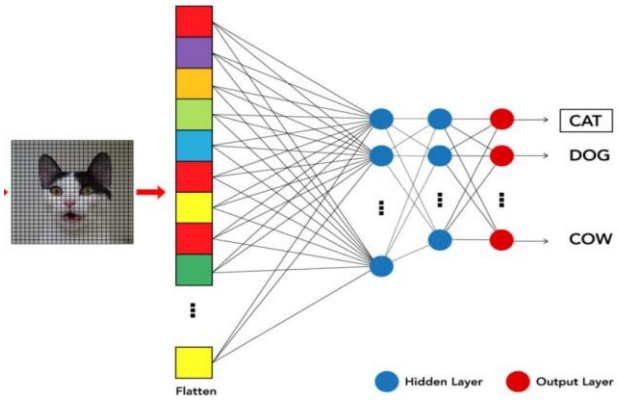


Figure 5: Image recognition using CNN

B. Image Recognition Challenges

After analyzing the techniques for image recognition, we estimate that we have found some shortcomings, which we have listed below.

a) CNN requires a large database which helps in image detection, comparison and recognition, but consequently comes with a higher cost.

b) False image recognition. This happens in many cases, for example, when the algorithm recognizes another person who has some features similar to the person who is really being searched for.
c) Racial bias. This is a strange issue, but it is worth noting that image recognition algorithms distinguish black people more easily than white people.
d) There are challenges in protecting personal data.

5. Conclusion

After evaluating the results obtained from the literature review, we estimate that a remarkable amount of work has been done in the last two decades. Both object detection and image recognition technologies have found a great application in practice, greatly facilitating some processes that were quite difficult until recently. There is a growing trend of organizations dealing with the development of complex algorithms, which will continue to facilitate this process. In general, the advantages of using these technologies are significantly greater than the shortcomings or drawbacks. Where, after the analysis of the results found, we recommend to work more with the image recognition technology, which has greater challenges due to the various factors that we mentioned above. We recommend that more work be done in the future in the field of health, where there is great potential for this very important sector for people to shine through the advanced use of these technologies. Also, one of the most sensitive sectors is undoubtedly privacy, the protection of personal data, since through the use of image recognition we have a high exposure of our personal images as well as our sensitive data. Therefore, more work is required in terms of protecting the data, code them, in order to be more confident in the use of various applications related to object detection and image recognition.

References:

[1] Felzenszwalb, D. et al. “Object detection with discriminatively trained part-based models”, IEEE Trans. Pattern Anal. Mach. Intell., pp. vol. 32, no. 9, p. 1627, 2010.
[2] Poggio, T. et al. “Example-based learning for view-based human face detection”, IEEE Trans. Pattern Anal. Mach. Intell, pp. vol. 20, no. 1, pp. 39–51, 2002.
[3] Dollár, P. et al., “Pedestrian detection: An evaluation of the state of the art”, IEEE transactions on pattern analysis and machine intelligence, pp. vol. 34, no. 4, p. 743, 2011.
[4] Kobatake H, Yoshinaga Y., “Detection of spicules on mammogram based on skeleton analysis”, IEEE Trans. Med. Imaging., pp. vol. 15, no. 3, pp. 235–245, 1996.
[5] Lowe, D. G., “Distinctive image features from scale-invariant key points”, Int. J. of Computer Vision, pp. vol. 60, no. 2, pp. 91–110, 2004.
[6] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection”, CVPR, 2005.
[7] R. Lienhart and J. Maydt, “An extended set of haar-like features for rapid object detection,” ICIP, 2002.
[8] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning”, Nature, pp. vol. 521, no. 7553, pp. 436–444, 2015.
[9] A. B. Amjoud and M. Amrouch, "Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review," *in* IEEE Access, *vol*. 11, pp. 35479-35516, 2023.
[10] P. Viola and M. Jones, “Robust real-time face detection,” Int. J. of Computer Vision, vol. 57, no. 2, pp. 137–154, 2004.
[11] K. C. Z. S. Zhengxia Zou, "Object Detection in 20 Years: A Survey," Fellow, IEEE, p. arXiv:1905.05055v3, 2023.
[12] D. G. Lowe, “Object recognition from local scale invariant features”, ICCV, pp.1150–1157, 1999.

[13] S. Belongie, J. Malik and J. Puzicha, “Shape matching and object recognition using shape contexts*”,* IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 4, pp. 509-522, April 2002.

[14] Zhong-Qiu Zhao, et al. “Object Detection with Deep Learning: A Review”, IEEE Transactions on Neural Networks and Learning Systems, April, 2019.

[15] Lin, T.Y. et al. “Feature pyramid networks for object detection”, CVPR, vol. 1, no. 2, p. 4, 2017.

[16] Gu, X., et al. Regularized Kernel Locality Preserving Discriminant Analysis for Face Recognition, Advanced Concepts for Intelligent Vision Systems. ACIVS 2010. Lecture Notes in Computer Science, vol 6475. Springer, Berlin, Heidelberg, 2010.

[17] M. C. Mayur Gadge, "Image Processing for Secure ATM", International Journal of Innovations in Engineering and Science, Vols. Vol. 7, No. 8, no. 2456-3463, pp. 132-136, 2022.

[18] K. Xu, J. Chen, Y. Ning and W. Tang, "Deep Learning in Image Classification: An Overview," *2023 4th International Conference on Computer Vision, Image and Deep Learning (CVIDL)*, Zhuhai, China, 2023, pp. 81-93, doi: 10.1109/CVIDL58838.2023.10167276.

[19] D. Al-obidi and S. Kacmaz, "Facial Features Recognition Based on Their Shape and Color Using YOLOv8," 2023 7th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkiye, pp. 1-6, 2023.

[20] K. Veljanovska, "Artificial Intelligence Implemented in Covid-19 Detection", 11th International Conference on Applied Information and Internet Technologies - AIIT, pp. 2-6, 2021.

[21] Y. Guan, Y. Han and S. Liu, "Deep Learning Approaches for Image Classification Techniques," 2022 IEEE International Conference on Electrical Engineering, Big Data and Algorithms (EEBDA), Changchun, China, pp. 1132-1136, 2022.

[22] Xiaoguang Li and S. Areibi, "A hardware/software co-design approach for face recognition," Proceedings. The 16th International Conference on Microelectronics, 2004. ICM 2004., Tunis, Tunisia, pp. 55-58, 2004.

An Example of Application for Custom Design Automation using SolidWorks Application Programming Interface

Hristijan Stojceski[1], Andrijana Bocevska[2], Igor Nedelkovski[3] and Nikola Rendevski[4]

[1,2,3,4] *University St Kliment Ohridski, 1 Maj nn, 7000 Bitola, R. North Macedonia*

hristijan.stojceski@uklo.edu.mk; andrijana.bocevska@uklo.edu.mk; igor.nedelkovski@uklo.edu.mk nikola.rendevski@uklo.edu.mk

Abstract:
In this paper an application which allows users with minimal CAD knowledge to interact with the software and make custom design automation is presented. The user-interface was developed using the Visual Basic for applications (VBA) programming language and the CAD software's built in application programming interface (API) command structure. This process allows the automation of selected embedded CAD productivity tools whereby the user is able to modify specific parameters to manipulate the shape of their design and automatically generate a 3D computer-aided model reflecting their specific modifications. This paper utilizes a solid body of four types of tool designs as an example. We describe the creation of the user interface, as well as specific limitation ranges and/or constraints that were placed on the parameters of the designs. The result was the development of an interactive design template that prevents design engineers from needing to design the same model's family repeatedly to suit the needs of the customer, thereby reducing the design time and mistakes while enhancing the design consistency.

Keywords:
Custom design automation, SolidWorks API, VBA

1. Introduction

With today's emerging markets and product variety, it is very important for industrial companies to explore product customization to capture customer attention and deliver true customer value. Product customization refers to enabling customers to personalize a product according to their needs and preferences [1]. Add-ons, exclusive functionalities, templates, and flexibility with product design all count as different forms of personalization.

Product customization is essential because it allows customers to feel like a part of the product development process. They can choose the features they want, the color they want, and even the style of the product. In addition, product customization helps businesses learn more about their customers. By understanding what customers want and need, businesses can create products that are more likely to sell. Customization also helps companies target specific markets.

A few of the benefits that businesses can enjoy when they offer product customization [2]:

Because Customers Like It - Product customization adds a personal touch to the product. When customers design it their way, the end product does not just remain a product but a feeling gets attached to it no matter whether they are doing it for their own or for someone else. Since customers like the personalization concept, it provides brands with an excellent opportunity to excel in the e-commerce market.

Product Customization builds Customer Loyalty - Customer loyalty is one of the significant benefits of product personalization. And there's no better way to earn a customer's loyalty than to give them complete control over the product's design. Customer success is the most crucial factor in gaining that loyalty [3]. Product customization creates a bond between customers and brands and significantly increases customer retention [4].

Higher Sales - Customers are willing to spend more for a personalized item because they regard it as more than simply a product; it is something unique to them. This can be backed up with a Deloitte

study claiming that one out of every five consumers would be willing to spend 20% more on an exclusive product [5]. Also, as per Invesp, 59% of the marketers are getting a good ROI after offering the product personalization facility [6].

Word-of-Mouth Marketing - According to the Harvard Business Review, buyers who can customize their products buy more and are more inclined to suggest the brand to their friends [7]. People tend to notice something that stands out. If customers are loyal and satisfied with the products, they'll probably tell others about them, resulting in more people buying them.

Get Better Insights into your Customers - Adding to the benefits of product customization is that a personalized purchase provides a more detailed analysis of your customer's preferences, likings, and taste than a normal purchase. When customers purchase customized products, brands can collect that information and utilize it to offer products based on their previous preferences. Efficient research and development lead to a competitive advantage [8].

Product Customization module by KnowBand - Knowband offers a Product Customization module for offering customization services to the customers. It adds a customization tab on the product pages and allows the customers to personalize a product as per their desire. From adding the desired images to putting up a customized text and QR code, the Product Designer module comes with a variety of features to help customers personalize the products exactly the way they want.

Customer involvement in new product development is currently a thriving activity implemented by companies in order to fulfill customers' needs [9].

In modern conditions, it is important for any company to reduce the time required to design new products. One way to do this is through the automation of design processes. Automation in SolidWorks is a huge opportunity to speed up and make design process dynamic. Custom applications using the SolidWorks API can be as simple or as complex as desired, for instance, the API can be used to create simple macros.

Therefore, the purpose of this paper was to create an application for producing customized products for customers with poor or no skill in CAD software. This paper focused on developing an interface for the customers to use to modify product appearance by integrating several software applications such as Visual Basic, SolidWorks and Excel. Visual Basic creates the user-interface form and integrates SolidWorks and Excel together using each software's API command structure, whereas SolidWorks for 3D solid model creation and Microsoft Excel are used for customer's parameter values. SolidWorks Macro/API allows the user to automate the designing tasks and eliminate the conventional and repetitive process of designing for different sizes and conditions.

The paper is organized as follows. Previous work is described in Section 2, whereas Section 3 discusses the features of the used technologies. Description and screenshots of the developed application are given in Section 4. Last section gives a brief review of the research, providing concluding remarks and directions for further work.

2. Previous work

Many researchers have used the API of various CAD systems in order to develop applications and tools and always the conclusions represent that the benefits and advantages of the API are huge.

An alternative methodology for real time object customization in a CAD system is suggested in [10]. The proposed program modifies the CAD model through Visual Basic programming and operates by controlling and manipulating the instructions. It realizes a real time customization where it allows users to adjust and change the parameters on the GUI instead of editing the model manually, thereby simultaneously modifying the shapes in the CAD interface.

The author of [11] elaborates how a complete design automation procedure can be made based on examples. The paper considers the exponential growth in design methodologies and depicts works on design automation of a cube on certain parameters. SolidWorks designing and modelling was used for writing the macro code for automation. Parameters like dimensions, colors and materials have been considered for automation. After the conceptual development and algorithm build up, an application has been made in order to make the product user-friendly.

The work presented in [12] uses Visual Basic to develop corrugated box parametric drawing system on SolidWorks platform on the basis of the established corrugated box access database. The users only

need to select the desired corrugated box carton and the key parameters, then can get the 2D expansion diagrams and 3D stereoscopic diagrams in accordance with the national standard, which significantly improve the efficiency of the corrugated box drawing.

A procedure to automate the design of metal scrap balers by integrating Visual Basic, Excel and SolidWorks is presented in [13]. Firstly, a prototype model of the mechanical system under development was designed in SolidWorks by GUI. The geometric dimensions considered for the iterative work were then entered to an Excel file. The data in the Excel file can be edited and changed for the iterative work. The modifying program developed in Visual Basic reads the data from the Excel file and automatically modifies the SolidWorks model using API. The modified model was analyzed and the analysis results were evaluated.

The authors of [14] investigate how design automation can be used and implemented to automate and improve sections of the order-to-delivery process for customized products at an industrial manufacturing company. Their objective was to develop a product configurator to automatically generate 3D models and documentation for production and sales support. This was achieved through the development of a product configurator which decreases ODP lead-time, prevents errors, reduces material waste and adds customer value.

The advantages of the applicable programming interface of CAD system in creating custom programs (macros) that can assist users in automating many of the tasks with the help of simple graphical user interface (GUI) are outlined in [15]. This procedure contributes to reduce design time and eliminate any potential errors. An application was developed using Microsoft Visual Studio and C# programming language and was integrated with SolidWorks CAD package through API. Output design result is passed to SolidWorks CAD package, which updates CAD models of car rim and manufacturing drawing.

3. Features of the used technologies

Key element for creating the application was a thorough understanding of the CAD tools available in SolidWorks to generate multiple configurations. SolidWorks allows the development of part configurations through the use of a design table. A design table uses an Excel spreadsheet to allow a user to enter part feature information to create a new configuration of a product by changing selected part dimensions or by suppressing part features. The Excel spreadsheet can be automatically linked to the SolidWorks software, so when a user edits the data in Excel spreadsheet externally, SolidWorks will create the new model or models based on the entered data. Each software uses a common programming structure called an application programming interface (API) which enables software users to create tools using Visual Basic for applications (VBA) programming language to interact with the software. This interaction could be accomplished using two basic approaches depending on the type of tool being developed and the type of user interaction desired. Those approaches included: VBA, Add-Ins (DLL or EXE), and standalone exe.

There are two ways to make design changes: customize a model from within SolidWorks or customize the model outside of the SolidWorks software. Customizing the model from within the CAD software wouldn't be very helpful since the user would need some level of functional knowledge of the CAD software. For those reasons the standalone exe. application was created as independent program that is capable to control other software and has its own interface. In this regard, Visual Basic software was used for the standalone exe. to interactively work within the SolidWorks and Excel software to customize a model. Once the user has submitted parameter values into the user-interface form, the parameter values will automatically be entered in Microsoft Excel, which in turn updates the SolidWorks 3D model automatically, and the result will be displayed back to the user-interface form for the customer to compare. The role for Visual Basic was to be able to control Microsoft Excel and SolidWorks using their API commands to work internally as if it were part of Visual Basic codes. To automatically enter the customer’s parameter value from the user-interface form into the Excel, Visual Basic must use Excel API commands. Once the model is customized, the picture box will show the updated model with the current values of the customized model. All the textboxes in the user-interface form are protected with error checking codes to make sure that the customer cannot enter any invalid input. The customers are only allowed to enter numeric points into the textbox parametric values. In

addition, the application creates an automatic drawing of the model if the checkbox for creating drawing is selected.

4. Description of the developed application

In the paper, an application is developed using SolidWorks API and VBA programming language. SolidWorks provides an API that allows external programs to interact with and control SolidWorks functionality. The API exposes a set of objects, methods, and properties that developers can use to access and manipulate SolidWorks models, assemblies, drawings, and other elements.

VBA is event driven language, embedded in SolidWorks that enables users to write custom macros and automation scripts. These VBA macros can be used to create custom features, automate repetitive tasks, and enhance the capabilities of SolidWorks.

With the help of VBA, users can create macros that automate various aspects of the design process in SolidWorks. For instance, you can write a VBA macro to generate complex geometric shapes, perform design validations, or automate the creation of drawings and reports. VBA also allows users to customize the SolidWorks user interface by adding custom toolbars, menus, and buttons. This level of customization enables users to have a more tailored and efficient workflow based on their specific needs. Lastly, through VBA, SolidWorks can communicate and exchange data with other Microsoft Office applications, such as Excel and Access. This integration allows for seamless data transfer between the design process and other parts of the workflow.

SolidWorks includes a built-in feature called "Macro Record", which allows users to record macros directly within the software. To use this feature, simply launch SolidWorks and select "Macro Record" from the toolbar. Once the recording is complete, you can stop the macro recording by clicking on "Macro Stop". The recorded macro will then be saved.

The application's primary objective is custom design automation of tools, such as spanners. By leveraging the recorded macros, users can streamline repetitive tasks and efficiently generate designs for various tools, enhancing productivity and reducing manual effort. Figure 1 shows the user interface of the application.

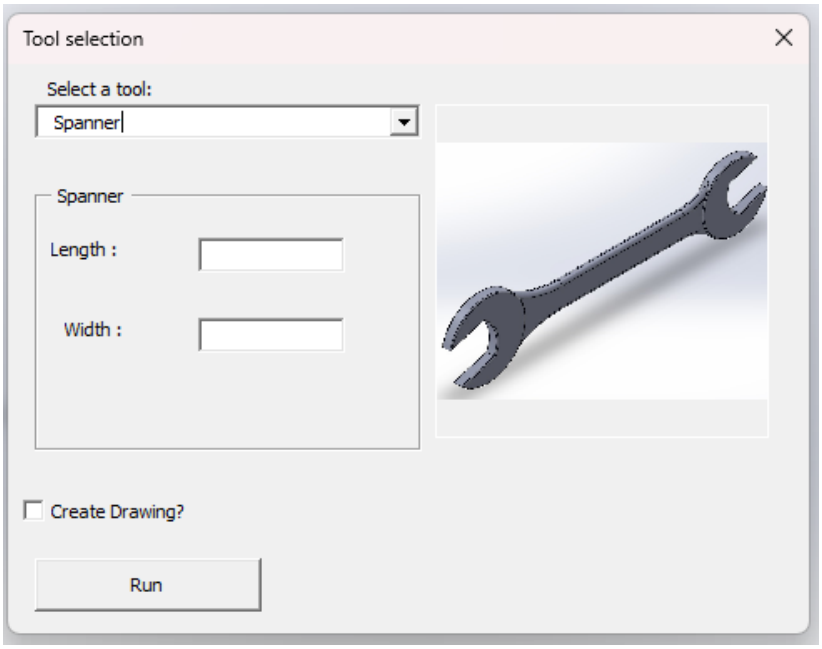


Figure 1: User interface of the application

The proposed graphic interface consists of a drop-down list, two textboxes, a checkbox and a "Run" button. A preview of the model is shown on the right side of the interface. Using the drop-down list, one of the following four tools can be chosen:

- Spanner

- Screwdriver
- Allen key
- Adjustable wrench

By default, spanner is selected, and the tool has couple of parameters that can be modified. Users can input desired tool dimensions through the textboxes, which will be utilized in the modelling process. By clicking the "Run" button, an algorithm executes, generating the tool based on the provided dimensions.

The interface's checkbox triggers the automatic creation of drawings for the newly designed tools. Users can opt to generate these drawings during the modelling process, providing comprehensive documentation for the models. To facilitate future reference and data analysis, an algorithm is implemented to store all tool dimensions in a Microsoft Excel spreadsheet. This feature ensures easy access to and review of the design specifications of previously created tools. The user can choose one out of four models which are available, as shown in Figure 2.

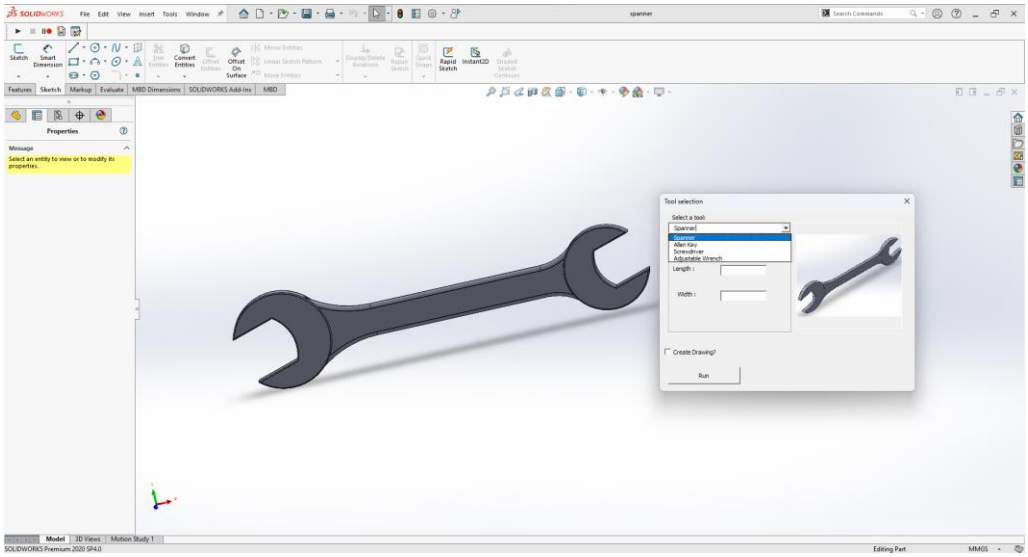


Figure 2: Available models

For each tool there are textboxes available which must be utilized before clicking on the "Run" button. Each textbox has validations for the minimal and maximal integer number for the selected tool - this prevents the user of creating a miniature spanner or an absurdly big one. If the user enters a faulty value, they will be warned by the system and informed what needs to be changed. They will also be provided with the min/max values or that particular dimension, as illustrated in Figure 3. The code for the spanner validation is shown in Figure 4.

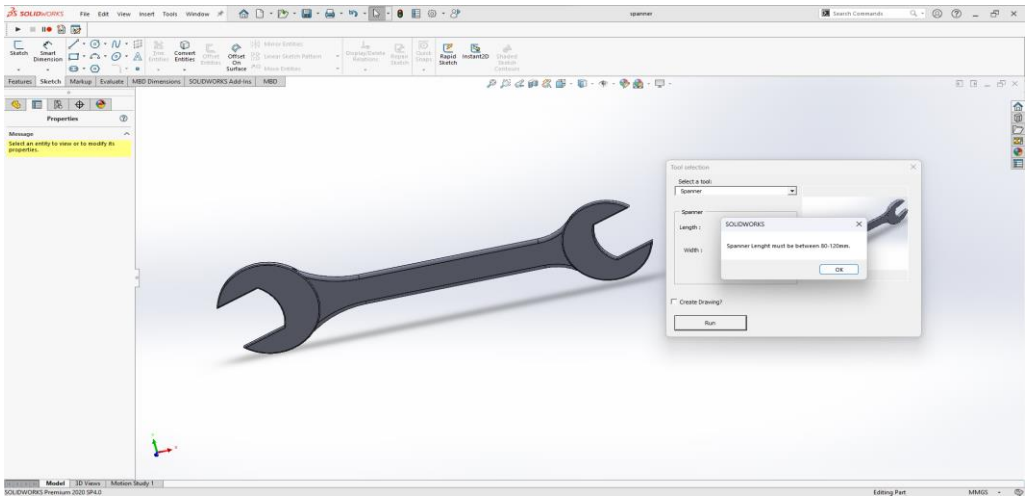


Figure 3: Validation for Spanner length

```
'Spanner Controls
If txtSpannerLenght.Value < 80 Or txtSpannerLenght.Value > 120 Then
MsgBox "Spanner Length must be between 80-120mm."
ControlValues = False
Exit Function
End If
```

Figure 4: Code for Spanner validation

After successful insertion of the values, the system will adjust the models as per requirement and it will call the function that writes the dimensions in a separate sheet for each model in an external Excel file, as given in Figure 5 and Figure 6.

```
Sub WriteToExcelFile()
Dim strDrawStatus As String

        If chkUpdateDrawing.Value = True Then
        strDrawStatus = "YES"
        Else
        strDrawStatus = "NO"
        End If

        Dim xlApp As Object
        Dim xlWorkbook As Object
        Dim xlWorksheet As Object


        Set xlApp = CreateObject("Excel.Application")
        Set xlWorkbook = xlApp.Workbooks.Open(excelPath)
        Dim lastRow As Long
        Select Case cmbModels.Text
        Case "Spanner":
        Set xlWorksheet = xlWorkbook.Sheets(1)

        lastRow = xlWorksheet.Cells(xlWorksheet.Rows.Count, "A").End(xlUp).Row + 1
        xlWorksheet.Cells(lastRow, "A").Value = spannerLenght
        xlWorksheet.Cells(lastRow, "B").Value = spannerWidth
        xlWorksheet.Cells(lastRow, "C").Value = strDrawStatus
```

Figure 5: Code for writing to external Excel file

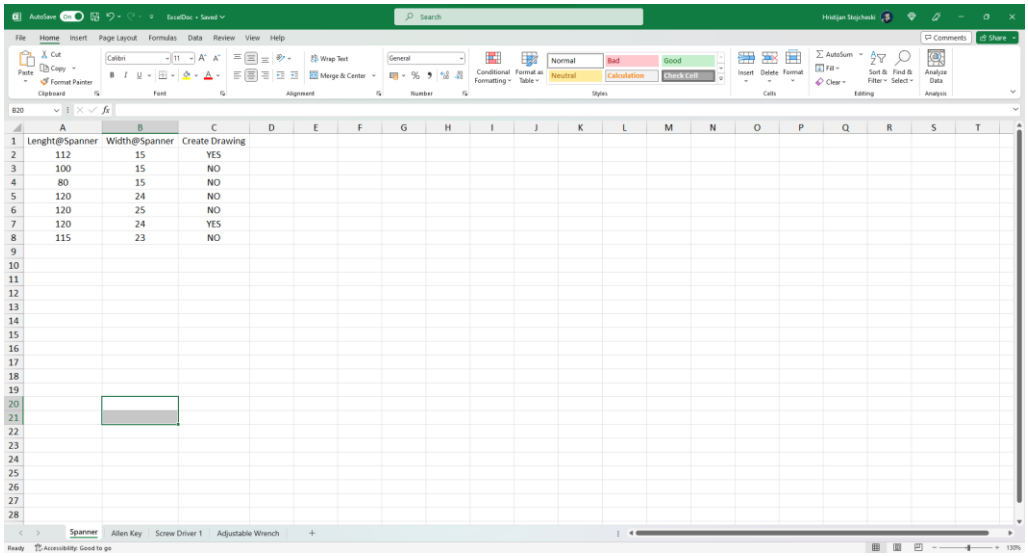


Figure 6: Excel data

Moreover, if the checkbox for creating drawing is selected by the user, the algorithm will create an automatic drawing of the model with the newly entered values by calling the method in Figure 7 in VBA. The created drawings of the spanner are illustrated in Figure 8.

```
Sub OpenAndUpdateDrawings()
Dim swDrawingModel As Modeldoc2
Dim drawingName As String
drawingName = folderPath & cmbModels.Text & "\" & cmbModels.Text & ".slddrw"
Set swDrawingModel = swApp.OpenDoc6(drawingName, swDocDRAWING, swOpenDocOptions_Silent, "", 0, 0)
swApp.ActivateDoc3 swDrawingModel.GetTitle, False, 0, 0
swDrawingModel.ForceRebuild3 True
'swApp.CloseDoc swDrawingModel.GetTitle
Set swDrawingModel = Nothing
End Sub
```

Figure 7: Code for creating drawing for a model

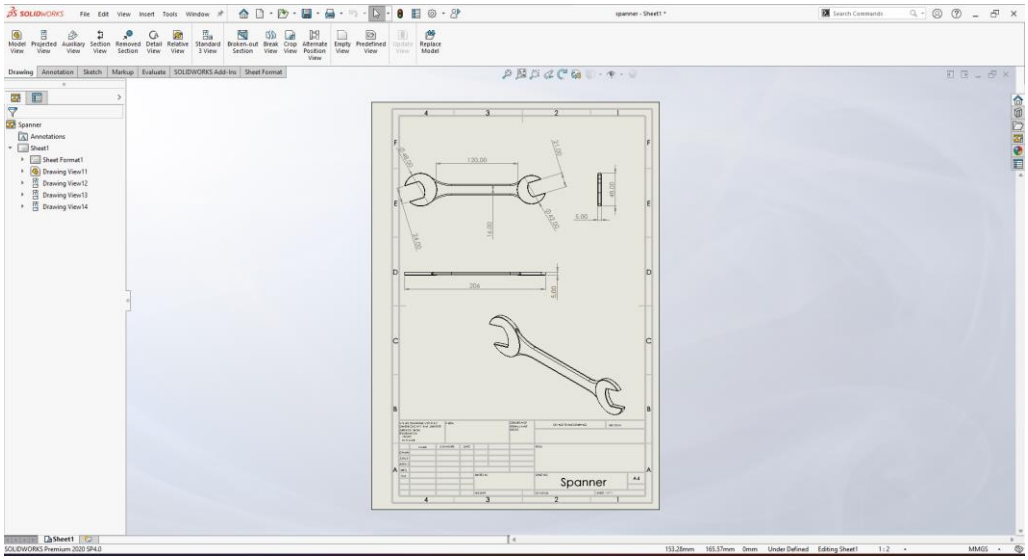


Figure 8: Drawing of a spanner

Finally, SolidWorks will rebuild the model and refresh the application so that the last entered values for the model are shown, Figure 9.

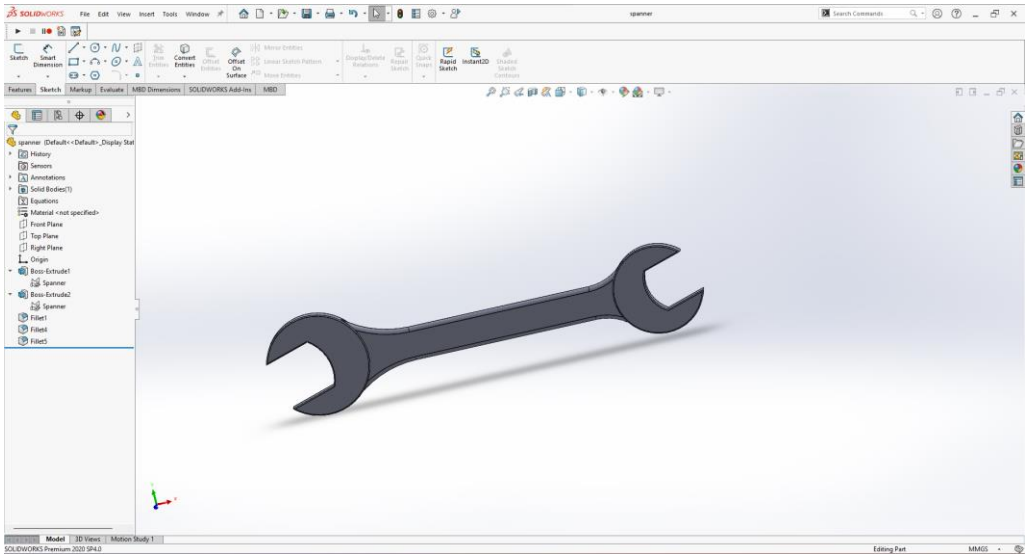


Figure 9: Final model

The process of creation of the three other models is the same as for the spanner. Dimensions are appropriately validated similarly as for the spanner and drawings can be created in the same manner.

5. Conclusions

Product customizations are significant for today's manufacturing industry, in order to gain a competitive edge among their competition. A better understanding regarding product customization and its impact on the customers can lead to a superior manufacturing process, which in turn will improve customer satisfaction. In this paper, an application was created to allow customers to be more involved in the design phase for product customization through the use of an interface that is seamlessly integrated with CAD's system API. The application provides customers with the ability to make design changes to a product via a user-interface form without the need to possess any type of CAD software skills. In the simplest sense, customers input values for pre-identified feature dimensions for a product model and the developed software program will automatically generate a new model based on the parameters. A VBA macro was written for design automation of repetitive and tedious tasks within SolidWorks. By involving customer in the new product development process, companies aim to get concepts and insights that allow them to improve an existing product or lunch a new product in the market. Customer involvement using same application in new product development is considered important for successful product development. The concepts developed for this paper can be used for any products with a broad scope of design opportunities. Directions for further work are to upgrade the existing algorithm, so that the user can select the material, color and texture of the tools, which would significantly improve this application.

References:

[1] A. Srikanth, Product Customization: Benefits, Examples, & Tips, Freshdesk Blog, April 2022, https://www.freshworks.com/freshdesk/general/product-customization-for-customer-satisfaction-blog/
[2] M. Barthwal, 6 Benefits of Product Customization, Knowband, August 2022, https://www.knowband.com/blog/ecommerce-blog/benefits-product-customization/
[3] The Ultimate Guide to Customer Success (2023 Updated), Chisel blog, 2023, https://chisellabs.com/blog/customer-success/
[4] J. Parker, 7 Effective Strategies to Improve Customer Retention, Knowband, August 2021, https://www.knowband.com/blog/ecommerce-blog/customer-retention-strategies/
[5] The Deloitte Consumer Review, Made-to-order: The rise of mass personalisation, Deloitte, 2019, https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/consumer-business/ch-en-consumer-business-made-to-order-consumer-review.pdf
[6] K. Saleh, Online Shopping Personalization – Statistics and Trends, invesp, https://www.invespcro.com/blog/online-shopping-personalization/
[7] W. C. Shih, The New Era of Industrial Policy Is Here, Harvard Business Review, August 2023, https://hbr.org/
[8] What is Competitive Advantage?, Chisel blog, July 2021, https://chisellabs.com/glossary/what-is-competitive-advantage/
[9] I. P. Nwaokorie, Customer Involvement in New Product Development: Investigation of customer involvement in the development of a product in a clean-tech company in Sweden , Master's Thesis, Uppsala Universitet, September 2020, https://uu.diva-portal.org/smash/get/diva2:1529472/FULLTEXT01.pdf
[10] Z. F. Z. Abidin, M. N. O. Zahid, "Real time object customization in CAD system", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 15, No. 2, August 2019, pp. 1066-1075, ISSN: 2502-4752, https://ijeecs.iaescore.com/index.php/IJEECS/article/view/19076
[11] D. R. Chowdhury, "Design Automation Methodology of a Cube using Solidworks and VB.net", International Journal of Engineering Research & Technology (IJERT), Vol. 9 Issue 06, June-2020, pp. 378-383, ISSN: 2278-0181, https://www.ijert.org/research/design-automation-methodology-of-a-cube-using-solidworks-and-vbnet-IJERTV9IS060352.pdf
[12] W. Mao, X. Zhang, S. Tantai, S. Shen, Z. Zhang, J. Wang, "Development of Parametric Drawing System of Corrugated Boxes Based on SolidWorks", IOP Conference Series: Materials Science

and Engineering, vol. 768, no. 7, 2020, p. 72053, https://iopscience.iop.org/article/10.1088/1757-899X/768/7/072053/pdf

[13] H. Karagulle, "Design Automation of Metal Scrap Balers by Integration of SolidWorks with Excel and VisualBASIC Language", International Journal of Engineering Research & Technology (Ahmedabad), vol. 7, no. 8, 2018, https://doi.org/10.17577/IJERTV7IS080095.

[14] M. Olander, A. Lüning, "Design automation in industrial order-to-delivery processes: Enabling mass customization of made-to-order products", 2021, http://liu.diva-portal.org/smash/get/diva2:1564234/FULLTEXT01.pdf

[15] A. Bocevska, S. Krsteski, S. Savoska, "An Example of Application Design using SolidWorks Application Programming Interface", 2020, pp. 159-166, https://ceur-ws.org/Vol-2656/paper16.pdf

Enhancing Spatial Exploration of Outdoor Object Recognition and Tracking with ARToolKit NFT Markers

Blagoj Nenovski, Igor Nedelkovski

University St Kliment Ohridski, 1 Maj nn, 7000 Bitola, R. North Macedonia

blagoj.nenovski@uklo.edu.mk; igor.nedelkovski@uklo.edu.mk

Abstract:
ARToolKit is an open source augmented reality toolkit that supports the recognition of fiducial markers and NFT (natural feature tracking) markers. There is significant research on optimizing and improving fiducial markers but an evident research gap on NFT makers. In this paper we provide a continuation of our previous research on creating NFT markers of outdoor objects by: choosing a source for the marker photo, comparing a range against the entire objects façade, address the level of initialization features and the level of tracking features in the NFT marker creation process.

Keywords:
ARToolKit, NFT, AR, Augmented reality, marker creation

1. Introduction

In our previous work [1] we defined an augmented reality end-to-end platform for spatial exploration with the time as an added component. This defining process was part of our goal to enable the exploration of past and future or houses or buildings that can be recognized by their natural features extracted from a photo of their façades. Since ARToolKit did not support cloud recognition we created a platform consisting of a smartphone app for recognizing the objects and displaying multimedia and a server for storage and distribution of the markers. By setting the bar for the platform to be available to as many users as possible we had to focus and develop guidelines for the best practices for creating markers and app settings that would allow for fast recognition and stable tracking. This translated to an app that would run on Android smartphones with low specifications, but thanks to the rapid development of smartphone processors and the included components such as camera, GPS and 4G connectivity we could focus on recognizing the objects by their natural features instead of using fiducial markers. We addressed the generating markers with the available parameters and the process of adding markers. We also addressed the importance of the camera calibration when using ARToolKit for marker recognition.

When recognizing outdoor objects there are factors in the environment such lighting, reflective surfaces and occlusion that can drastically affect the user experience. In a setup where the before mentioned conditions are similar, the methodology of marker creation as well as the app parameters are of key importance. We have worked on resizing the photo before creating the marker, emphasized the importance of camera calibration as well as the camera resolution on the recognition speed and quality tracking of outdoor objects [2].

2. Related work

There is a comparative study of planar fiducial markers [3] that analyzes the literature, describes the differences and limitations and conducts detailed experiments to compare the sensitivity, specificity, accuracy, computational cost and performance under occlusion.

Research has been done on comparing systems like ARTag, AprilTag and CALTag on the reliability and detection rate when occlusion of various types and intensity is present [4]. ARToolKit markers have been compared with similar systems like ARTag on the reliability, detection rates, and immunity to lightning and occlusion [5].

Fiducial marker optimizer is presented [6] in order to optimize the design attributes of ARToolKit markers, including black to white ratio, edge sharpness, information complexity and to reduce inter-marker confusion. There are multiple factors [7] that are important when designing and tracking ARToolKit fiducial markers. Each of the factors can affect the accuracy, detection speed and inter-marker confusion. The specific distribution of tracking accuracy and its dependency on the distance and the angle between the camera and the fiducial marker is addressed in [8].

In regard to ARToolKit specifically, there is research on the effect of edge sharpness, noise and markers distinction on markers reliability with a developed specialized algorithm for designing sharp-edged, de-noised and distinct markers [9]. Fiducial markers in ARToolKit have been explored in terms of marker sizes, marker distance from the camera, marker speed, the brightness in environment, the contrast level of lighting, as well as the correlation between marker size and distance [10].

Work has been done on solving the tracking failure problem on partially occluded marker in multi marker environment with addition of codebook based foreground detection model for detecting hand region in unexpected background environment [11]. There is a path generation algorithm [12] that automatically identifies fiducial markers in a building in order to create a path for user navigation. The algorithm has been implemented in an android application and internal mechanism for database creation and guidance system has been discussed.

Even relatively new research on the development of a network camera system for long distance use of augmented reality function using ARToolKit [13] focuses on using fiducial markers.

Another research focuses on tailoring paper media markers, improving recognition accuracy via integrated single-response matrix and optimized image matching for real-time tracking. Enhancing ARToolKit SDK's image segmentation by simulating scene changes with a 45° marker card rotation relative to the camera is addressed in [14].

3. Methodology

To determine the recognition speed we used ARToolKit’s feedback on the state of a marker being loaded and a marker being recognized. We achieved this by subtracting the marker load time from the marker recognition time. Our focus was to create better markers that could easily be recognized and tracked by entry level devices so we used Samsung J3(2017) and Samsung J4+ (2018) as validation devices. To create the markers we used different devices to eliminate the advantage in a scenario where the marker creation device is the same with our validation smartphones. While testing the recognition speed we put both smartphones in a fixed position and run each test for 5 times. For quantification of the marker tracking we simulated the use of the app in a perspective of a user that is using the app for the first time: holding the smartphone in a natural position, pointing it at the object until the marker is recognized after which we simulated various intensity phone movements. All of the tests were done with sampleRate set to 30 and cutoffFreq parameter set to 15.

To remove the parts that are not needed for object recognition and to enable quality display of multimedia over the recognized object, prior the marker creation we extracted the objects façades from the photos. Based on the conclusion of our previous work the images were resized to 1000 pixels before the marker creation. Visualized results of the recognition speed are the representation of subtracting the marker recognition time form a baseline of 5000ms as our defined limit for good user experience.

4. Choosing a source for the marker photo

First we wanted to determine the difference in recognition and tracking of markers created with different devices. For this we took a photo of the object for recognition, with a mobile phone (different from the validation devices) and with a digital camera (Nikon D3300). We initially scaled both images down to 1000 pixels and created markers from them with DPI values set to 96, 48 and 24. The created markers resulted in the following files (expressed in size) and the following number of initialization and tracking features.

Table 1:
File size and number of features

		File size in KB				features	
DPI	Source	iset	fset	fset3	Sum (KB)	fset	fset3
96	Smartphone	94	3	76	173	142	596
96	Digital camera	77	3	77	157	136	588
48	Smartphone	28	1	67	96	39	523
48	Digital camera	26	1	68	95	44	517
24	Smartphone	8	1	40	49	10	294
24	Digital camera	8	1	38	47	11	307

From table 1 we can see that the size of the files, as well as the number of initialization and tracking features have close values when using a smartphone and a digital camera for the photo from which the markers were created.

We performed additional analysis of the extracted features for each of the markers and found that for the most part the features are repeated among the markers with the same DPI value. However, markers created from a mobile phone and a digital camera photo are not identical and that they contain non-repeating features or features that are located in different positions.

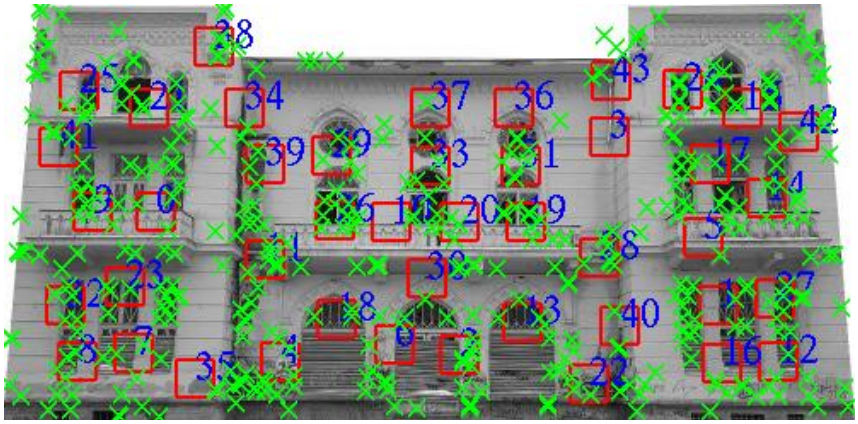


Figure 1: Extracted features, source: smartphone, 48 DPI

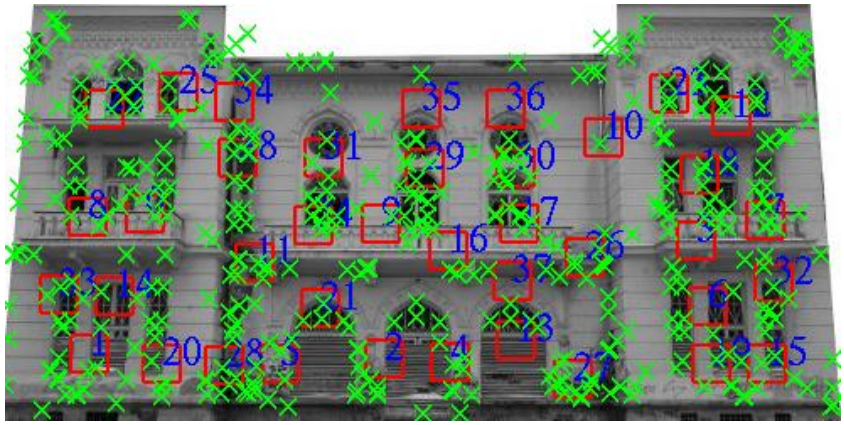


Figure 2: Extracted features, source: digital camera, 48 DPI

From the obtained results, we can see that although the original photo taken with a digital device has a higher resolution and a greater number of details, by reducing the photo to 1,000 pixels, regardless of the source, the recognition speeds in both cases are almost identical.

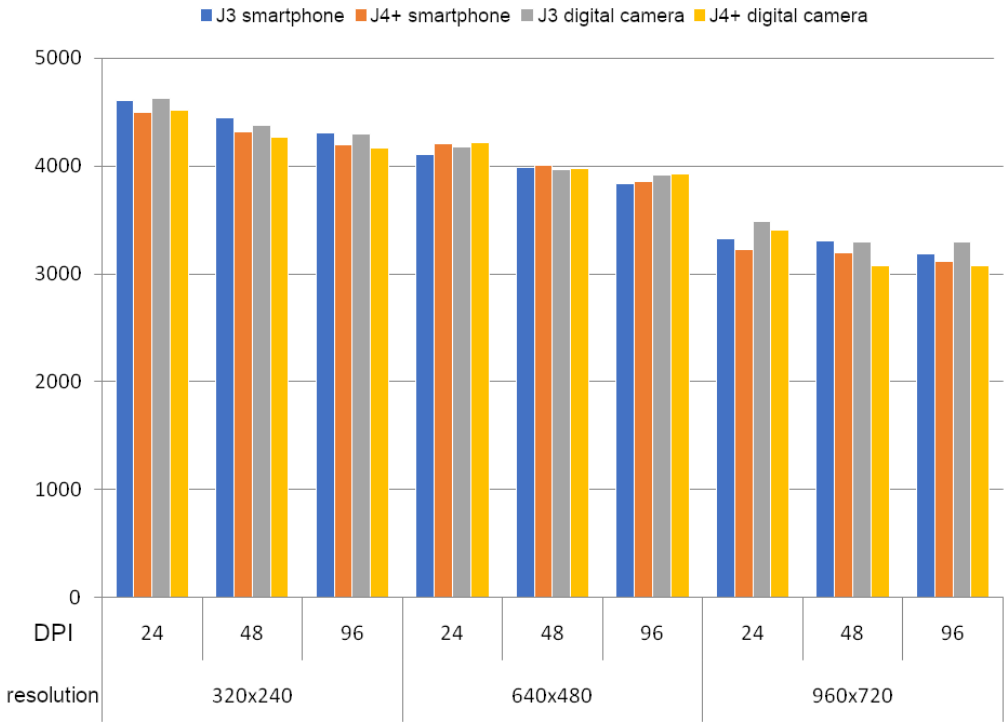


Figure 3. Recognition speed of markers created from smartphone and a digital camera photo

By further checking the quality of the tracking, we concluded that using each of the resolution values and each of the DPI values of the created markers, we obtained quality tracking, regardless of whether the photo to create the markers was taken with a mobile phone or a digital camera.

5. Range of the façade to create the markers

Here we focused on the range of the façade photo from which the markers are created. The reason we implemented this step is that the conditions do not always allow for a complete view of the objects. These conditions can be: the interference from additional objects or nature; many reflective surfaces; insufficient space to move away from the object.

In this step we tested the recognition speed as well as the tracking quality. In doing so, we compared the marker from the previous step (size: 1,000px, source: phone) with a marker with a different range from the same photo (in this case of the middle part of the façade). Our platform automates a large part of the steps, so for this marker we extracted the middle part of the façade from the original photo from the same photo as the previous marker. Then we reduced the image to a size of 1000 pixels and created markers with a DPI value of 96, 48 and 24. As a result we got the following markers and number of features.

Table 2:
File size and number of features

		File size in KB				features	
DPI	Range	iset	fset	fset3	Sum (KB)	fset	fset3
96	Wide	94	3	76	173	142	596
96	Middle	142	5	79	226	225	605
48	Wide	28	1	67	96	39	523
48	Middle	48	2	64	114	58	491
24	Wide	8	1	40	49	10	294
24	Middle	14	1	46	61	21	353

From table 2 we can see that the number of initialization features is almost the same for 96 DPI (wide 596, middle 605), higher for wide (523) than middle (491) for 48 DPI, and higher for middle (353) compared to wide (294) for 24 DPI. However, the difference in these initialization feature numbers is not large as seen from the recognition speed results. Here we should state that the features from the middle marker are located in one part of the camera view when exploring the object. We can see the extracted features of this marker in the following figure.

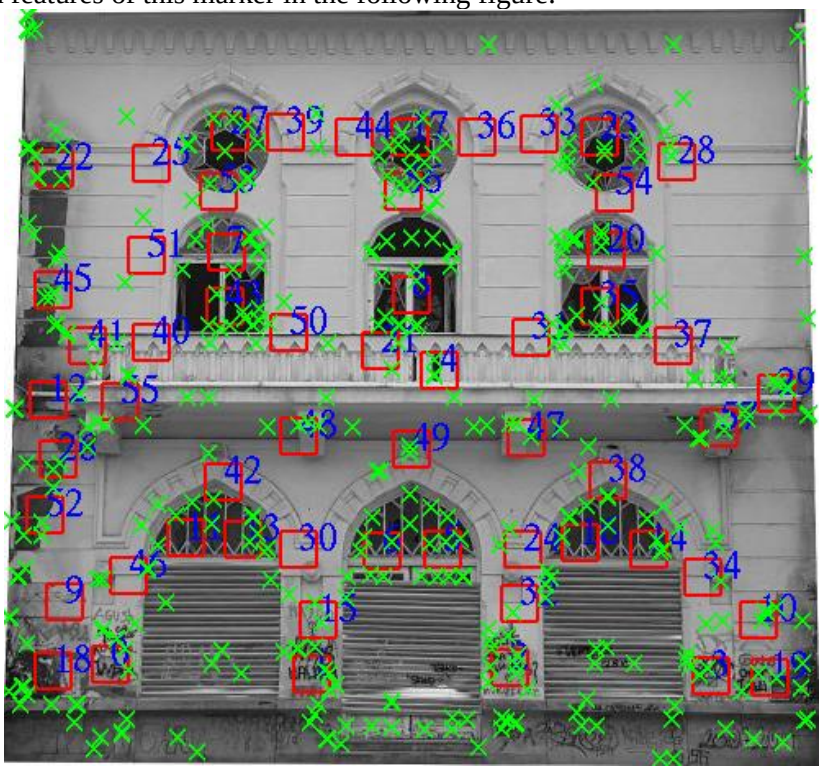


Figure 4. Extracted features, 48 DPI

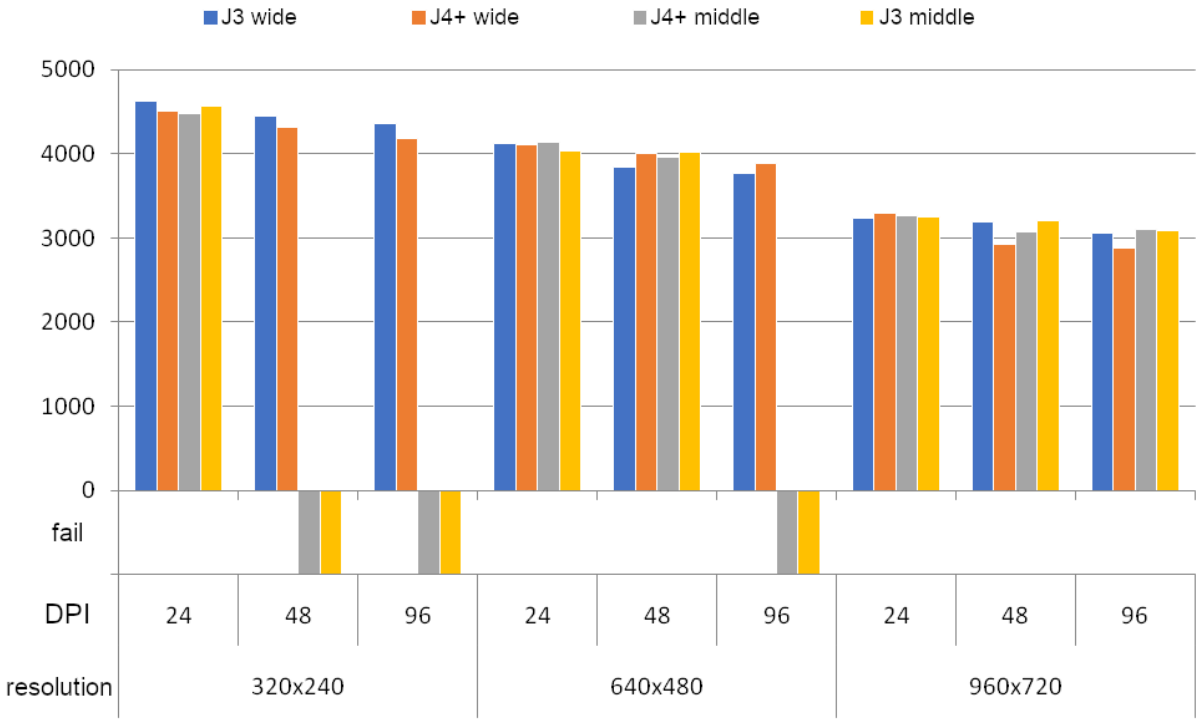


Figure 5. Recognition speed of markers created with different ranges of the photo

From the obtained results, we realized that in the cases where we had successful recognition with both markers, the speed of recognition was almost identical. On the other hand, since the testing was done with a camera setting that had a view of the entire object, we got 3 cases in the two phones: 320 x 240 camera resolution (48 DPI and 96 DPI) and 640 x 480 camera resolution (96 DPI) where we had failed recognition. With these cases, we confirmed the relationship between the DPI value and the resolution value of the camera.

Since with the marker with a DPI value of 24 we had a smaller number of features, we got a successful recognition with the camera resolution at 320 x 240. At the same resolution we had a failed recognition for the markers created with DPI values of 48 and 96. The same situation repeated for the 640 x 480 camera resolution for the marker with a value of 96 DPI. Already at a resolution of 960 x 720, enough details were obtained from the camera to recognize all three values for the middle marker. Of course, if the phone was placed closer to the middle of the object, then we would get different results, but in such a case the display image would cover the entire screen, so the object could not be explored by changing the position of the photo in relation to the position of the marker in the camera view.

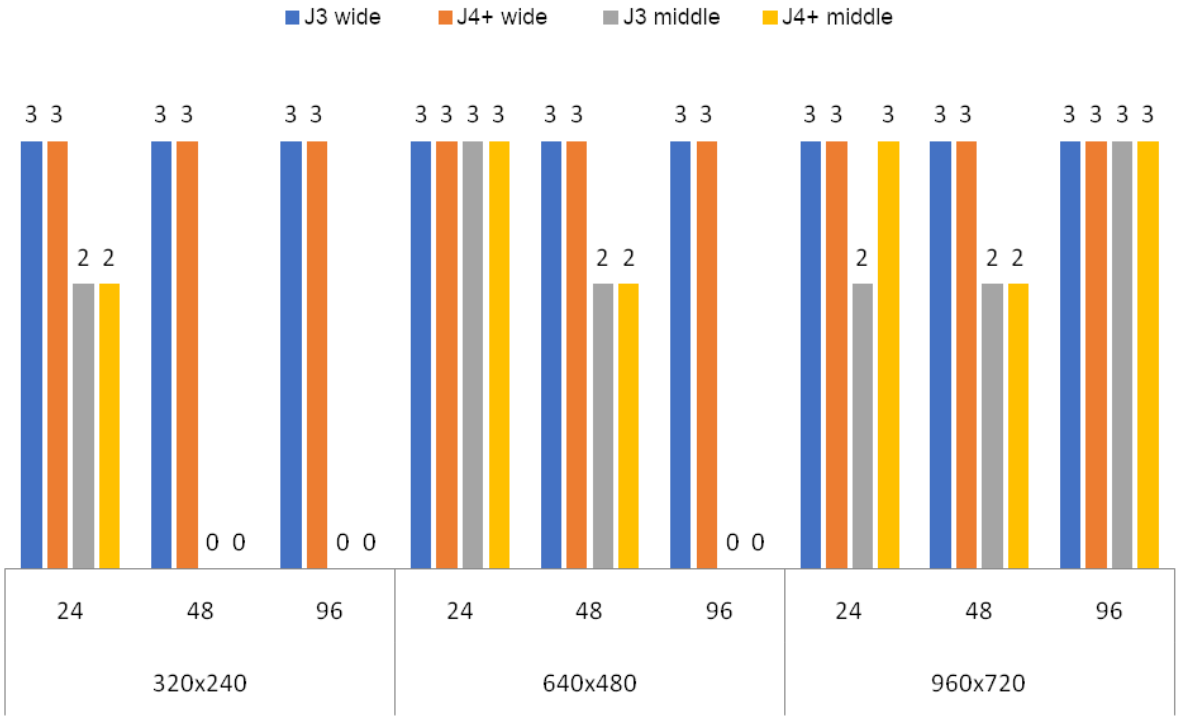


Figure 6. Tracking quality of the markers created with different ranges of the photo

In the tracking quality check we again failed to achieve recognition in the same cases that got the failed results in the recognition speed validation. In two cases (resolution 640 x 480, 48 DPI and resolution 960 x 720, 96 DPI) we had high tracking quality with both markers and both phones. In all other cases, either one or both phones consistently had lower tracking quality when the middle marker was used. As we previously emphasized, this condition is due to the distance from which the object is being explored. Since in the case of the middle marker, only the middle part of the building's façade is used and the same photo (of the entire building) is used for display, we had a lower tracking quality with this marker. This happened both when the phone is held steady and during similar movement, as for the wide marker.

6. Level of the initialization features

In the fourth step we created individual markers with a DPI value of 48, a default value of 2 for level of the tracking features, but a different value for the initialization features level from 0 for the lowest level, up to 3 for the highest level of initialization features.

For each of the four individual markers we got files of equal size, i.e. iset 28 KB, fset 1 KB, and fset3 67 KB. For the default level (2) of tracking features, we obtained the same number of features (44) for each of the markers. On the other hand, despite the fact that for each of the four markers we had set a different level of initialization features, for each of the markers we obtained the same number of features (517).

In the five repetitions of individual tests for the marker, we got results that were close to each other, which we can see in the graphical representation of the average value from the two phones.

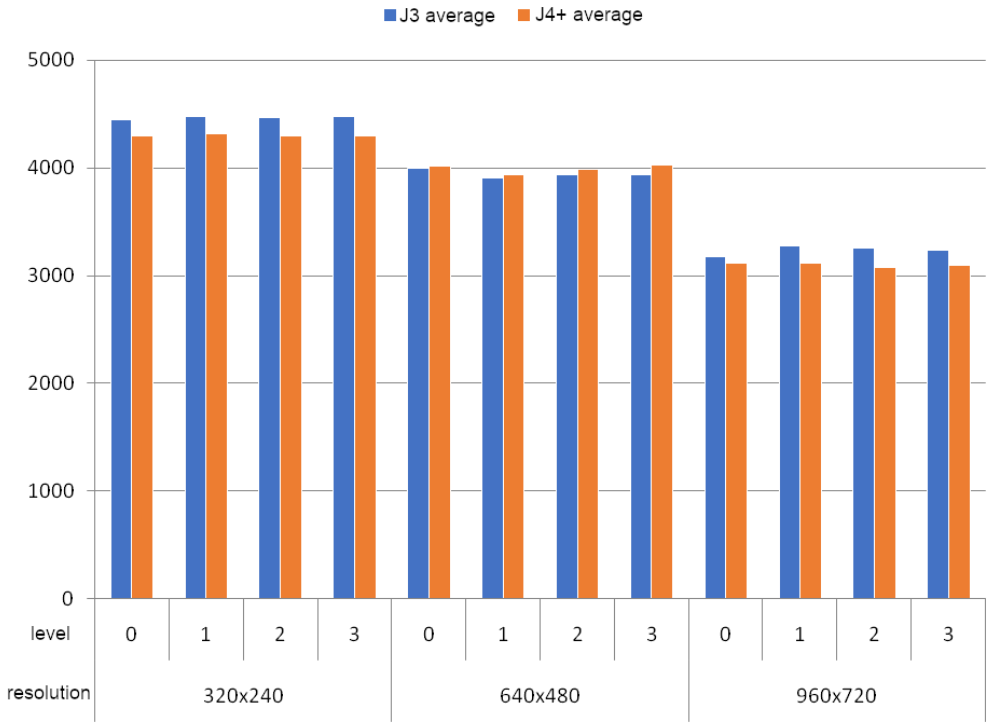


Figure 7. Recognition speed at different levels of initialization features

As a result of the identical markers, we got identical results, regardless of the set level of the initialization features. From the display in the graph, we can see the difference in the recognition speed with respect to the change in the camera resolution of the phones. Again we can see the trend of decreasing recognition speed when increasing camera resolution. This difference is smaller between values 320 x 240 and 640 x 480, where the difference is less than half a second, and larger between values 640 x 480 and 960 x 720, where the difference in recognition speed is about one second.

Getting the same result for the initialization features regardless of the setting is due to the setting not being supported by ARToolKit 5.3. Customization was part of the SURF extractor used in V5.2, but switching to the FREAK detector in 5.3 regardless of the settings gives the same result.

In the tracking quality test with each of the markers and with each resolution value, we obtained a high level of tracking quality, noting that in real world scenarios, at the time of not recognizing an already recognized marker, the speed of re-recognition effects the user experience. This means that at 320 x 240 and 640 x 480 resolution values, the photo display is significantly faster and provides a more realistic experience compared to the 960 x 720 camera resolution.

7. Level of tracking features

In this step, we tested the speed of marker recognition and its tracking with different levels of tracking features. In the previous step, we observed that regardless of the level set for initialization features, we consistently got the same files and the same number of features, resulting in equal recognition speed. Since there are a total of 5 levels of tracking features, in this step we created a total of 5 markers and set a variable level for the tracking features while keeping the same level for the initialization features. The different numbers of tracking features are shown in the following table.

Table 3
Number of features at different set level of tracking features

level	fset3 features	fset features
0	17	517
1	22	517
2	44	517
3	44	517
4	68	517

.

As expected the recognition speed is the same as in the previous step. When analyzing the tracking quality, we obtained stable tracking for each of the markers in each of the resolution values. Only, when using the 320 x 240 resolution and tracking features at level 0, we got a lower tracking quality, but such a difference is only noticeable when directly comparing the different markers.

8. Conclusion

We can conclude that when following the principle of reducing the photo size prior creating the NFT marker, using a digital camera for taking the photo does not differ from taking the photo with a smartphone in regards to both the recognition speed and the tacking quality.

With successful recognition of NFT markers when using a part of an objects façade for marker creation we can expect similar speed to when using the entire façade. But there are instances when using part of the façade that result with failed recognition as well as the problem with overlaying multimedia while tracking the marker.

We noticed that although there is a setting for selecting the level of initialization features, created markers with variable setting for initialization features bring identical file sizes and the same number of initialization features. Choosing the level of tracking features plays a small role in the tracking quality, except for when using the lowest level while coupled with 320 x 240 for the camera resolution.

References:

[1] B. Nenovski and I. Nedelkovski, "Defining a feature-rich end-to-end augmented reality platform for spatial exploration," in Proceedings/8th International Conference on Applied Internet and Information Technologies, vol. 8, no. 1, pp. 103-108, "St Kliment Ohridski" University-Bitola, Faculty of Information and Communication Technologies-Bitola, Republic of Macedonia, 2018.

[2] B. Nenovski and I. Nedelkovski, "Recognizing and tracking outdoor objects by using ARToolKit markers," International Journal of Computer Science & Information Technology (IJCSIT), vol. 11, no. 6, pp. 21-28, 2019.

[3] D. Jurado-Rodriguez, R. Muñoz-Salinas, S. Garrido-Jurado, and R. Medina-Carnicer, "Planar fiducial markers: a comparative study," Virtual Reality, pp. 1-17, 2023.

[4] Sagitov, K. Shabalina, R. Lavrenov, and E. Magid, "Comparing fiducial marker systems in the presence of occlusion," 2017 International Conference on Mechanical, System and Control Engineering (ICMSC), St. Petersburg, Russia, 2017, pp. 377-382, doi: 10.1109/ICMSC.2017.7959505.

[5] M. Fiala, "Comparing ARTag and ARToolkit Plus fiducial marker systems," in IEEE International Workshop on Haptic Audio Visual Environments and their Applications, pp. 6-pp, IEEE, 2005.

[6] D. Khan, S. Ullah, D.-M. Yan, I. Rabbi, P. Richard, T. Hoang, M. Billinghurst, and X. Zhang, "Robust tracking through the design of high quality fiducial markers: an optimization tool for ARToolKit," IEEE Access, vol. 6, pp. 22421-22433, 2018.

[7] D. Khan, S. Ullah, and I. Rabbi, "Factors affecting the design and tracking of ARToolKit markers," Computer Standards & Interfaces, vol. 41, pp. 56-66, 2015.

[8] D. F. Abawi, J. Bienwald, and R. Dorner, "Accuracy in optical tracking with fiducial markers: an accuracy function for ARToolKit," in Third IEEE and ACM International Symposium on Mixed and Augmented Reality, pp. 260-261, IEEE, 2004.
[9] D. Khan, S. Ullah, and I. Rabbi, "Sharp-Edged, De-Noised, and Distinct (SDD) Marker Creation for ARToolKit," in Information and Software Technologies: 20th International Conference, ICIST 2014, Druskininkai, Lithuania, October 9-10, 2014. Proceedings 20, pp. 396-407, Springer International Publishing, 2014.
[10] I. Rabbi, S. Ullah, M. Javed, and K. Zen, "Analysis of ARToolKit fiducial markers attributes for robust tracking," in International Conference of Recent Trends in Information and Communication Technologies (IRICT'14), University Technology Malaysia, pp. 281-290, 2014.
[11] H. Wang, J. Qin, and F. Zhang, "A new interaction method for augmented reality based on ARToolKit," in 2015 8th International Congress on Image and Signal Processing (CISP), pp. 578-583, IEEE, 2015.
[12] Ullah, S., Rahman, I. U., & Rahman, S. U. "Systematic augmentation of ARToolKit markers for indoor navigation and guidance." Proc. Pakistan Academy of Sciences: A. Physical and Computational Sciences, vol. 56, no. 1, pp. 1-8, 2019.
[13] K. Ogata, S. Nakatani, and S. Yamaguchi, "A network camera system enabling long-distance use of augmented reality functionality using ARToolKit," International Journal of Innovative Computing, Information and Control, vol. 17, no. 2, pp. 655-669, 2021.
[14] W. Zhang, W. Zhang, M. Xu, Y. Zheng, and H. Wang, "Design and implementation of augmented reality system for paper media based on ARToolKit," in Applications of Digital Image Processing XLIV, vol. 11842, pp. 423-430, SPIE, 2021.

NFT Marker Recognition in Multi-Marker Environment and Media Integration in ARToolKit

Blagoj Nenovski, Igor Nedelkovski

University St Kliment Ohridski, 1 Maj nn, 7000 Bitola, R. North Macedonia

blagoj.nenovski@uklo.edu.mk; igor.nedelkovski@uklo.edu.mk

Abstract:
ARToolKit is an open source augmented reality toolkit that supports the recognition of fiducial markers and NFT (natural feature tracking) markers. Placing fiducial markers for recognition of outdoor objects can result in economic, legal and logistical challenges. NFT markers are software based and in our platform are distributed by filtering of an area around the user's location. Since there are instances for either multiple markers or a single marker that can be recognized from multiple sides we addressed the effect on NFT maker recognition and tracking in multi marker environment. After successful recognition we focus on displaying various multimedia.

Keywords:
ARToolKit, NFT, AR, Augmented reality, AR multimedia

1. Introduction

To address the problem of ARToolKit not supporting cloud recognition we defined a platform [1] that consists of an app executed on a smartphone and a server that filters the markers and their corresponding multimedia based on an area around a user's location. By targeting as many devices as possible we had to rely only on the smartphones camera for optical recognition. We have addressed the comparison of natural features tracking compared to fiducial markers and the benefit of NFT markers in regards to recognizing outdoor objects, making NFT much larger markers. This enables for recognition from further distances while eliminating the need to place fiducial markers which would occlude the objects. We provided guidelines for generating markers, for resizing the photo prior marker creation, discussed the level of extraction features and emphasized the need for camera calibration. We have addressed NFT marker creation [2] with suggestions on separating the objects façade from the photo and resizing the photo prior the marker creation and their effect on the recognition speed and tracking.

Filtering the markers by user location enables for downloading multiple markers for objects that are in the area of the user's location as well as scenarios where a single object can be explored from multiple sides of its façade in which case we need as many NFT markers as the object's available sides for exploration. For such cases we will address the effect on the recognition speed and tracking quality with increasing number of NFT markers.

In this paper we will point out the evident lack on research on the effects on load times, recognition speed and tracking quality in NFT multi-marker environment. We will determine the load times, the recognition speed and display of a photo with increasing number of NFT markers. After that we will address the display of a photo, a video, sound reproduction, a 3D object and display of web link. We will also address changing the position and orientation of the display multimedia.

2. Related work

There is a comparative study of augmented reality SDK's [3] where the ability for multi marker recognition is one of the aspects that are addressed. Real-time camera tracking method using multiple markers with free camera movement is presented in [4]. ARToolKit can combine several co-planar fiducial markers into multi marker set. Single marker and multi marker tracking is tested in [5]. Multi marker tracking in ARToolKit is implemented by tracking all loaded markers separately and then

combining all the tracking results. Tracking multi markers is slower and comes with higher computational cost compared to single independent markers. Multi marker recognition is addressed in [6] with focus on finding the optimal values for multiple marker attributes such as: marker size, marker distance from camera, marker speed, environmental brightness, contrast level and the correlation between the marker size and the distance from the camera. Multi-marker approach for increased tracking robustness is addressed in [7] and research on determining a confidence factor for tracking multiple markers with ARToolKit in [8]. Multi marker tracking is suggested in [9] for ensuring sufficient accuracy by adding single markers to multi-marker setup. An automation method for calibration of multiple fiducial markers in order to obtain stable relations among markers is presented in [10].

All of the before mentioned research is based on fiducial markers and research on the effects such as load times, recognition speed and tracking quality with NFT markers in a multi-marker environment is practically non-existent.

3. Methodology

To determine the recognition speed we used ARToolKit's feedback on the state of a marker being loaded and a marker being recognized. We achieved this by subtracting the marker load time from the marker recognition time. Our focus when creating markers was to create better markers that could easily be recognized and tracked by entry level devices. That is why we used: Samsung J3(2017) and Samsung J4+ (2018) as test devices. We also added an additional flagship device: Samsung S9+ (2018) so we could compare the load times and the recognition speed. To eliminate the advantage in a scenario where the marker creation device is the same as the smartphone used in our tests we used a different smartphone to take the marker photos. To remove the parts that are not needed for object recognition and to enable quality display of multimedia over the recognized object, prior the marker creation we extracted the objects façades from the photos. Based on the conclusions of our previous work the images were resized to 1000 pixels before creating the markers.

When testing we put the smartphones in a fixed position and run the tests simultaneously. In the real world test we simulated the use of the app in a perspective of a user that is using the app for the first time: holding the smartphone in a natural position and then pointing it at the object until the marker is recognized. After that we simulated various intensity phone movements. All of the tests were done with sampleRate set to 30 and cutoffFreq parameter set to 15.

Visualized results of the recognition speed are a representation of subtracting the marker recognition time form a baseline of 5000ms, a limit we defined for a good user experience.

4. Time required for: loading the markers, recognizing the markers and displaying a photo when using additional number of markers

Our defined platform is initially expected to contain a small number of markers at a specific location. By enriching the content on the server itself, there will be markers for objects that are at a geographical distance smaller than the values set for filtering the markers. The impact of a different number of markers that can be used to recognize the same object on different mobile phones can determine the degree of a quality user experience.

Our goal was to measure the time required for loading the markers and their recognition speed.

We also provide an overview of the time it took to load different number of markers to determine whether it can affect the overall user experience. For the purposes of this validation, in addition to the marker that we used for recognition (1), additional 5 markers (2, 3, 4, 5, and 6) were created. All markers used a resized photo to 1000 pixels, a DPI value of 48, 2 for tracking features, and 1 for initialization features. We created scenarios for each of the set markers in a separate directory. The first test loaded only the marker that was used for recognition. After the first test was run, the recognition marker and an additional marker were loaded from the directory in numerical order. In this test, the same marker was recognized again, but two markers are loaded. The tests continued by

adding one additional marker for each of the scenarios. Table 1 shows the file size and number of features for each of the markers.

Table 1:

File sizes and number of features for each of the markers

	File sizes in KB				features	
Marker	**iset**	**fset**	**fset3**	**Sum**	**fset**	**fset3**
1	28	1	67	96	44	517
2	35	2	62	99	39	450
3	36	1	71	108	75	563
4	34	2	55	91	75	423
5	17	2	73	92	28	543
6	32	1	59	92	56	478

From the test results we noticed that loading an additional marker takes up to a maximum of 10 milliseconds. The difference between the load times of one marker and the additional 5 markers does not exceed 50 milliseconds for low-end phones. These values are low and do not play a big role in the overall user experience.

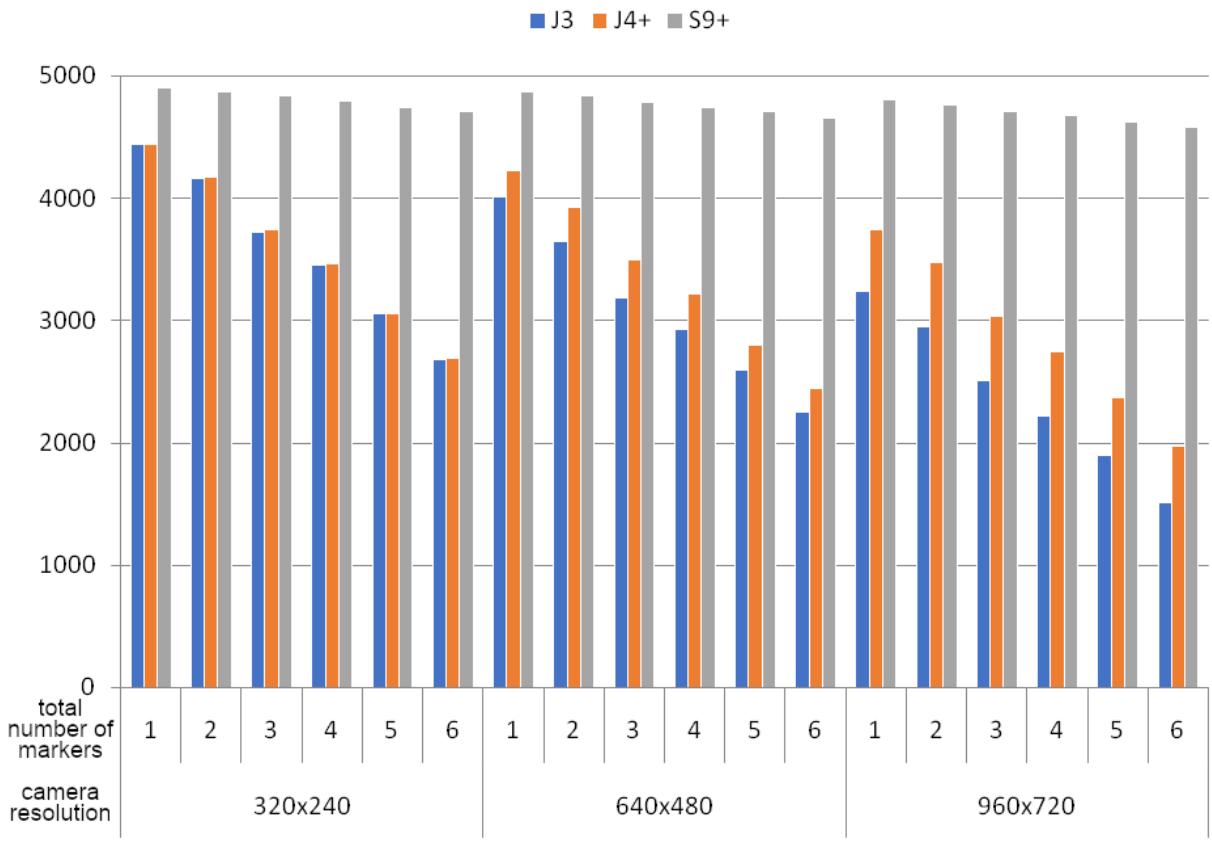


Figure 1. Recognition speed with different number of total markers

From figure 1 we can see the time required for marker recognition when multiple markers are included. We can see that the total number of loaded markers is directly proportional to the recognition time. This trend is repeated for the three camera resolution values tested, i.e. adding markers further reduces the recognition speed by an almost identical time for each additional marker.

Analyzing the data from the high-end phone, we can see that regardless of the resolution used, even in the case with the recognition marker plus additional 5 markers, we obtained high recognition speed. The customization in the process of creating the markers, as well as the selection of the camera

resolution, makes our platform suitable for low-end devices, but also allows for increasingly better results when using high-end phones.

In addition to the recognition speed we did additional field analysis on the tracking quality. We noticed that once a marker is recognized, in each of the cases regardless on the number of additional number of markers, the tracking was stable. From this we can confirm that the working principle of ARToolKit is to focus on the recognized marker as long as it is visible in the camera view.

With real word use of the app by the end users, phone movement is expected with changes in the camera view in order to explore the object from different perspectives. Various factors, such as rapid phone movement or a recognized marker leaving the camera view can lead to the marker not being recognized. In such a case, it should be taken into account that the speed of its re-recognition is in the range of the recognition speed obtained from the tests. This time can be reduced by setting a buffer space of several seconds. In such a case, upon stopping the recognition of a particular marker within the given buffer period, the application would only concentrate on recognizing the last recognized marker. The version of ARToolKit we were working with didn't allow dynamic loading of markers while the camera view is displayed, which prevented us from implementing that idea.

5. Multimedia content display in ARToolKit

Augmented reality applications can be rich in multimedia content. Since the available demos of the ARToolKit Android app only include a pre-generated cube, next we worked on the display of a photo, a video, sound reproduction, display of a web link and a 3D object.

5.1. Photo

When displaying content from the past, most of the time it would be presented in a form of a photo. To display a photo we drew a plane with 2 triangles in OpenGL [11] and applied the selected photo as a plane texture. The default width of the plane is set to 300 pixels, and the height is obtained from the aspect ratio of the photo. ARToolKit yields a transformation matrix that contains the marker's position, orientation, and skew relative to the camera. This transformation is given to OpenGL and even though the photo is in 2D, it is displayed in a 3D environment. This gives the effect that the image is part of the physical world. A more realistic experience can be achieved when using transparent photos in png format. We strongly suggest erasing parts that are not related to the object that we want to display, such as the background, trees, obstacles or side objects. To contribute to a more realistic experience we suggest applying a gradient between the transparent part and the object being displayed.

Figure 2. Representation of an object's past with a photo

5.2. Video

For displaying a video we use the same method of drawing a plane as in the case of displaying a photo. We initially tried continuous extraction of a single frame from a video and plotting the result onto the plane [12]. Such a method proved to be inefficient due to the high use of computational resources. We managed to optimize the app to a degree but even in cases where we had a reduction in the reproduction interval the results didn't provide realistic user experience. By giving the plane an external texture [13], we got an effective rendering of the videos.

5.3. Sound

In addition to a picture and a 3D object, a sound or a song can also be played. This option is removed when a video is selected because the sound is part of the video itself. Only a sound can be played without coupling it with a photo or a 3D object. We use the MediaPlayer [14] component from Android to play sounds. Sound can be paused when an object is not being recognized and resumed on re-recognition, or can be paused manually.

5.4. 3D object

To display a 3D object we needed a 3D engine. To display 3D objects we initially included the ArToolKitJpctBaseLib [15] library which includes concepts such as: camera, 3D-object and textures and supports the following formats: 3DS, OBJ, MD2, ASC and XML. The library is under the same license as ARToolKit.

Our initial display approach was the same as with a photo display, i.e. the 3D object was loaded once the marker was recognized. Usually the 3D objects are larger compared to photos and might need to be built before the marker is recognized. Real world tests with this library showed us that building and rendering the objects after the marker is recognized resulted in a display delay. We made an effort to optimize the application to load and build the 3D objects the moment it receives the configuration file. When a marker is recognized the 3D object would be ready for display after recognition, reducing the delay time between the markers recognition and the 3D object display. Real testing with this approach showed us that the time required to build the model is long and does not provide a quality user experience.

Figure 3. Display of a 3D object

Next we tried 3D Model Viewer [16], a 3D object display engine available as open source software. Android 3D Model Viewer is a demo of OpenGL ES 2.0 and has the ability to load

Wavefront OBJ, STL and DAE files. The application does not use additional libraries and the nature of its code being open source allows for robust application. It is available on the Play Store, making it suitable for testing and previewing a model before it is added to the platform.

By testing with several different models, we realized that the building process of the models and their display are significantly faster, and the possibility of transformations, such as changing the size, position and orientation, gave us a good synergy with our platform. The built model is given the pose of the marker in which the model is displayed. As with other visible multimedia contents, we have the possibility to change the size, position and orientation with 6 degrees of freedom.

5.5. Web Link

To receive additional information about the recognized objects we added the ability to display web links from the platform itself or other web pages. Web links are opened as a WebView component that is part of the Android operating system.

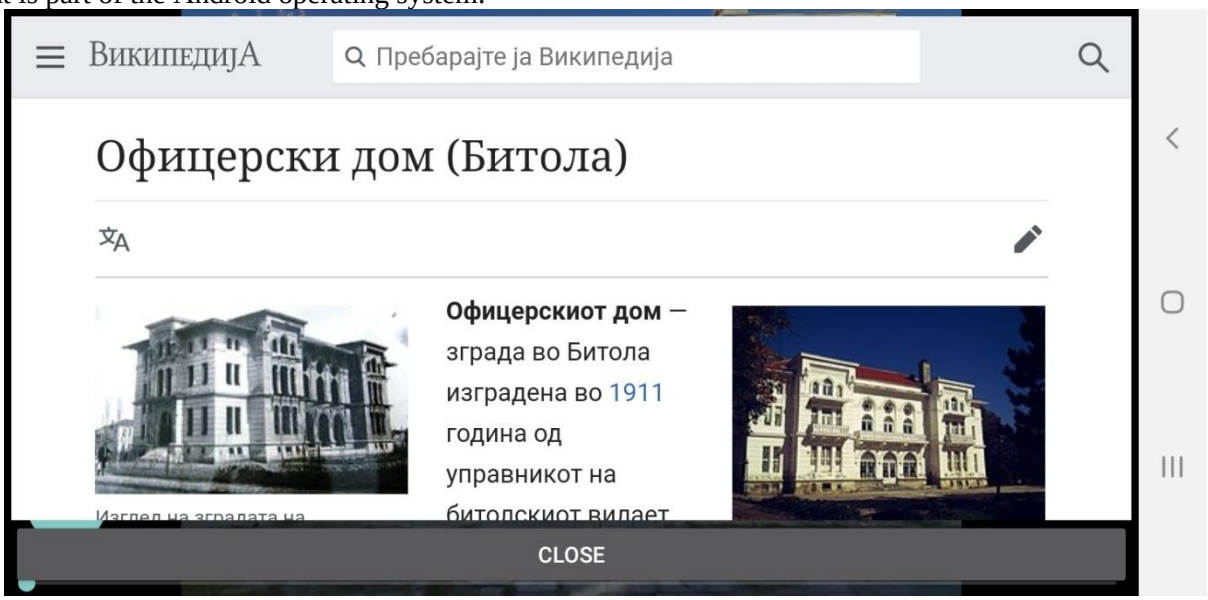


Figure 4. Web link

5.6. Changing the position and orientation of multimedia

Real world tests showed us that after adding a marker and its display objects, when displaying the multimedia over the recognized marker, there is usually the need to change the size, position and orientation of the displayed multimedia. The default setting for the size of the plane on which photos and videos are displayed is set to 300 pixels in width and a variable height in proportion to the content being displayed. By using 640 x 480 pixels as the default camera resolution, choosing a value of 300 pixels is an adequate size for a display that does not take up majority of the screen. However, the display content may not be the correct size relative to the marker. By increasing or decreasing the size of the plane, we can adjust the size to be equal as the size of the object in the camera view.

Since we have 3D registration, to address the need of 6 degrees of freedom, we can move the plane along the X, Y, and Z axes, or rotate it along any of the axes. In that way, by repositioning the objects, we also allow for robustness in cases where instead of the facade itself, we use a door or another/auxiliary object for recognition.

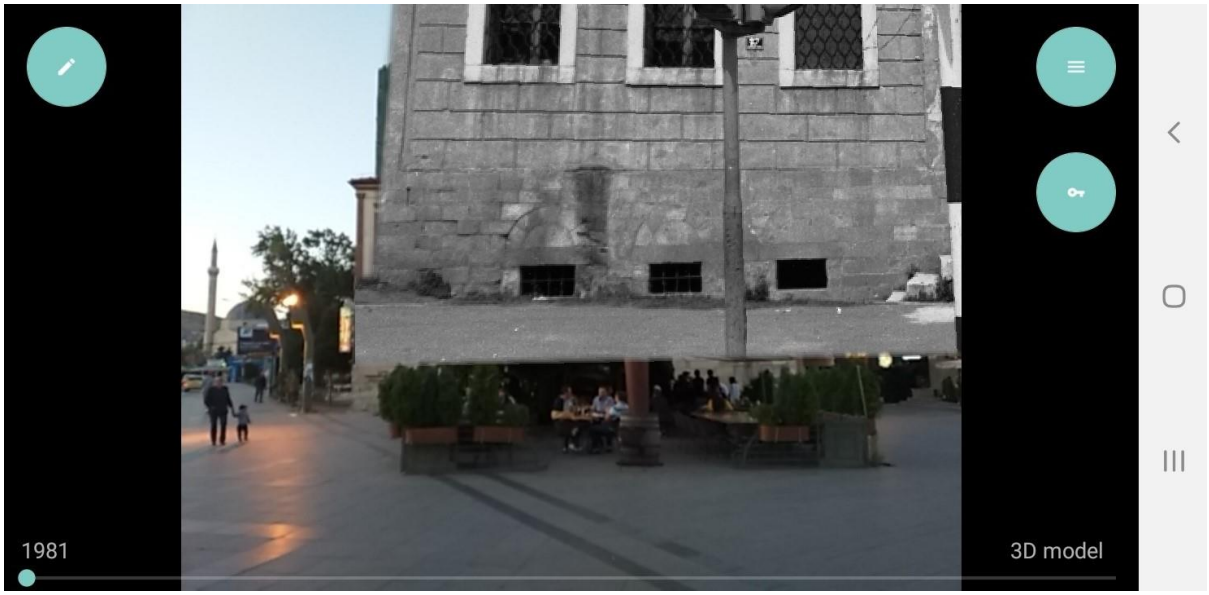


Figure 5. Initial photo display before resizing and positioning

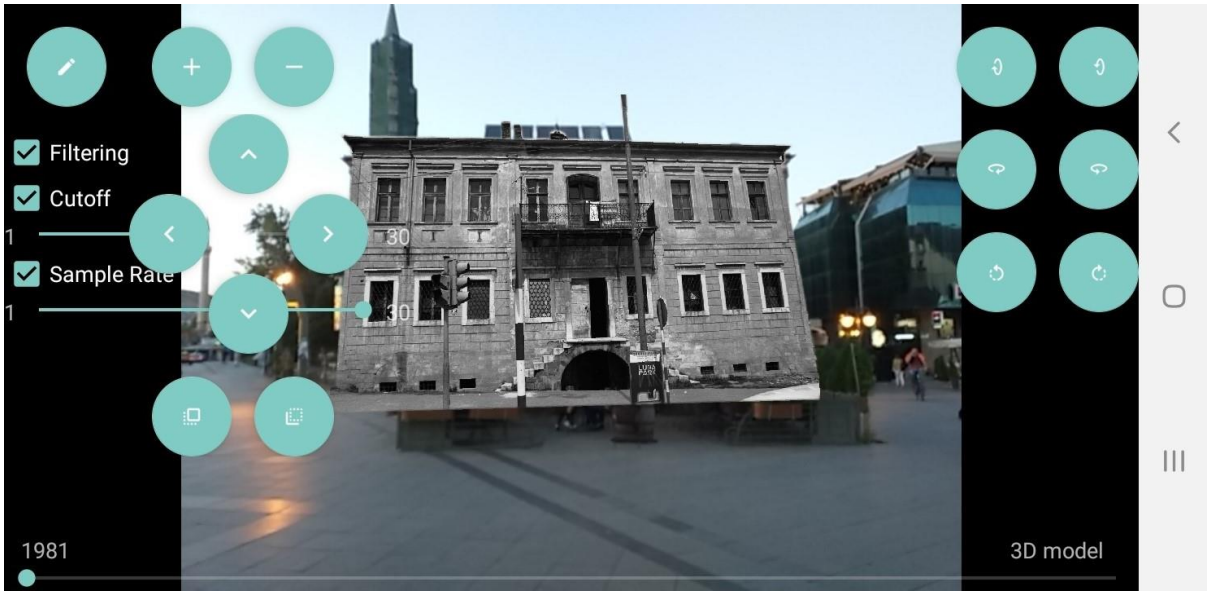


Figure 6. Adjusted size and position of the photo

6. Conclusion

In this paper we addressed the impact of different (total) number of markers when recognizing NFT markers on the load times and marker recognition speed and tracking. Load times of an additional marker take up to 10 milliseconds on low end smartphones. In a scenario with 5 additional markers, 50 milliseconds do not play a big role in the overall user experience. The total number of loaded markers is directly proportional to the time required for recognition for each of the tested camera resolutions. When loading additional markers with similar file sizes and similar number of natural features the recognition speed is reduced by an almost identical time. When using a high end smartphones, high recognition speed is obtained even in cases with 5 additional markers. On site testing proved us that tracking is not affected with additional markers, but we have to emphasize the effect that additional markers have on the recognition speed. In real world use, once a marker is not recognized, re-recognition time is dependent on the total number of additionally loaded markers. ARToolKit's inability for dynamic marker loading does not allow us to implement a buffer technique for certain a time interval for marker re-recognition.

For display of a photo we drew a plane and applied the photo as a texture. For more realistic experience we suggest using transparent photos of only the display objects and a gradient between the object and the background. For video play we gave external texture to the drawn plane and for sound we used MediaPlayer from Android. For 3D objects we incorporated 3D Model Viewer which provided much better results compared to the "default" ArToolKitJpctBaseLib library. Display of photos, videos, and 3D objects usually need resizing, repositioning or changing the orientation. We implemented an in app functionality to address this issue.

References:

[1] B. Nenovski and I. Nedelkovski, "Defining a feature-rich end-to-end augmented reality platform for spatial exploration," in Proceedings/8th International Conference on Applied Internet and Information Technologies, vol. 8, no. 1, pp. 103-108, "St Kliment Ohridski" University-Bitola, Faculty of Information and Communication Technologies-Bitola, Republic of Macedonia, 2018.
[2] B. Nenovski and I. Nedelkovski, "Recognizing and tracking outdoor objects by using ARToolKit markers," International Journal of Computer Science & Information Technology (IJCSIT), vol. 11, no. 6, pp. 21-28, 2019.
[3] D. Amin and S. Govilkar, "Comparative study of augmented reality SDKs," International Journal on Computational Science & Applications, vol. 5, no. 1, pp. 11-26, 2015.
[4] J.-H. Yoon, J.-S. Park, and C. Kim, "Increasing camera pose estimation accuracy using multiple markers," in International Conference on Artificial Reality and Telexistence, pp. 239-248, Springer Berlin Heidelberg, 2006.
[5] D. Wagner and D. Schmalstieg, "ARToolKitPlus for pose tracking on mobile devices," in Conference: Proceedings of 12th Computer Vision Winter Workshop CVWW07, 2007.
[6] I. Rabbi, S. Ullah, M. Javed, and K. Zen, "Analysis of ARToolKit fiducial markers attributes for robust tracking," in International Conference of Recent Trends in Information and Communication Technologies (IRICT'14), University Technology Malaysia, pp. 281-290, 2014.
[7] K. Anagnostou and P. Vlamos, "Square AR: Using augmented reality for urban planning," in 2011 Third International Conference on Games and Virtual Worlds for Serious Applications, pp. 128-131, IEEE, 2011.
[8] I. V. Vista, P. Felipe, D. J. Lee, and K. T. Chong, "Remote activation and confidence factor setting of ARToolKit with data association for tracking multiple markers," International Journal of Control and Automation, vol. 6, no. 6, pp. 243-252, 2013.
[9] D. F. Abawi, R. Dörner, and S. Reinhold, "Interactive specification of virtual objects' poses in an AR system using multi-marker tracking."
[10] D. C. B. de Oliveira, F. A. Caetano, and R. L. D. S. Da Silva, "A method to automate the calibration of a multiple fiducial marker setup," in 2014 XVI Symposium on Virtual and Augmented Reality, pp. 89-95, IEEE, 2014.
[11] The graphics pipeline, URL: https://open.gl/drawing.
[12] MediaMetadataRetriever, URL: https://developer.android.com/reference/android/media/MediaMetadataRetriever.html#getFrameAtTime(long).
[13] GLES11Ext, URL: https://developer.android.com/reference/android/opengl/GLES11Ext.html.
[14] MediaPlayer, URL: https://developer.android.com/reference/android/media/MediaPlayer.html.
[15] ArToolKitJpctBaseLib, URL: https://github.com/plattysoft/ArToolKitJpctBaseLib
[16] Android 3D Model Viewer, URL: https://github.com/andresoviedo/android-3D-model-viewer.

Content-Based Image Retrieval: Contemporary Trends and Challenges

Buen Bajrami[1], Zoran Kotevski[2] and Kostandina Veljanovska[3]

[1,2,3] *"St. Kliment Ohridski" University - Bitola, Faculty of Information and Communication Technologies - Bitola, Studentska bb., 7000 Bitola, North Macedonia*

[1]*bajrami.buen@uklo.edu.mk;* [2]*zoran.kotevski@uklo.edu.mk;* [3]*kostandina.veljanovska@uklo.edu.mk*

Abstract:
Content-Based Image Retrieval (CBIR) is a process that enables finding similar images in large sets of databases based on an image query. The purpose of CBIR engines is to mimic humans in the image classification process, thus a high computational cost is required to reach the proper selection of features, which must be as unique as possible. In the last two decades, the research related to CBIR has increased, but even today it is considered as a difficult process. For this reason, the researchers in this field have developed various models and techniques that help in rendering the image, trying to make the results as accurate as possible. Some of these techniques include Local Binary Pattern (LBP) histogram, Local Difference Binary (LDB), Local Tetra Pattern (LTrP), etc. In this paper, we review the latest research in the field of CBIR, we compare their performance based on several factors, such as calculation time, image acquisition time and accuracy of results and we conclude the paper with a discussion about which models have shown the best performance and what are their advantages and disadvantages for which we make some recommendations for the future research.

Keywords:
Image Retrieval, Content-Based Image Retrieval

1. Introduction

When we discuss image retrieval we have two technologies in mind: Text-Based Image Retrieval (TBIR), which enables image retrieval based on the metadata that the images contain, and the newer and more advanced Content-Based Image Retrieval (CBIR), which enables image retrieval through various processes based only on the visual appearance of the image, extracting their details such as color, fractures, shapes, etc. CBIR is a process that enables finding similar images in large sets of databases based on an image query. The most basic need in CBIR is to search and sort images from a given archive through human-machine communication. So the purpose of CBIR engines is to mimic humans in the image classification process. To increase the accuracy rate the requirements of the end user must be as concrete as possible. Also in this aspect, a high computational cost is required to reach the proper selection of features which must be as unique as possible, otherwise, the wrong choice of the model and then of the image features negatively affects the result by providing images that are not so similar to the query image. In CBIR, all high-level visuals are represented by feature vectors consisting of numerical values. Image feature vectors are used as input to Machine Learning (ML) algorithms through training and testing models. In the last two decades, the research related to CBIR has increased rapidly, but even today it is treated as a complex paradigm. In this manner, various models and techniques are developed that help in rendering the image, trying to make the results as accurate as possible. Some of the techniques used for CBIR include Local Binary Pattern (LBP) histogram, Local Difference Binary (LDB), Local Tetra Pattern (LTrP), Local Neighbor Pattern (LNP), Color Volume Histogram (CVH), Deep Learning (DL) and Color Moments (CM). We have also researched several models and frameworks such as Hybrid+SVM+RF, Hybrid+CFBPNN+RF, GMAF, LSH, Color Histogram, GMM, etc. We have analyzed the way they work, and we have compared the results obtained from different experiments based on several factors, such as calculation time, image acquisition time, and accuracy of results. Most of these techniques and models have improved CBIR technology by increasing the level of reliability and accuracy, providing satisfactory results. The

remainder of this paper is organized as follows. Section 2 presents state of the art in CBIR technologies and models. Section 3 concludes the paper with a discussion about the performance of the presented CBIR paradigms.

2. Latest research in CBIR

Leung, Ma and Zhang [1] propose adaptive multimedia indexing using Naïve Bayes classification retrieval of multimedia objects. The core of this search framework is based on capturing human judgment based on user queries to develop semantic indexes related to search terms. Thus, we can say that the lesson is concluded as a success or failure depending on whether the number of positive labels or negative labels, whichever of them reaches the majority. The authors estimate that stochastic methods should be used in multimedia information searches, because if they are missing we will always get the same results. A multi-agent framework is proposed where positive and negative labels are performed by agents and the result will be based on a stochastic method of agents competing with each other. All these tags occur independently. As a future challenge, it remains to do other research related to similar stochastic methods and to add Markov dependence as part of the analysis to increase the accuracy of the process.

Alrahhal and Supreethi [2] presented a CBIR method to find an image from a set of databases. This method covers several areas such as image segmentation, extracting features from the image, and converting these features into semantic features. In this paper, the authors have focused on extracting low-level (color, texture and shape) and high-level features from images. They perform CBIR in two main phases. The first phase enables offline search where the system extracts the feature from all images and stores them in the DB. The second phase is the online phase, where the user enters the query image and the system extracts features from this image and measures the similarity by calculating the distance between the query image (feature vector) and all images in the DB. The authors have analyzed several models that enable the classification of results, four of them have been analyzed in this paper: LBP, LDB, LTrP and LNP. For image classification, the authors use four types of algorithms, including Linear Discriminant, Support Vector Machine, K-Nearest Neighbor classification and Ensemble Classifiers. Also, an experiment was done with three databases, including Color database (Corel 1k), Texture database (Vistex database) and Faces database. The results obtained in the performed experiments show that the most accurate model is LNP compared to other models, in terms of average recall. The development of the LNP model in video retrieval remains an unsolved future work.

Hua et al. [3] developed a feature descriptor that enables conversion from the RGB color space to the HSV color space. The HSV color space closely mimics human color perception and can be interpreted as a cylinder. The authors find it easier to calculate the volume of the cylinder, so the most color areas are highlighted using the volume of the cylinder. Commonly used color spaces include RGB, Lab, LUV, YUV, HSV, YIQ, and YCbCr. Of these, RGB is the most widely used, but it is not a uniform color space. Therefore, the authors have proposed a new visual descriptive method called color volume histogram (CVH) which is based on visual perception. According to the experimental results, the authors claim that this histogram provides greater accuracy than other local histograms or text models. LBP and Multitext histogram (MTH) were used for comparison to validate the performance of this method. The authors also claim that CVH provides a better representation of edges, as well as a better spatial representation of colors compared to LPB, but in the representation of local structural information LPB performs better.

Hou and Wang [4] proposed an improved Gaussian Mixture Model (GMM) and presented a framework of standard image retrieval, where after the system accepts a query image, it enables image retrieval and starts preprocessing. After this stage, the image is divided into blocks. Then the classification according to complexity begins where different methods are used for image extraction and feature vectors are coded by Fisher vectors. In the end, the image acquisition is realized through the similarity index of the feature vectors. The authors have made improvements in several aspects of the Gaussian Mixture algorithm including the feature encoding algorithm, Gaussian mixture model initialization method and new Gaussian distribution generation. After the experiments carried out in PASCAL VOC2012 where a comparison was made between the standard Gaussian algorithm and the

improved one, it is proven that there is a significant improvement in image restoration through the improved algorithm. To compare these two frameworks, the Friedman test was used, which is non-parametric, but takes the performance rank of the algorithms as value. Therefore, we conclude that the framework proposed by the authors in this paper provides greater accuracy, takes less computation time, and increases efficiency.

Saha et al. [5] proposed a new descriptive model of CBIR, which is enabled using DL techniques. Briefly, this process works by collecting the input image and processing it to remove noise. Then, the processed image is extracted and classified using a feed-forward convolutional neural network. According to the experiments carried out in the paper, the technique proposed by the authors, where the processed image is extracted and classified, shows an accuracy performance of 95% and a precision of 79%.

Dhingraa and Bansal [6] proposed two CBIR models. First, Color moment is used, which provides spatial features of an image, and then LPB contains information about scale, noise resistance and brightness variability. To increase the accuracy of this system they use techniques that speed up the process of classifying the similarity between images. Two databases are used in this paper. 1) Corel 1-K: which contains 1000 images divided into 10 categories such as buses, buildings, beaches, food, mountains, etc. and 2) Oxford Flower, which contains 1360 images in JPEG format, divided into 17 categories, each of them containing 80 images. These data are examined through two proposed models: 1) Hybrid + Support Vector Machine (SVM) + Relevance Feedback (RF); and 2) Hybrid + Cascade Forward Back Propagation Neural Network (CFBPNN) + Relevance (RF). These are two innovative hybrid models, in which intelligent techniques have been incorporated. The SVM model is based on ML, while the CFBPNN model is based on deep learning. According to the results obtained from the experiments carried out with the two databases, the authors concluded that the CFBPNN model gives more accurate results but is slower in calculation than the SVM model.

Wagenpfeil et al. [7] presented a general framework that unites existing algorithms for image and video retrieval in a unified model of indexing, annotation and semantic retrieval. The authors developed a design concept based on "User Centered System Design", which is presented using Unified Modelling Language (UML) and use case design for the main activities. In this use case, the Multimedia Feature Vector Graph (MMFVG) is added, which helps in adding other metadata to digital content. General Multimedia Analysis Framework (GMAF) is a framework that combines and uses existing Multimedia processing systems for image, video, and textual information. New algorithms can be integrated into GMAF very easily. The actual results of GMAF processing show that the level of detail is greatly increased due to the recursive application of the algorithms. Furthermore, MMFVG is designed to be presented in Resource Description Framework (RDF) and queries can be written in the official SPARQL language.

Magliani et al. [8] employed a kNN graph, which makes a connection of all the nodes with each other, but takes a lot of time as a process. In this case, the Locality Sensitive Hashing (LSH) method was used, which saves time in building the graph compared to brute-force or divide-and-conquer methods. In this kNN graph, which is built based on database images, diffusion is applied, which helps distinguish nodes, assigning them different values depending on the similarity of the images. The diffusion applied in this paper is similar to the Google PageRank algorithm where a graph is selected with the help of diffusion which is used iteratively. To make comparisons between the query image and the data set image, the Euclidean distance is used. Each individual is assigned a probability, including genes. The best individuals are assigned from a buffer, and by comparing generations of individuals, in the end, only the individuals with the highest probability remain. This genetic algorithm is implemented using DEAP1. The authors used three image datasets for the experiment. Oxford5k contains 5063 images belonging to 11 classes. Paris6k contains 6412 images belonging to 12 classes. Flickr1M [19] contains 1 million Flickr images used for large. By comparing the results, it was proven that the genetic algorithm gives the same or better results compared to other algorithms such as random search, grid search, and PSO. The method proposed in this paper, utilizing kNN graphs, a diffusion process, and a genetic algorithm, has applications in CBIR technology, enhancing both time and quality optimization.

In [9], Tzelepi and Tefas present a new model of RF which uses deep Convolutional Neural Networks (CNNs). So, the goal is to use the user feedback from the CNN, which modifies its structure and provides better representations in the image return. The reason for using deep CNN is that recently a great advancement has been seen in image classification, digit recognition and pedestrian detection.

CNNs belong to deep learning algorithms which are based on deep neural network architecture. Also to improve the retrieval quality, the authors proposed retraining the convolutional architecture on a data set with relevant image statistics and the tested database classes. The presented method for the improvement of deep CNN affects the permanent improvement of the CBIR system. This is based on two stages. In the first stage, the system collects information from the feedback of various users and stores it. This information consists of questions and images relevant and irrelevant to these questions. Then, in the second phase, the system builds targets for each image based on user queries. In this paper, two image retrieval datasets are used for the experiment: the 102 category flower dataset, consisting of 8189 images divided into 102 categories, and the Inria Holidays 3 dataset, consisting of 1491 images, divided into 500 classes. After obtaining results from the experiments, the authors claim that there is a great improvement in the proposed model.

Shikha et al. [10] proposed a unique hybrid system that is based on content and extracting different attributes such as texture, color, and shape with the help of a Gray-level co-occurrence matrix (GLCM). For the development of this hybrid model, the authors have analyzed several different techniques such as Color Moment, Gray Level Co-occurrence Matrix, Region-props Process, Extreme Learning Machine, and Relevance Feedback. As for Color Moment, it has also been used as a technique to extract color regardless of angle, rotation, or scale. While extracting the shape features, the Mass, Centroid, Mean, Variance, and Dispersion parameters were calculated. Four large databases are used in this paper, and most of them are used in most of the experimental works. They are Corel-1K, Corel-5K, Corel-10K and GHIM-10. In forming a query or input image, all images of all datasets are used. If the acquired images are similar or consistent with the input image, then we say that this system is effective and the image retrieval system has been successful. The similarity between the input image and the output images is found by comparing all the images of the four databases. Similarity is then calculated using three distance metric techniques. Usually, the Euclidean distance is used as a basis because it gives a much more accurate and faster result compared to the Manhattan distance and the Minkowski distance, which very often give false results. Finally, the authors conclude that this Hybrid model is very effective and provides a concrete CBIR solution.

In [11], Magliani et al. use four large datasets to perform experiments, including Holidays with 1491 high-quality images, Flickr1M with 1 million images, SIFT1M consists of 1 million 128D Sift descriptors, GIST1M consists of 1 million 960D Gist descriptors. First, an analysis of the results obtained after the experiment was carried out for the first two datasets, i.e. Holidays + Flickr1M, where they refer to the acquisition time and acquisition accuracy. As the authors report, LSH and Multi-probe LSH obtained the best results but required a huge average query time. PP-index reduced the retrieval time but with a loss in terms of accuracy. The FLANN technique is also used, which is an open-source library for ANN, which is very successful in nearest-neighbor matching. It has achieved an accuracy rate of 83.97%. In the second scenario, the PP-index has achieved a withdrawal of 94.32%, but it takes a lot of time, i.e. 17 sec. LOPQ has a low accuracy of 19.93%, with a computation time of only 3msec. FLANN has achieved better results than the LOPQ technique but is still weaker than the PP-index.

Arun et al. [12] deal with the categorization of image extraction techniques that are part of the Bag of Visual Words (BoVW) model. In this model, the visual dictionary is built through K-Means clustering which is obtained from local image descriptors. Then these images are categorized using visual word histograms. The BoVW-based representation is derived from the given query and each image dataset is then classified to determine the degree of similarity between the images. The distance function is one of the most frequently used metrics for BoVW-based similarity determination. Six datasets are used throughout this paper. INRIA holiday dataset, Scene-15 dataset, Oxford dataset, GHIM-10K dataset, IAPR TC-12 dataset, and SUN-397 dataset. Depending on the capacity of the datasets, samples were also taken. To more easily extract the features of the images of interest, the authors have proposed the Hessian-Affine detector and the Scale Invariant Feature Transform (SIFT) descriptor which provide reasonable solutions. For obtaining the similarity between two histograms, it has been observed that the best metric is Chi-square. After reviewing the literature, the authors have proven that three approaches that influence the increase in the effectiveness of image restoration: (i) approaches that minimize the quantization error, (ii) approaches that minimize the semantic loss and (iii) approaches that incorporate spatial information of visual words. One of the challenges is that during the visual learning of the dictionary and the local encoding of the descriptor, some information is lost during the process. Each of the existing approaches tries to address only one specific limitation of the

BoVW model. A unique BoVW-based system is missing which would address all the issues that are still unresolved.

In [13], Devi and Parmar present various techniques that are part of the CBIR technology. Here they explain that there are four stages of image retrieval from a database based on geometric properties in the input image. They are the image database generation phase, input images from scanner-thinning editing, outline-based image retrieval phase, and identifying global shape similarity. The authors also give brief descriptions of CBIR features which are Texture Features, Color Features, Spatial Location Features, Shape Features, and Local Image Features.

Kumar and Esther [14] address three image feature extraction techniques, Gabor, Wavelet, and Histogram. Color Histogram is one of the most used techniques in CBIR, which is based on RGB space. In Matlab, the maximum number of bins used by Color Histogram is 256. The similarity between the Query image and the database image is calculated through distance metrics. The Gabor technique, as we have mentioned above in other papers, extracts information from an image. This technique is a multi-scale, multi-resolution filter. This two-dimensional filter is presented as a sinusoidal signal. By applying the Wavelet method, they divide the image into four sub-images with three bands: diagonal, vertical, and horizontal. These sub-images contain information about the texture.

Atlam et al. [15] performed a comparison between three techniques for extracting image features including Color Histogram, HSV Color Histogram, and Color Histogram Equalization. The authors have experimented using the WANG database, which is a subset of the Corel database, using 1000 images, divided into 10 categories of 100 images each. Euclidean distance and the correlation coefficient were used as similarity metrics, while the requirement was that 20 similar images be returned for each query image. Time, accuracy, and error rate were also calculated. After the experiment they concluded that HSV color histogram and color histogram give the best results but with different calculation times. As a metric for measuring similarity, the correlation coefficient turned out to be more accurate than the Euclidean distance.

3. Conclusions

After reviewing the literature, some conclusions can be drawn about CBIR. In recent years, significant progress has been made in this field, leading to an increase in the number of large databases designed for CBIR. Many authors have developed specialized methods aimed at enhancing the ease of use and accuracy of image retrieval, while also reducing computation time. Below is a list of some of the main methods for CBIR that proved to be quite practical and recommended to use.

1. LNP model - after experiments and comparisons with other similar models such as LBP, LDB, and LTrP, the most accurate model is LNP.
2. Regarding the HSV color space we have analyzed two models from which CVH can describe color, texture, shape, and other spaces. We can also say that in the representation of colors, and edges, as well as the spatial representation of colors, the CVH method is better than LPB, but in the representation of local structural information LPB is better than CVH.
3. Gaussian Mixture is also a fairly developed method in image restoration.
4. In the CBIR technology based on the conducted experiments, it turns out that the genetic algorithm gives the same or better results compared to other algorithms such as random search, grid search, and PSO.
5. Deep CNNs have also influenced the development of CBIR. So a new model of Relevance Feedback is proposed which uses CNN, where the goal is to use feedback from CNN, which modifies its structure and provides better representations in the image return.
6. A hybrid system developed in [10] analyzes different techniques such as Color Moment, Gray Level Co-occurrence Matrix, Region-props Process, Extreme Learning Machine, and Relevance Feedback, and turns out to be quite accurate in image retrieval.
7. One of the other widely used techniques in CBIR is Color Histogram, which is based on RGB space.

These are seven of the most important developments that have had an immediate effect on the improvement and expansion of image retrieval in general. In the end, we have managed to show a very

real state of development of CBIR, also presenting the orientation of this technology which is multidimensional. We have listed the most commonly used techniques in CBIR, through which many advances have been made that increased the quality and use of CBIR. However, in each of them, we have listed the shortcomings as well as research challenges that should be addressed in future research.

References:

[1] Clement H. C. Leung, R. F. Ma & Jiayan Zhang (2020). Adaptive Multimedia Indexing Using Naïve Bayes Classification. ADVCOMP 2020: The Fourteenth International Conference on Advanced Engineering Computing and Applications in Sciences, pp. 12-17.

[2] Maher Alrahhal, K.P. Supreethi (2019). Content-Based Image Retrieval using Local Patterns and Supervised Machine Learning Techniques. CSE Dept., JNTUH Collage of Engineering, IEEE, pp. 118-124.

[3] Ji-Zhao Hua, Guang-Hai Liu, Shu-Xiang Song (2019). Content-Based Image Retrieval Using Color Volume Histograms. International Journal of Pattern Recognition and Artificial Intelligence, Vol. 33, No. 11.

[4] Yong Hou, Qingjun Wang (2018). Research and Improvement of Content-Based Image Retrieval Framework. International Journal of Pattern Recognition and Artificial Intelligence. Vol. 32, No. 12, 1850043.

[5] Nilanjana Saha, P. Thiruvannamalai Sivasankar, Rahama Salman, Badria Sulaiman Alfurhood, Aishwary Awasthi, & Sukhwant Singh Bindra. (2022). Novel Technique in Content-Based Image Retrieval using Classification by Deep Learning in Artificial Intelligence. International Journal of Intelligent Systems and Applications in Engineering, 10(2s), pp. 256-259.

[6] Shefali Dhingraa, Poonam Bansal (2022). Designing of a rigorous image retrieval system with amalgamation of artificial intelligent techniques and relevance feedback. Journal of Intelligent & Fuzzy Systems. vol. 42, no. 2, pp. 1115-1126.

[7] Stefan Wagenpfeil, Felix Engel, Paul Mc Kevitt, Matthias Hemmje (2021). AI-Based Semantic Multimedia Indexing and Retrieval for Social Media on Smartphones. Information. Vol. 12, Issue 1: 43.

[8] Federico Magliani, Laura Sani, Stefano Cagnoni, Andrea Prati (2019). Genetic algorithms for the optimization of diffusion parameters in content-based image retrieval. Proceedings of the 13th International Conference on Distributed Smart Cameras, pp. 1-6.

[9] Maria Tzelepi, Anastasios Tefas (2018). Deep convolutional learning for Content-Based image retrieval. Neurocomputing. Vol. 275 pp: 2467-2478.

[10] B Shikha, P. Gitanjali, D. P. Kumar (2020). An extreme learning machine-relevance feedback framework for enhancing the accuracy of a hybrid image retrieval system. International Journal of Interactive Multimedia and Artificial Intelligence. Vol. 6, No. 2, pp: 15-28.

[11] Federico Magliani, Tomaso Fontanini, Andrea Prati (2021). Bag of indexes: a multi-index scheme for efficient approximate nearest neighbor search. Multimedia Tools and Applications. Vol. 80, pp. 23135-23156.

[12] K. S. Arun, V. K. Govindan, S. D. M. Kumar (2020). Enhanced bag of visual words representations for Content-Based image retrieval: a comparative study. Artificial Intelligence Review. Vol. 53, pp. 1615–1653.

[13] Pooja Devi, Mahesh Parmar (2017). A survey on cbir techniques and learning algorithm comparison. International Journal of Latest Trends in Engineering and Technology. Vol. 8, No. 1, pp. 197-205.

[14] Ashok D. Kumar, J. Esther (2011). Comparative study on CBIR based by color histogram, Gabor and wavelet transform. International Journal of Computer Applications. Vol. 17, No. 3, pp. 37-44.

[15] Hany Fathy Atlam, Gamal Attiya, Nawal El-fishawy (2013). Comparative Study on CBIR based on Color Feature. International Journal of Computer Applications. Vol. 78, No. 16, pp. 9-15.

Latest Advances in Video Indexing and Retrieval

Nora Pireci Sejdiu[1], Zoran Kotevski[2], Blagoj Ristevski[3] and Kostandina Veljanovska[4]

[1,2,3,4] *"St. Kliment Ohridski" University - Bitola, Faculty of Information and Communication Technologies - Bitola, Studentska bb., 7000 Bitola, North Macedonia*

[1]*pireci.nora@uklo.edu.mk;* [2]*zoran.kotevski@uklo.edu.mk;* [3]*blagoj.ristevski@uklo.edu.mk;* [4]*kostandina.veljanovska@uklo.edu.mk*

Abstract:
This paper provides a review of recent advancements in the field of multimodal video indexing and retrieval. The presented papers cover various topics including multimodal fusion, keyframe extraction, semantic analysis, scalability, and evaluation. The reviewed studies propose innovative methods and techniques to address the challenges associated with these areas. However, several unresolved issues and opportunities for future research are identified. These include the effective integration of multimodal information, improvement of keyframe extraction accuracy, semantic understanding of video content, scalability of video retrieval systems, and standardized evaluation protocols. By addressing these challenges, future research can contribute to the development of more robust and efficient video retrieval systems, enabling better access and utilization of large-scale video collections across diverse domains.

Keywords:
Video indexing, video retrieval, semantic video indexing, keyframe extraction

1. Introduction

Multimodal video indexing and retrieval have emerged as crucial research areas due to the exponential growth of video data in various domains such as surveillance, entertainment, education and social media. With the increasing availability of multimodal content, including text, audio, and visual information, there is a pressing need to develop efficient and effective methods for indexing and retrieving videos based on their content. The ability to automatically analyze and annotate videos enables improved search capabilities, content recommendation, and information extraction.

In recent years, several studies have made significant contributions to multimodal video indexing and retrieval by proposing innovative approaches and techniques. These studies address various challenges and issues related to different aspects of video analysis. For instance, the fusion of multiple modalities plays a critical role in capturing the comprehensive representation of videos [15, 16]. By leveraging the complementary nature of textual, audio, and visual features, multimodal fusion techniques aim to improve the accuracy and robustness of indexing and retrieval systems.

Another important aspect is the extraction of keyframes, which are representative frames that capture the essence of video content. Keyframes play a vital role in summarizing and representing videos effectively [8, 13]. Several studies have proposed methods for automatic keyframe extraction using feature extraction, shot boundary detection, and clustering techniques. These approaches aim to identify salient frames that encapsulate the content of a video and facilitate efficient retrieval.

Semantic analysis and annotation of video content have also been addressed in the literature [12]. The challenge lies in bridging the semantic gap between low-level visual features and high-level semantic concepts associated with videos. By utilizing approaches such as correspondence-Latent Dirichlet Allocation (corr-LDA) and information extraction techniques, researchers aim to automatically annotate videos with textual descriptors, enabling more accurate and meaningful retrieval of video content.

Scalability and retrieval performance are significant concerns in the context of large-scale video databases [4, 5]. Efficient indexing structures, optimization algorithms, and parallel processing techniques are being explored to improve the scalability and retrieval speed of video retrieval systems. Additionally, the development of standardized evaluation protocols and benchmark datasets is essential for fair comparisons and objective assessments of different methods [3].

Despite the significant progress made in multimodal video indexing and retrieval, several challenges and unresolved problems persist. Future work should focus on addressing these challenges and exploring new directions for research. These include enhancing multimodal fusion techniques, improving keyframe extraction accuracy and adaptability, advancing semantic understanding of video content, enhancing scalability and retrieval performance, and establishing standardized evaluation frameworks.

In this paper, we provide a comprehensive review of recent advancements in multimodal video indexing and retrieval. We analyze and summarize the contributions of various studies in addressing the aforementioned challenges. By examining the proposed methods and techniques, we identify the key issues that need further exploration and discuss potential avenues for future research. Through this review, we aim to provide insights and directions for researchers in the field of multimodal video indexing and retrieval.

The rest of the paper is structured as follows. Section II describes the latest advances in video indexing and retrieval. Section III discusses the presented papers, which address various challenges and propose innovative solutions for multimodal video indexing and retrieval, including multimodal fusion, keyframe extraction, semantic analysis, scalability, and evaluation. Along with the discussion about their contribution, this section comments ongoing research that is needed to tackle the unresolved issues and improve the efficiency of video retrieval systems. Concluding remarks are highlighted in the last section.

2. Latest advances in video indexing and retrieval

The field of video indexing and retrieval has witnessed significant advances in recent years, driven by the exponential growth of video data and the increasing demand for efficient video search and analysis. Researchers have made notable progress in various aspects of video indexing and retrieval. In the domain of multimodal fusion, Podlesnaya and Podlesnyy have explored deep learning-based techniques for semantic video indexing and retrieval [15], while Hamroun et al. introduced a novel method employing machine learning and semi-automatic annotation on large video collections [16]. For keyframe extraction, Priya, Lakshmi, and Domnic have contributed to the field by developing shot-based keyframe extraction methods tailored for ecological video indexing and retrieval [13]. In the realm of semantic analysis, Safadi and Quénot proposed re-ranking by local re-scoring techniques to enhance video indexing and retrieval results [12]. Lee, Park, and Yoo addressed scalability challenges in surveillance video indexing and retrieval by introducing a data cube model [5]. Moreover, Wang, Huang, Wang, Zhang, and Tian have discussed video indexing and retrieval based on key frame extraction, providing insights into effective techniques for organizing and accessing video content [3]. Multimodal fusion techniques have been developed to effectively integrate and exploit the complementary information from different modalities such as text, audio, and visual features. Keyframe extraction methods have improved the accuracy of identifying representative frames that summarize video content. Semantic analysis approaches have aimed to bridge the gap between low-level visual features and high-level semantic concepts for more meaningful video annotation and retrieval. Scalability challenges have been addressed through the development of optimized indexing structures and parallel processing techniques. Furthermore, standardized evaluation protocols and benchmark datasets have been established to facilitate fair comparisons and objective assessments of video indexing and retrieval methods. These recent advancements lay a solid foundation for further research and innovation in this field, opening up new opportunities to enhance the efficiency and effectiveness of video indexing and retrieval systems.

The semantic pathfinder, a generic approach for video indexing that aims to bridge the semantic gap in multimedia archives is presented in [1]. The approach is based on the understanding that videos are the result of an authoring process. The semantic pathfinder utilizes various detector types, multimodal analysis, hypothesis selection, and machine learning to select the most appropriate path through content analysis, style analysis, and context analysis. Experimental results using a lexicon of 32 semantic concepts demonstrate the effectiveness of the semantic pathfinder in enabling generic video indexing and validating the value of the authoring metaphor in the indexing process. However, the precision for

some concepts at 100 performances is still relatively low, which may be sufficient for tasks involving selecting illustrative footage but not accurate retrieval tasks. Further improvements are needed to enhance the accuracy of retrieval tasks.

Paper [16] introduces a novel method for multimodal indexation using machine learning and semi-automatic annotation. The authors propose a genetic algorithm-based approach to detect concepts from the text in videos. They enrich basic concepts using a method called DCM and utilize semantic and enriched concepts to improve multimodal indexation and construct an ontology. The method's effectiveness is evaluated on the TRECVID 2015 dataset.

Paper [15] focuses on deep learning-based approaches for semantic video indexing and retrieval. The system utilizes features extracted by convolutional neural networks (CNNs) to enable efficient indexing and retrieval. The paper provides implementation details and empirical evaluation results to showcase the effectiveness of the proposed approach.

In [14], the authors propose a framework for multimodal video indexing and retrieval using a similarity measure called shrinkage-optimized directed information assessment (SODA). The directed information captures the direction of information flow in videos and is applied to audio-visual features over successive frames. This approach leverages the natural characteristics of videos and aims to improve indexing and retrieval performance.

Lakshmi and Domnic in [13] present a new automatic shot-based keyframe extraction technique for video indexing and retrieval applications. They utilize feature extraction and continuity value construction steps of shot boundary detection to cluster frames into shots. Keyframes are then extracted based on inter-cluster similarity analysis. This method provides a more efficient and accurate way of representing video content for indexing and retrieval purposes.

A re-ranking method to enhance the performance of semantic video indexing and retrieval systems is proposed in [12]. The proposed method re-evaluates the scores of shots based on their homogeneity and their relationship with the overall video content. By considering the temporal sequence and the distribution of video shots, this method improves the ranking accuracy and provides a framework for re-ranking.

A specific approach for content-based video indexing and retrieval using the Correspondence-Latent Dirichlet Allocation (corr-LDA) probabilistic framework is presented in [11]. The method utilizes the semantic relations between video content and text to improve indexing accuracy. The authors focus on the audio components of video recordings and compare the results with a support vectore machines (SVM)-based approach.

In [10], a Knowledge Distillation framework called Distill-and-Select (DnS) for video indexing and retrieval is introduced. The framework involves training Student Networks with different retrieval performance and computational efficiency trade-offs. A Selector Network is also trained to direct samples to the appropriate student network at test time. This approach maintains high retrieval performance while ensuring computational efficiency.

Paper [9] presents a probabilistic Bayesian belief network (BBN) method for the automatic indexing of excitement clips in sports video sequences. The excitement clips are extracted using audio features, and multiple subclips corresponding to different events are detected and classified. The proposed method utilizes hierarchical classification and provides an efficient approach to indexing sports videos.

Paper [8] proposes a novel algorithm for shot detection and keyframe determination based on gradient fields. The algorithm aims to find a compact set of keyframes that accurately represents a video segment. By considering the structural similarity of frames, the proposed approach provides an effective method for content-based video indexing and retrieval.

Paper [7] presents an approach for automatic lecture video indexing based on video OCR technology. It introduces a video segmenter for slide video structure analysis and a weighted DCT-based text detector. The text occurrence information and analyzed text content are used for indexing. The proposed method leverages video OCR technology to automate the indexing process, improving the accessibility and searchability of lecture videos.

Paper [6] presents a novel framework for video surveillance indexing and retrieval. The proposed framework consists of three main modules: pre-processing, query processing, and retrieval processing. It aims to overcome the challenges of video surveillance data management and enhance the efficiency and effectiveness of video retrieval systems in surveillance applications.

Paper [5] proposes a data cube model, SurvCube, for the multi-dimensional indexing and retrieval of surveillance videos. The proposed method utilizes the data cube structure to provide multi-dimensional analysis of interesting objects in surveillance videos based on chronological view, events, and locations. It offers functionalities such as retrieval at different levels of abstraction, tracing object trajectories, and summarization of surveillance videos.

Paper [4] addresses the challenges of managing large-scale video databases by utilizing spatial indexing and querying of field-of-views (FOVs). The authors propose a class of new R-tree-based index structures that effectively harness FOVs' location, orientation, and view-distance properties. The proposed method optimizes filtering and query optimization by considering the geographical properties of FOVs, providing efficient indexing and retrieval of geo-tagged video databases.

Paper [3] presents a query-by-concept approach for video retrieval. The proposed method focuses on key frame extraction and achieves better results compared to existing methods. The evaluation shows a mean average precision (MAP) of 0.68 for the video retrieval model.

3. Discussion

The presented papers cover a wide range of approaches and techniques for multimodal video indexing and retrieval. These studies contribute to the advancement of video retrieval systems by addressing various challenges and proposing innovative solutions. However, there are still several issues and unresolved problems that need to be addressed in future research.

One of the challenges in video indexing and retrieval is the effective utilization of multimodal information, including text, audio, and visual content as show in Figure 1 [14]. Several papers propose methods for extracting and integrating these modalities to improve the indexing and retrieval process [14, 15, 16]. However, achieving a robust and accurate multimodal fusion remains an ongoing research problem.

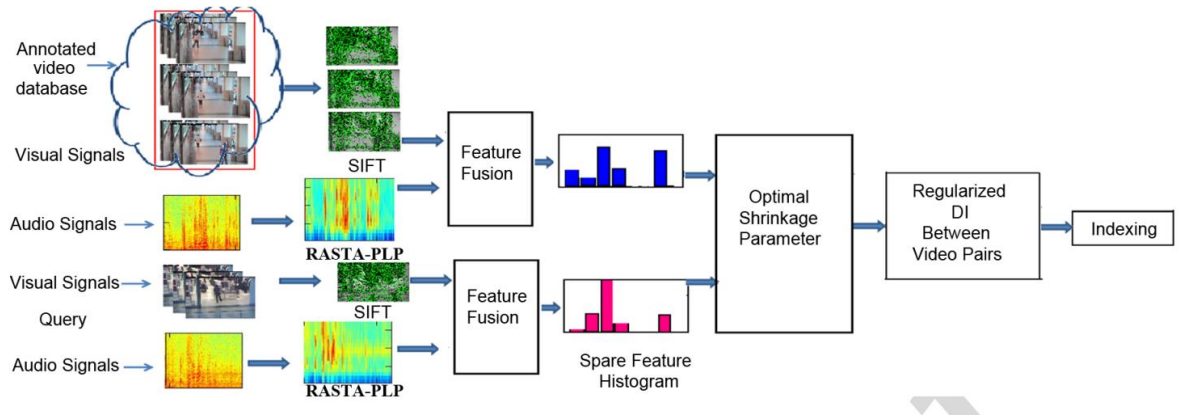


Figure 1: Block diagram of shrinkage optimized directed information (SODA) for fusion of audio and visual features for video indexing

Another area of research focuses on the extraction of keyframes and shots from videos as presented in Figure 2 [8]. Keyframe extraction plays a crucial role in representing the content of a video efficiently. While the proposed methods demonstrate advancements in automatic keyframe extraction, there is still room for improvement in terms of accuracy and adaptability to different video genres and styles [8, 13].

Semantic analysis and annotation of video content are also addressed in paper [5]. These approaches aim to bridge the semantic gap between the low-level visual features and the high-level semantic concepts associated with videos. However, achieving precise and comprehensive semantic understanding of video content remains a challenge due to the inherent complexity and subjectivity of video semantics.

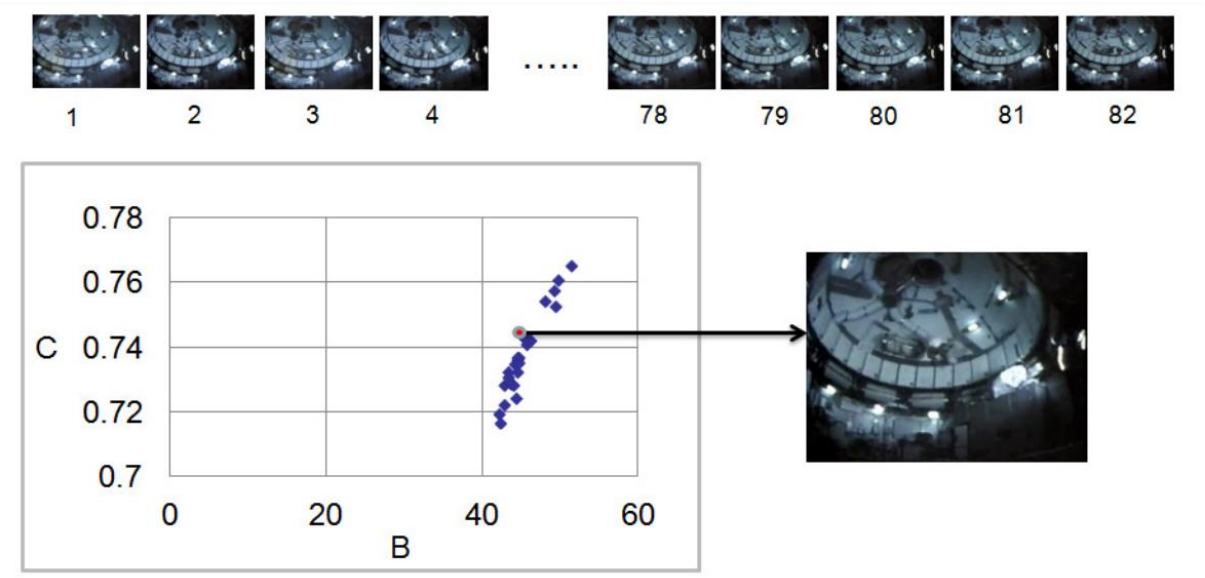


Figure 2: Illustration of key-frame

Efficient indexing and retrieval of large-scale video databases pose significant challenges [4, 5]. The scalability and retrieval performance of video retrieval systems need to be further improved to handle the ever-increasing volume of video data. Additionally, the integration of spatial and temporal information in video indexing and retrieval remains an area of research that requires further exploration.

Another important aspect is the evaluation and benchmarking of video retrieval systems. While some papers mention the evaluation of their proposed methods using specific datasets [16], there is a need for standardized evaluation protocols and benchmark datasets that cover diverse video genres and capture real-world scenarios.

In conclusion, the discussed papers contribute valuable insights and advancements in multimodal video indexing and retrieval. However, several challenges and issues, such as multimodal fusion, keyframe extraction, semantic analysis, scalability, and evaluation, require further attention and research efforts in order to develop more robust and efficient video retrieval systems.

4. Conclusions

In conclusion, the presented papers on multimodal video indexing and retrieval have made significant contributions to the field by proposing innovative methods and techniques. These studies address various challenges related to multimodal fusion, keyframe extraction, semantic analysis, scalability, and evaluation. However, there are still unresolved issues and opportunities for future work.

The challenges of effectively integrating multimodal information, including text, audio, and visual content, remain open research problems. Future work can focus on developing advanced fusion techniques that leverage the complementary nature of different modalities to improve the accuracy and robustness of video indexing and retrieval systems.

Keyframe extraction continues to be an important task for representing video content efficiently. Further research is needed to enhance the accuracy and adaptability of keyframe extraction methods across different video genres and styles. Additionally, exploring novel approaches that consider temporal dynamics and semantic relevance can lead to more comprehensive and meaningful keyframe representations.

Semantic analysis and annotation of video content remain challenging due to the complexity and subjectivity of video semantics. Future work can explore advanced machine learning and deep learning techniques to bridge the semantic gap and enable a more precise and comprehensive understanding of video content. Developing ontologies and knowledge graphs that capture the semantic relationships among video concepts can also contribute to more effective indexing and retrieval.

The scalability and retrieval performance of video retrieval systems require continuous improvement to handle the ever-increasing volume of video data. Future research can focus on developing efficient indexing structures, optimization algorithms, and parallel processing techniques to enhance the scalability and retrieval speed of large-scale video databases.

The evaluation and benchmarking of video retrieval systems need standardized protocols and benchmark datasets that cover diverse video genres and real-world scenarios. Future work can focus on developing comprehensive evaluation frameworks and datasets to enable fair comparisons among different methods and facilitate progress in the field.

References:

[1] Blanken, Henk M., et al., eds. “Multimedia retrieval.” Springer Science & Business Media. (2007).
[2] H. Yang, M. Siebert, P. Luhne, H. Sack and C. Meinel, "Lecture Video Indexing and Analysis Using Video OCR Technology," 2011 Seventh International Conference on Signal Image Technology & Internet-Based Systems, Dijon, France, pp. 54-61, doi: 10.1109/SITIS.2011.20. (2011).
[3] Wenshi Wang, Zhangqin Huang, Weidong Wang, Shuo Zhang, and Rui Tian. “Video Indexing and Retrieval based on Key Frame Extraction.” [J]. Int J Performability Eng, 14(8): 1824-1832. (2018).
[4] Lu, Y., Shahabi, C. and Kim, S.H. “Efficient indexing and retrieval of large-scale geo-tagged video databases.” Geoinformatica 20, 829–857. (2016).
[5] H. Lee, S. Park and J. -H. Yoo. "A data cube model for surveillance video indexing and retrieval." 2013 International Conference on Signal Processing and Multimedia Applications (SIGMAP), Reykjavik, Iceland, pp. 163-168. (2013).
[6] Fereshteh Falah Chamasemani, Lilly Suriani Affendey, Norwati Mustapha and Fatimah Khalid. “A Framework for Automatic Video Surveillance Indexing and Retrieval.” Research Journal of Applied Sciences, Engineering and Technology 10(11): 1316-1321. (2015).
[7] H. Yang, M. Siebert, P. Luhne, H. Sack and C. Meinel. "Automatic Lecture Video Indexing Using Video OCR Technology." *2011 IEEE International Symposium on Multimedia*, Dana Point, CA, USA, pp. 111-116, doi: 10.1109/ISM.2011.26. (2011).
[8] Asatryan, David, and Manuk Zakaryan. "Novel approach to content-based video indexing and retrieval by using a measure of structural similarity of frames." International Journal" Information Content and Processing 2.1, 71-81. (2015).
[9] Kolekar, Maheshkumar H. "Bayesian belief network based broadcast sports video indexing." Multimedia Tools and Applications 54, 27-54. (2011).
[10] Kordopatis-Zilos, Giorgos, et al. "DnS: Distill-and-Select for Efficient and Accurate Video Indexing and Retrieval." International Journal of Computer Vision 130.10, 2385-2407,(2022).
[11] Iyer, Rahul Radhakrishnan, et al. "Content-based video indexing and retrieval using corr-lda." arXiv preprint arXiv:1602.08581. (2016).
[12] Safadi, Bahjat, and Georges Quénot. "Re-ranking by local re-scoring for video indexing and retrieval." Proceedings of the 20th ACM international conference on Information and knowledge management. (2011).
[13] Priya, GG Lakshmi, and S. Domnic. "Shot based keyframe extraction for ecological video indexing and retrieval." Ecological informatics 23.107-117. (2014).
[14] Chen, Xu, Alfred O. Hero III, and Silvio Savarese. "Multimodal video indexing and retrieval using directed information." IEEE Transactions on Multimedia 14.1. 3-16. (2011).
[15] Podlesnaya, Anna, and Sergey Podlesnyy. "Deep learning based semantic video indexing and retrieval." Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016: Volume 2. Springer International Publishing. (2018).
[16] Hamroun, Mohamed, Karim Tamine, and Benoît Crespin. "Multimodal video indexing (mvi): A new method based on machine learning and semi-automatic annotation on large video collections." International Journal of Image and Graphics 22.02. 2250022. (2022).

VR as a Tool for EVs Maintenance Training

Naile Emini[1], Konstantin Veljanovski[2] and Nikola Rendevski[3]

[1,2,3] *University St Kliment Ohridski, 1 Maj nn, 7000 Bitola, R. North Macedonia*

emini.naile@uklo.edu.mk; konstantin.veljanovski@uklo.edu.mk; nikola.rendevski@uklo.edu.mk

Abstract:
The fast adoption of electric vehicles (EVs) into the automotive scene requires the development of novel maintenance training approaches. Due to its ability to create immersive and interactive learning environments, virtual reality (VR) is a powerful training tool across various domains. Trainees can visually interact with various components of an EV, practice maintenance methods, and replicate real-world scenarios in a risk-free environment by using VR's capabilities. The main benefits of using VR for EV maintenance training include hands-on practice, the capacity to replicate complicated EV systems, scalability for educating several personnel at the same time, and the possibility of interactive and adaptable learning modules. VR may provide a new sensory experience by merging haptic input and realistic visual and auditory cues, boosting trainees' spatial knowledge. The research also underlines the potential benefits of VR training, such as increased engagement, cost-effectiveness, scalability, and the ability to standardize training content. This study calls for the use of VR technology as a teaching tool for EV repair, emphasizing its potential to alter the way personnel are trained and skilled in this changing automotive scene. The findings highlight the importance of preparing a skilled workforce capable of meeting the demands of sustainable mobility.

Keywords:
VR, training, EV, sustainable, electric vehicle

1. Introduction

The early 2000s marked the rise in popularity of electric vehicles. In 2004, Tesla Motors began production of the Tesla Roadster, which was delivered to customers in 2008. The Roadster was the first fully electric vehicle powered by a Li-ion battery that was deemed road-ready, with a driving range of 320 km on a single charge. In December 2010, the Nissan Leaf appeared on the Japanese and American markets. It was the first modern electric vehicle without any emissions that fell under Nissan's "Zero Emission" brand. When the Li-Ion battery arrives at the repair center, it typically goes through the following phases: Diagnostics and testing, where the battery gets checked for damage and tested for faults. The diagnostics team will determine the next steps that need to be taken. Next, a team of specialists trained to work with high-voltage components repairs or replaces the parts highlighted by the diagnostics team, and finally, the battery assembly is reattached to the vehicle. Li-ion batteries contain toxic metals such as cobalt, nickel, and manganese, and so the process of training EV battery repair technicians can involve risks. A technician without the proper training for dealing with EVs can experience life-threatening risks. The following are some of the things that may go wrong during the repair process: Wrong Diagnosis: Without the proper knowledge of dealing with EVs, a technician can easily make a wrong diagnosis, which leads to a loss of money and time. Safety Concerns: EVs have high-voltage systems, and if the technician isn't prepared on how to safely deal with the situations, he might be in danger. High voltage may lead to electric shock, fire, and a number of other life-threatening situations. Due to all these stated reasons, training in VR is an ideal solution for all companies that manufacture EVs as well as all repair shops.

2. Health risks

The increase in the number of electric vehicles (EVs) being sold is causing a greater need for a special training program to work on EVs in the future. This section will explain why the training sector

is facing new challenges. The main difference between electric cars and regular cars is that electric cars have a high-voltage battery. High voltage means a certain range of electric current that is considered dangerous. According to European rules (ECE R100) [2] and American rules (SAE J1766-2014) [3], for alternating current (AC), a high voltage is between 30 volts and 1000 volts. For direct current (DC), a high voltage is between 60 volts and 1500 volts. The battery in electric vehicles (EVs) usually has a voltage of 355 volts (DC) [4]. However, some EVs, such as the Nissan Leaf [5], have a higher voltage of 360 volts [6]. Based on [7], car companies are designing batteries for electric vehicles (EVs) that can have voltages of up to 850 V DC in the future. Therefore, using high voltages in cars can be dangerous because you could get an electric shock or burn yourself. Electric vehicles (EVs) have built-in safety features to prevent people from experiencing an electric shock. One of these safety measures is the service disconnect. This measure will turn off the high-voltage loop. However, employees in the vehicle repair industry are required by law to attend specialized training programs (e.g., [1], [8]). For instance, they must understand where the safety features are located in the car and how they operate. There are some dangers to your health when working on electric vehicles because they have a lot of electrical parts and a high-voltage system. So, you need to take extra precautions. The health dangers for people working with EVs are not small. For instance, if a car mechanic accidentally touches a cable with too much electricity or a battery with a problem, they could get a shock. When you get an electrical shock, you might have a dangerous heart problem called ventricular fibrillation or get burned. This means it is very important for the people who work on electric vehicles (EVs) to have a lot of knowledge and skills. This includes, among other things, knowing where the parts are located, understanding the dangers of high voltages and high currents for our health, knowing how to use protective equipment correctly, knowing how to provide first aid, following safety rules (for example, wearing protective gear) and turning off the powerful electrical system.

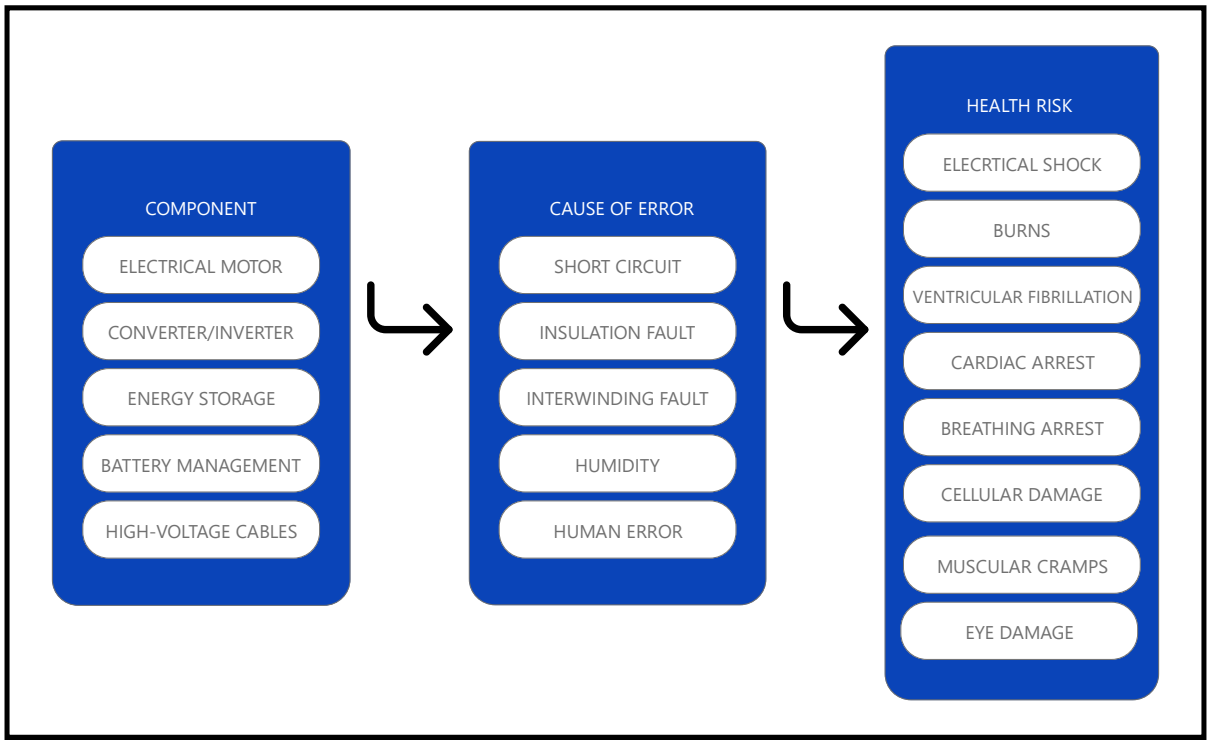


Figure 1. Health risk of some selected components

3. Advantages of VR in Electrical Vehicle training and maintenance

The main advantage of VR in EV maintenance is the opportunity to provide specialized training for technicians. VR can help technicians upgrade their skills and be prepared for real-world challenges. Replicated scenarios of the product lines enable the identification and diagnosis of issues without the

need to stop production. This decreases the wait time and improves the quality of the production capacity. The benefits of VR training are the following:

- Increased safety and ability for the technicians to get familiar with the various tools and machines without the risk of injury,
- Increased efficiency that comes from VR being able to simulate various scenarios and states that mandates usage of different types of tools in the correct way,
- Saves on costs that are needed to get all of the tools required for a trainee to successfully deal with all possible situations,
- The ability to better understand and repeat the process, which leads to greater self-confidence,
- The ability to create and replicate situations that are more complex or that are impossible to replicate in the real world helps the technicians better deal with the rare scenarios that can happen.
- Technicians can train in different locations without needing to be in the same place at the same time, which is a benefit for a company that has multiple training locations.

4. Training procedures in VR

The training can be grouped into general lectures for safety and equipment usage as well as repair procedures that can be standardized and/or custom tailored to a company's needs based on their training program. VR allows for the transmission of clear and precise instructions (Figure 2) to the trainee to make sure that they follow the training procedure while maintaining safety practices.

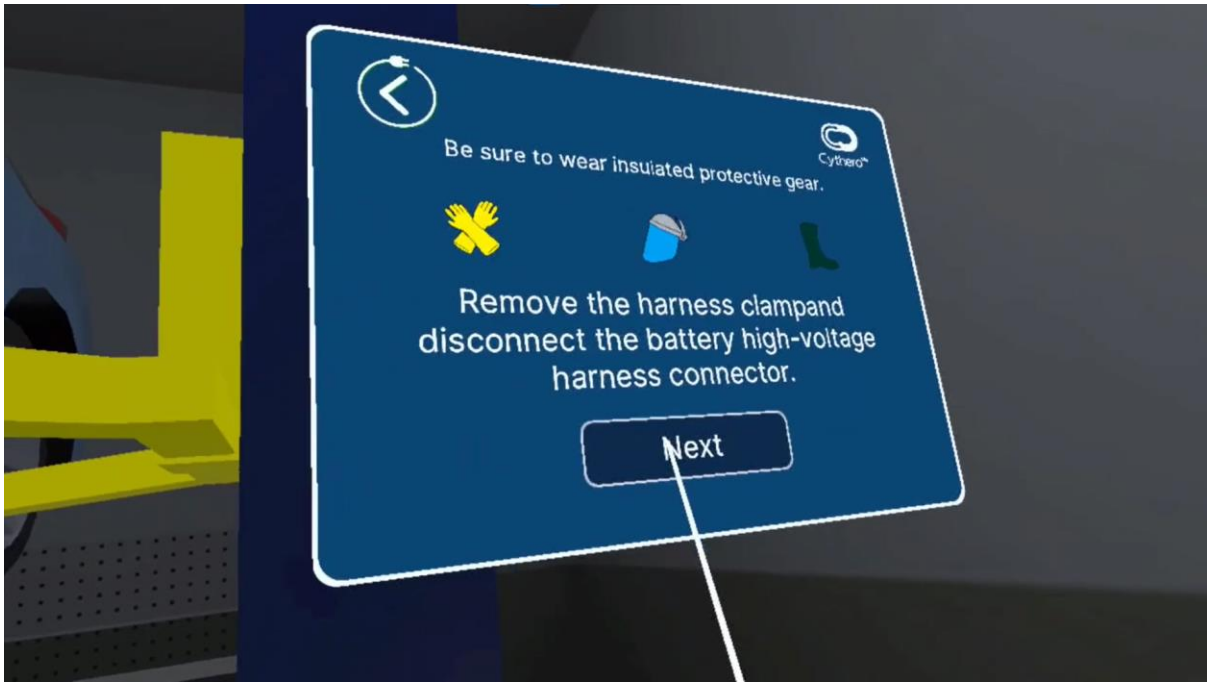


Figure 2. Instructions in VR

The training in VR is accomplished by using visual and auditory cues such as highlights (Figure 3), arrows, pings to guide the trainee to interact with the environment in order to complete a task.

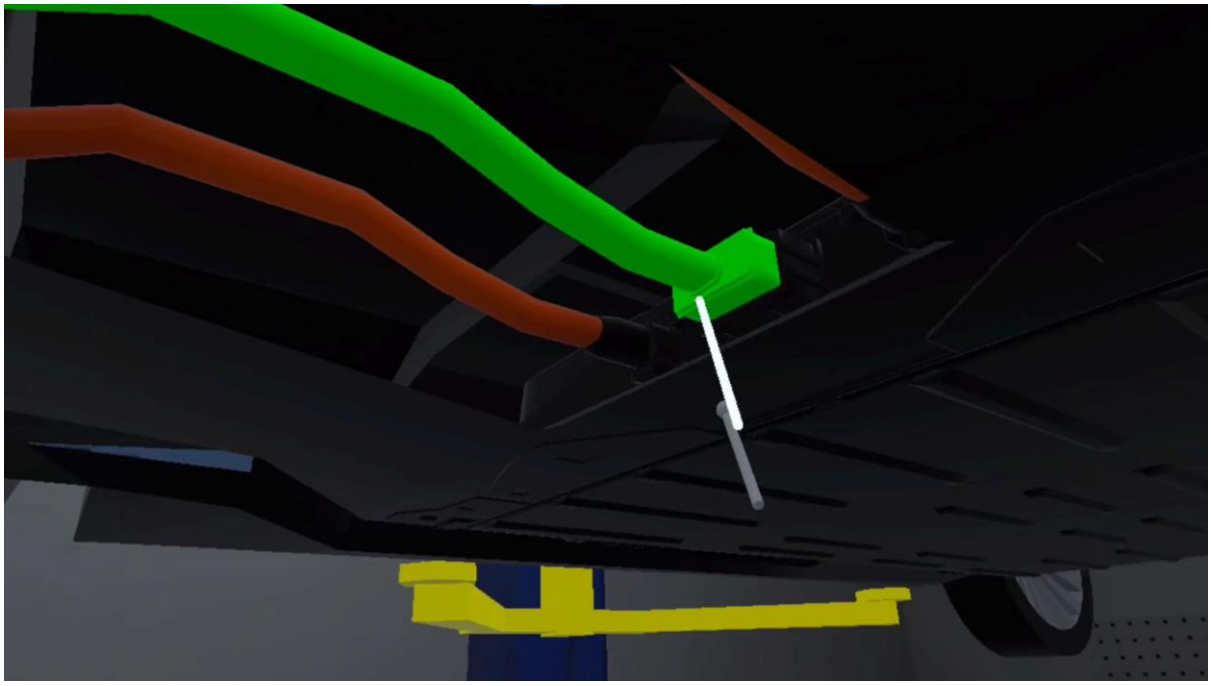


Figure 3. Visual Cues

General lectures can have the trainee learn about their equipment through quizzes, prompts, and interactions with their environment such as picking the right equipment for a job, identifying faulty equipment and learning about general hazards and how to avoid them.

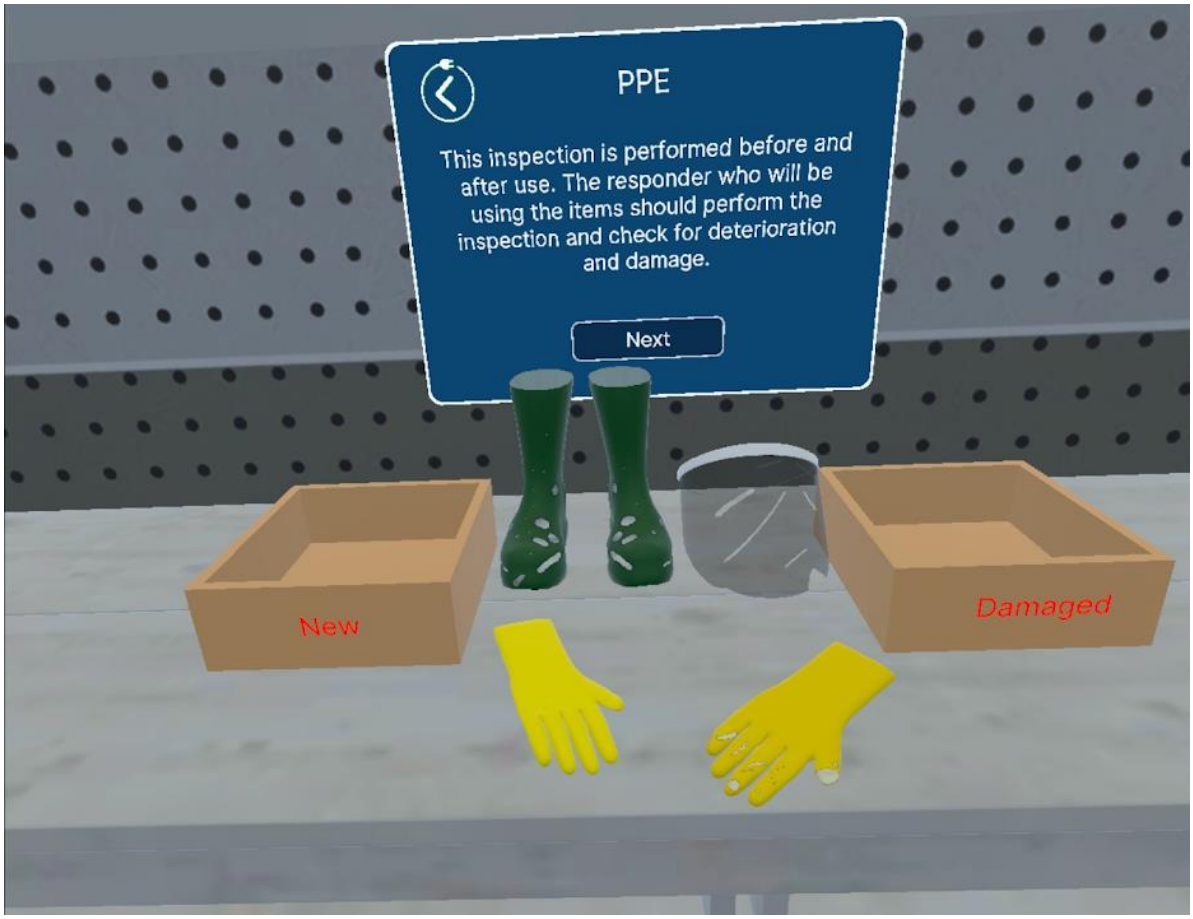


Figure 4.Equipment identification

5. Conclusions

As both the EV and VR industries grow and become the main stream, most companies should look into implementing VR into their training programs, as it can wildly improve the training time needed for technicians to be field-ready, cut down on the cost of training, and avoid the risks of personnel harm. VR can be a great tool for companies to do decentralized training of new and existing technicians, as it doesn't require the trainer to be at the same place as the trainee. Training applications can be equipped with a system to record and save training sessions for further review and performance tracking, which can help pick out outstanding trainees as well as help out trainees struggling with certain topics.

References:

[1]Kultusministerkonferenz, „Rahmenlehrplan für den Ausbildungsberuf Kraftfahrzeugmechatroniker und Kraftfahrzeugmechtaronikerin," KMK, 2013.

[2] Economic Commission for Europe of the United Nations (UN/ECE), ECE R100: Uniform provisions concerning the approval of vehicles with regard to specific requirements for the electric power train, Official Journal of the European Union, 2011.

[3] Society of Automotive Engineers, SAE J1766: Recommended Practice for Electric, Fuel Cell and Hybrid Electric Vehicle Crash Integrity Testing, 2014.

[4] B. Scrosati, J. Garche and W. Tillmetz, Advances in Battery Technologies for Electric Vehicle, Cambridge: Elsevier Science & Technology, 2015.

[5] B. Crothers, "Tesla 'Most Popular' U.S. Electric Vehicle in Q1, While Norway Leads Globally, Says IHS," Forbes, July 13, 2015. [Online]. Available: https://www.forbes.com/sites/brookecrothers/2015/07/13/tesla-most-popular-u-s-electric-vehicle-in-q1-while-norway-leads-globally-says-ihs/?sh=29c8e196380e [Accessed: July 13, 2015].

[6] Nissan, "2015 LEAF – First Responder`s Guide," 2014.

[7] ZVEI - German Electrical and Electronic Manufacturers´ Association, "Voltage Classes for Electric Mobility," Frankfurt am Main, Germany, 2013.

[8] Deutsche Gesetzliche Unfallversicherung, Training for work on vehicles with high voltage systems, Berlin, Germany: DGUV, 2012.

[9] Brian T. Horowitz, "How AR & VR in Healthcare Enhances Medical Training", Health Tech Magazine, December 15, 2022, Available: https://healthtechmagazine.net/article/2022/12/ar-vr-medical-training-2023-perfcon [Accessed: December 15, 2022]

Automation and Monitoring on Integration ETL Processes while Distributing Data

Aneta Trajkovska, Tome Dimovski, Ramona Markoska, Zoran Kotevski

University "St. Kliment Ohridski" – Bitola, Faculty of Information and Communication Technologies, Studentska bb, Bitola, Republic of Macedonia

aneta.trajkovska@uklo.edu.mk, tome.dimovski@uklo.edu.mk, ramona.markoska@uklo.edu.mk, zoran.kotevski@uklo.edu.mk

Abstract:
This paper describes an approach for the automation of processes, leveraging the capabilities of various technologies and diligent data monitoring. The continuous advancement of technology presents new opportunities for development and offers a superior and more precise alternative for decision-making across various domains, not only IT areas. From an analytical standpoint, such a system is characterized by its ability to operate in a manner where multiple computers or computing devices on a network share various components responsible for distinct operations. By dividing the workload, these devices coordinate their capabilities to efficiently accomplish the intended task, resulting in enhanced performance compared to a single device handling it alone. This is precisely what SnapLogic, as an iPaaS platform, excels at. It simplifies management through APIs, which are utilized in its ETL (Extract, Transform, and Load) processes, thereby expanding communication possibilities with other systems. By using this integration platform, businesses can eliminate slow and error-prone manual methods while achieving a high level of automation for their processes. One notable feature of this platform is its process orchestration capability. It empowers users to monitor and manage data pipelines originating from various platforms and systems, allowing for seamless application integrations and API calls, all executed efficiently within a single system.

Keywords:
Integrations, IPaaS, ETL (Extract, Transform and Load) processes, Snaplogic, pipelines, API, AWS IaC, AWS CloudFormation, ELK (Elasticsearch, Logstash and Kibana)

1. Introduction

With the exponential progress of technologies and the growing need to distribute a variety of data formats (while they are easily accessible and presented with good visualization), there is also the need to use processes that that enhance efficiency while providing a high level of automation for various operations. The ever-growing volume of data within organizations necessitates finding ways to ensure that their transactions are executed in a timely manner, enabling them to be processed and transferred to the appropriate destination promptly. The systems themselves have the capability to coordinate the execution of operations and tasks in a timely manner, thereby enhancing efficiency and significantly reducing the time required to complete tasks. This is precisely one of the objectives of integration iPaaS (Integration Platform as a Service) systems, ETL(Extract, Transform and Load) processes, and distributed systems.

iPaaS is a set of automated tools that integrate software applications that are deployed in different environments [1]. Typically, an iPaaS platform provides pre-built connectors, business rules, maps and transformations that facilitate the development of applications and orchestrate integration flows [2].

The utilization of ETL processes facilitates a straightforward and standardized approach to meet the integration requirements, enabling large organizations and their businesses to process data effectively and achieve the desired outcomes. ETL, which stands for extract, transform and load, is a data integration process that combines data from multiple data sources into a single, consistent data store that is loaded into a data warehouse or other target system [3].

In this paper it was used automation on integration, infrastructure in AWS, and monitoring on ETL processes. The remaining of this paper is organized as follows. Section 2 presents the challenges of utilization of the integration processes in the organizations. In section 3 are elaborated: ETL processes and their advantages, Snaplogic as an integration platform that is running ETL processes ELK components and their functions, AWS CloudFormation service as an Infrastructure as a Code (IaC). Section 4 describes the proposed solution with utilization of the mentioned services and technologies, while section 5 concludes the paper with a summary of the research.

2. The challenge of using integration processes in the organization

The utilization of integration processes within organizations poses certain challenges that need to be addressed effectively. In the past, companies integrated their business processes through custom programming, enterprise middleware or enterprise application integration (EAI) implementations, such as service-oriented architecture (SOA). These integration solutions worked but were expensive and time-consuming to create [4].

iPaaS offers a unified and streamlined solution for data exchange and integration across all applications within an organization, regardless of whether they are hosted on-premises or in the cloud. This enables seamless and consistent data integration processes throughout the entire organization (Figure 1).

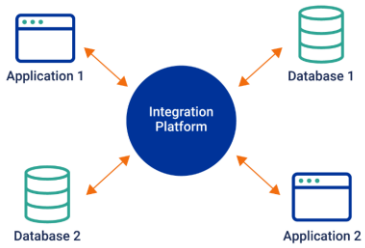


Figure 1: How integration platforms drive the flow of data between applications and databases

An organization can implement iPaaS more rapidly and cost-effectively compared to traditional technologies, offering comprehensive integration across various use cases such as cloud-to-cloud, cloud-to-on-premises, and on-premises-to-on-premises scenarios. iPaaS plays a vital role in driving digital transformation by providing several notable advantages [5], [6],[7],[8]:

- Faster deployment and cost efficiency,
- Business and IT agility,
- Flexibility for growth and scalability,
- Enhanced Connectivity,
- Simplified Management and Monitoring.

3. Automation of ETL integration processes using Snaplogic platform, AWS infrastructure and ELK monitoring

Increasing operational efficiency through streamlined data integration brings a lot of advantages. Exploring the integration of ETL processes with the Snaplogic platform, AWS infrastructure, and ELK monitoring yields a transformative approach to data management and optimization.

3.1. ETL (Extract, Transform and Load)

Modern information technology relies extensively on ETL processes to achieve enhanced data processing and secure data distribution. ETL stands for Extract, Transform, and Load, and it refers to the procedure of transferring data from its source to a designated data warehouse [9], [10].

During the extraction phase, data is gathered and extracted from various sources such as databases, APIs, or files. The data is then subjected to transformation, which involves cleaning, filtering, and structuring it to meet the requirements of the target system or data model. Lastly, the transformed data is loaded into the data warehouse, where it can be efficiently stored, organized, and analyzed (Figure 2).

Figure 2: Phases of ETL processes

The utilization of ETL processes plays a crucial role in modern information technology, enabling organizations to effectively manage and leverage their data assets for improved decision-making, analytics, and business insights [11],[17].

ETL and ELT (Extract, Load and Transform) are both viable solutions for data movement and transformation but are best suited to different business use cases [12].

3.2. Snaplogic – integration platform that runs automated ETL processes

Today, organizations are driven to achieve enhanced performance in terms of speed and quality. SnapLogic is an integration platform (iPaaS-integration-platform–as-a-service) that is continuously coming into use when solving any business solutions. It is renowned as one of the fastest and most efficient platforms and services available in the market. SnapLogic's Intelligent Integration Platform (IIP) automates all phases of IT integration projects—design, development, deployment, and maintenance—whether in cloud or hybrid environments. It makes the data integration process faster and easier.

The SnapLogic iPaaS includes 700+ pre-built connectors, called Snaps and pre-built end-to-end integrations, called templates, for common business processes. Pipeline templates enable every user to quickly create the integrations for new, improved workflows and outcomes [15]. The power of this platform is of really great importance to many organizations, as it has huge possibilities that it offers, and at the same time, easy management, and an understandable interface for users, which makes it irreplaceable and gives the impression of an integration platform that will be constantly upgraded. It is a simple platform that consists of many powerful features [18], such as:

- Unification (One integration platform connects more than 700 snaps),
- Productivity (Self Service Integration for everyone),
- Modern (Flexible Deployment for Any – sized Enterprise),
- Artificial intelligence (for speed and quality),
- Interface – easy access through a browser to the platform itself.

The products it offers are a simple way (with little or no code) to automate applications and data integrations. Performs cloud data mobilization using visualization and ETL processes to further test them at the application level.

3.3. Snaplogic Architecture

Snaplogic's architecture is designed to address the complexities of modern data integration, providing a scalable and flexible platform to handle data from on-premises systems, cloud applications, APIs, and IoT devices. At its core, Snaplogic employs a unique visual design approach, empowering users with little to no coding experience to create complex data pipelines effortlessly. The architecture of Snaplogic revolves around three fundamental components [19]:

- Snaps (Snaps are pre-built connectors that facilitate the seamless integration of various data sources and targets),
- Snaplex (The Snaplex is a distributed execution grid that executes the data pipelines created using Snaps. It operates both in the cloud and on-premises, providing the flexibility to deploy integration workloads in the most suitable environment.),
- Snaplogic Manager (The Snaplogic Manager serves as the control center for designing, monitoring, and managing data integration processes.).

3.4. ELK (Elasticsearch, Logstash and Kibana)

ELK, an acronym for Elasticsearch, Logstash, and Kibana, is a comprehensive data processing and visualization platform designed to help organizations efficiently collect, store, analyze, and visualize vast amounts of data in real-time [20]. Elasticsearch's underlying data structure, based on JSON documents, enables lightning-fast searches, aggregations, and complex queries, making it an indispensable component for performing data analysis at scale [21]. Logstash serves as the data processing engine in the ELK Stack. It provides a flexible and extensible framework for collecting, parsing, and enriching data from various sources before forwarding it to Elasticsearch [22]. Kibana is a powerful data visualization tool, allows users to create interactive and customizable dashboards, reports, and visualizations, enabling them to explore data insights effortlessly [23].

3.5. AWS CloudFormation

AWS CloudFormation revolutionizes the way infrastructure is managed by treating it as code. Instead of manually configuring resources through the AWS Management Console or using AWS Command Line Interface (CLI), users can codify their infrastructure in JSON or YAML templates. This paradigm shift enables version control, peer review, and automated deployments, ensuring that infrastructure changes are consistent, documented, and auditable [24]. By leveraging declarative templates, resource drift detection, stack management, and integrations with various AWS services, CloudFormation streamlines the process of building scalable and consistent infrastructures, thereby enabling businesses to focus on innovation and growth while maintaining optimal control over their cloud environments.

4. Proposed solution

Within this paper, I will present a solution that offers a significant contribution to how the automation itself introduced into the process can be a key tool that enables simpler maintenance and optimization of resources. By harnessing the power of the SnapLogic platform and the businesses approach of utilization of the ETL processes, fortified by the scalability of AWS infrastructure, and enhanced by ELK monitoring (system logs, application logs, system metrics), we have engineered a comprehensive system that revolutionizes data workflows.

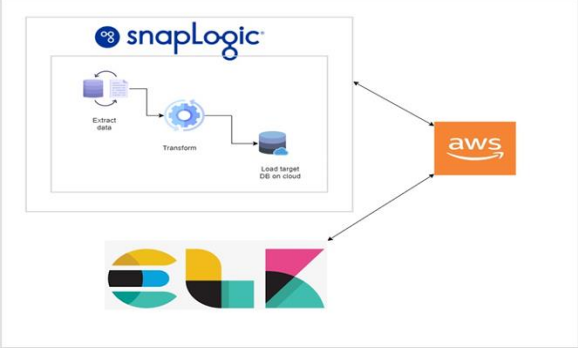


Figure 3: High overview of the proposed automated solution and used technologies

To present the overall setting of the solution in relation to the subject that is dealt with in this paper I have set up four stages of workflow: code preparation (as the first phase – staging phase), code development (the second phase of the automated ETL processes and the design of the infrastructure itself), code review (the third phase is to the review of the code itself within the prepared scripts and the verification of accuracy) and the code promotion (is the final stage of the prepared automated processes where the environments are active and the execution of the ETL processes runs smoothly).

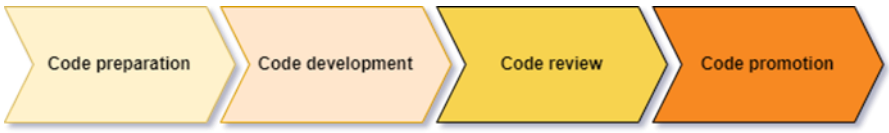


Figure 4: Workflow stages

4.1. Infrastructure of the solution in AWS

Based of the automation in the infrastructure I have used the AWS service CloudFormation, which is a service of Amazon Web Services (AWS) that enables the automatic creation and management of the infrastructure in the cloud by coding the infrastructure as code.

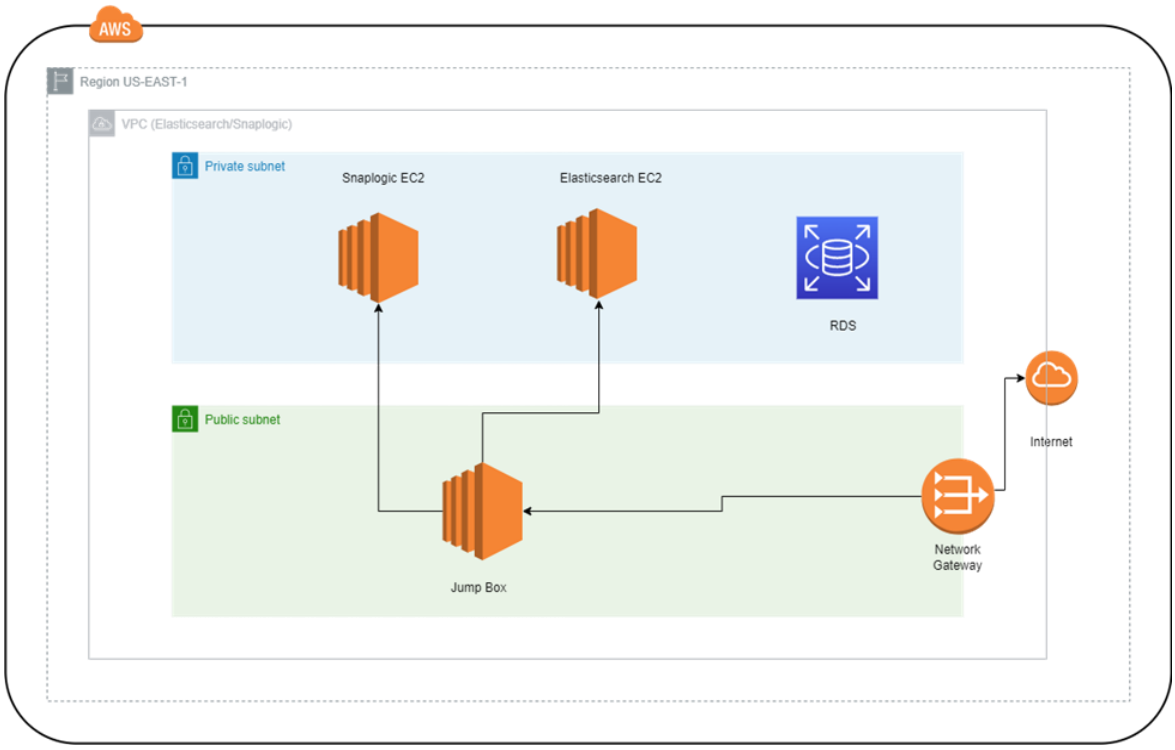


Figure 5: Infrastructure of the solution on AWS

The solution's architecture is outlined as follows: within the cloud environment, an initial work region is designated, leading to the creation of an isolated virtual private cloud within the broader AWS public cloud. This virtual private cloud is subdivided into two distinct subnets - one public and one private. A unique aspect is observed in the placement of two EC2 instances within the private network, and an additional two within the public network.

The essential contribution in selecting this configuration is in the communication between these instances. A Network Gateway instance is established to facilitate access to the Jump Box instance,

which resides alongside other components within the same public network. By utilizing the public IP address, a connection is established with the Jump Box instance. Subsequently, the connection proceeds from the Jump Box EC2 instance to access Snaplogic and Elasticsearch, both housed within the private network. This connection is facilitated via the employment of the private IP address.

Moreover, the proximity within the same subnet allows seamless access to the created RDS database, underscoring the cohesive integration of components within this comprehensive solution.

4.2. Automated ETL integration processes

The Snaplogic integration platform is very powerful and can achieve the processing of a large amount of data in just a few seconds, if the processes configured in it are created optimally.

The following scenario is composed of four processes, one of which is the main process, and the rest are sub-pipelines that are adequately called depending on where the data flow is. At the beginning, the process log information from the start of the execution. Then it pulls data from the API to further map and transform them, to finally record certain data in the MYSQL database.

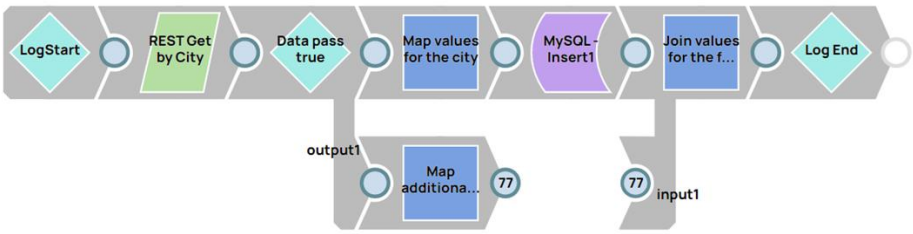


Figure 6: An integration process that aims to log into the database, read data from the API, transform it and write it to the database

The second scenario of the integration ETL process provides information about the state and activity of the snaplex itself. It works in a way that makes a call to a public API, where this time I use Snaplogic as the generator of the necessary data, making a call:

GET https://{pod_path}/api/1/rest/public/snaplex?{query_parameters}

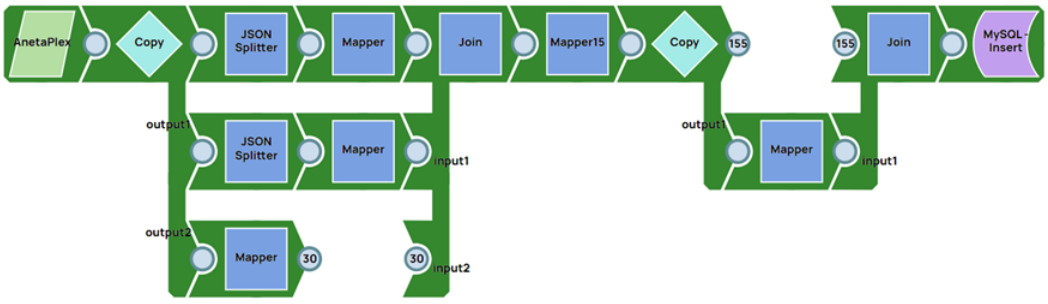


Figure 7: Successfully completed integration process that writes statistics data into the database from the information on the Snaplex

4.3. Monitoring of the data with utilization of ELK

ELK can be used to read metrics and monitoring data from a variety of sources, such as system logs, application logs, system metrics, and others. The monitoring itself is of great importance not only for the users of the data but also for the development of the business itself within the organizations.

Monitoring integration ETL processes is important to ensure that they are working properly, and that data is being transferred and transformed in a way that is appropriate for business needs. Monitoring can help detect process problems and warn them before serious consequences occur. This can be done by constantly monitoring the performance of the ETL process, as well as by reviewing the logs and occasionally testing the data.

Monitoring plays a pivotal role in elevating business efficiency. If process problems can be detected and resolved, it can lead to faster and better decision-making, as well as increase the quality of the data used to run the business.

Figure 8: Visualization of the temperature, pressure and air humidity changes for the city of Bitola in Kibana, based on the data in the table

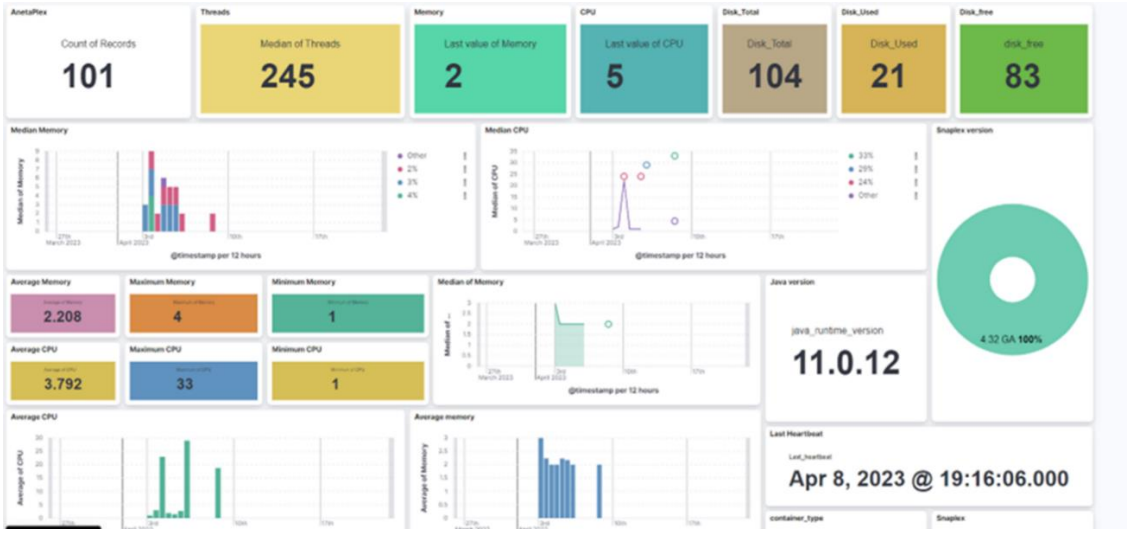


Figure 9: Dashboard that is showing data metrics from both tables SnaplexHealthTableInfoStats and SnaplexHealth

5. Conclusions

Automation control in process management is one of the most important areas of science and technological development. The automation and adaptability of integrated systems themselves

represent opportunities for better process control. The primary objective of this project was to enable the automation of the ETL process and implement monitoring for data flow through the Elastic Stack.

In this paper, were explored the numerous advantages of using automation in ETL integration processes within the AWS infrastructure. Through extensive research and practical implementation, we have observed significant benefits from utilizing this approach.

The project has successfully achieved its goals, allowing users to enhance database performance, improve infrastructure creation efficiency, and reduce the time required for manual data processing by implementing this solution.

The results of the research were successful, achieving the automation of the ETL process and providing real-time data flow monitoring through Elasticsearch and Kibana. The automated ETL process significantly improved data management efficiency, while the data flow monitoring empowered users with a comprehensive, real-time view of their data.

This project serves as compelling evidence that leveraging automated integration for ETL processes, along with AWS CloudFormation infrastructure, and ELK stack for monitoring, can yield improved data management, increased accuracy, and enhanced process efficiency.

References:

[1] Fred Churchville and Stephen J. Bigelow " What is iPaaS? Guide to integration platform as a service" Cloud infrastructure design and management, June (2021)
[2] Fred Churchville and Stephen J. Bigelow " Guide to integration platform as a service" Cloud infrastructure design and management, June (2021)
[3] IBM Cloud Education, IBM Cloud Education, " ETL (Extract, Transform, Load)", 14 December 2021
[4] IBM Cloud Education, " The enterprise integration challenge", 2022
[5] SnapLogic, iPaaS - Integration Platform as a Service Explained, https://www.snaplogic.com/blog/what-is-ipaas-integration-platform-as-a-service , 28 October 2023
[6] Asad Zulfahri, "16 Benefits of iPaaS That Show How It Can Transform Your Business", https://blog.hubspot.com/website/ipaas-benefits , 2021
[7] Stephen J. Bigelow, " Review these top enterprise iPaaS benefits and challenges", https://www.techtarget.com/searchapparchitecture/tip, May 2021
[8] IBM Cloud Education, " Benefits of iPaaS", 2022
[9] Snaplogic, The Modern ETL Guide, 16 May 2022
[10] IBM, ETL (Extract, Transform, Load), https://www.ibm.com/topics/etl
[11] Shanika Wickramasinghe, " Data 101: Extract, Transform & Load (ETL) Explained" January 26, 2023, https://www.splunk.com/en_us/blog/learn/etl-extract-transform-load.html
[12] Snaplogic, ETL vs. ELT: What's the Difference Between These 2 Processes, https://www.snaplogic.com/blog/etl-vs-elt-whats-the-difference , 13 September 2022
[13] Snaplogic documentation https://docs-snaplogic.atlassian.net/wiki/spaces/SD/overview
[14] Snaplogic, Data Pipeline, https://www.snaplogic.com/glossary/data-pipeline
[15] Snaplogic, " Intelligent Integration Platform", https://www.snaplogic.com/products/intelligent-integration-platform
[16] Martin Flower, ISBN 0321127420, Patterns of Enterprise application architecture, 11th October 2006
[17] Snaplogic, "ETL Data Pipelines", https://www.snaplogic.com/blog/all-about-etl-data-pipelines , May 2023
[18] Snaplogic, Enterprise Integration Architecture, Cloud Service Integration (snaplogic.com)
[19] Snaplogic, https://www.snaplogic.com/blog/cloud-integration-platform
[20] Elastic Stack, https://www.elastic.co/what-is/elk-stack
[21] Elastic Stack, https://www.elastic.co/what-is/elasticsearch
[22] Elastic Docs, https://www.elastic.co/guide/en/logstash/current/introduction.html
[23] Elastic Stack, https://www.elastic.co/what-is/kibana
[24] AWS Docs, https://docs.aws.amazon.com/AWSCloudFormation/latest/UserGuide/Welcome.html

Finding the Eigenspaces of a Matrix with GeoGebra

Sonja Mančevska[1], Elena Karamazova Gelova[2] and Mirjana Kocaleva Vitanova[2]

[1] *University "St. Kliment Ohridski"-Bitola, Faculty of Information and Communication Technologies-Bitola, Studentska nn, Bitola, R. North Macedonia*
[2] *University "Goce Delcev" Shtip, Faculty of computer science, Goce Delcev 89, Shtip, R. North Macedonia*

sonja.manchevska@uklo.edu.mk; elena.gelova@ugd.edu.mk; mirjana.kocaleva@ugd.edu.mk

Abstract:
In this paper we present the design of two GeoGebra applets for solving problems involving eigenspaces of a given matrix. One applet features the steps of the standard procedure for finding the eigenvalues and the corresponding linearly independent eigenvectors which will span the eigenspaces. As an additional result, this applet gives an answer to the question whether a given matrix can be diagonalized (by displaying its modal and spectral matrix), or not. The other applet is based on the GeoGebra's commands for the eigenvalues and the Jordan canonical form of a given matrix. It automates the extraction from the similarity matrix in the Jordan decomposition the basis for each of the eigenspace.

Keywords:
Matrices, diagonalizable matrices, defective matrices, eigenvalues, eigenvectors, eigenspaces, Jordan canonical form, Jordan decomposition, GeoGebra

1. Introduction

The standard procedure for solving problems related with the eigenspaces of a given $n \times n$ matrix A involves two steps. The first one is to determine the set of all eigenvalues i.e., the spectrum of A, which comes down to finding all the roots of its characteristic polynomial $P(\lambda) = \det(\lambda I_n - A)$. If λ is an eigenvalue for A, the next step is to determine the eigenvectors corresponding to λ, by solving the matrix equation $A\mathbf{x} = \lambda\mathbf{x}$, with $\mathbf{x} \in \mathbb{C}^n$ as an unknown vector. The eigenspace $E(\lambda)$ corresponding to λ is the set of all solutions of this equation and it is expressed in a form of subspace of $\mathbb{C}^n$ spanned by a finite set of linearly independent eigenvectors for λ. Depending on the size of the matrix and the values of its entries, the calculations that should be performed for each step can be quite extensive and can be properly handled if a suitable software is employed.

GeoGebra is an educational software that includes a computer algebra system which is sophisticated enough to be successfully used for solving various problems when dealing with matrices with moderate sizes. In most of the cases, the CAS specific commands for the eigenvalues and the eigenvectors will display the output in just a few seconds. But, in some cases, the command for the eigenvectors will display a question mark as an output although, for the same matrix, the command for the eigenvalues and the Jordan decomposition will display appropriate results. This situation occurs whenever the matrix is defective i.e., when the matrix is not diagonalizable. A priori, we cannot determine whether a given matrix is defective or not. This can be verified only after the comparison of the algebraic and geometric multiplicities of its eigenvalues. In a case of a defective matrix, we can proceed in two ways. One way is to use other GeoGebra's commands for performing each of the steps of the standard procedure. An applet that is based on this procedure is described in the next section. The other way is to use the result obtained via the Jordan decomposition of the matrix and extract from it all the required information about its eigenvectors and eigenspaces. An applet that automates the extraction of the bases for the eigenspaces is presented in the third section.

The theoretical backgrounds of the spectral theory of matrices, along with variety of solved problems involving the calculation of spectrum and the corresponding eigenvectors and eigenspaces, are usually covered in most of the undergraduate textbooks in linear algebra or operator theory, and their application. For textbooks and articles that also cover the concepts of Jordan canonical form, generalized eigenvectors, and Jordan chains as well, we refer to [1] – [10].

2. GeoGebra applet for the standard procedure

The first applet is based on the output of the **Factors( )** command and works *only if* this command displays a proper factorization of the characteristic polynomial. This means that if the factorization of this polynomial is of form

$$P(x) = (x-\lambda_1)^{k_1} \dots (x-\lambda_m)^{k_m}(x-\mu_1)^{s_1}(x-\bar{\mu}_1)^{s_1} \dots (x-\mu_r)^{s_r}(x-\bar{\mu}_r)^{s_r}, \quad (1)$$

where $\lambda_i \in \mathbb{R}$, for every $i \in \{1, \dots m\}$, $\mu_j \in \mathbb{C}\backslash\{\mathbb{R}\}$ and $\bar{\mu}_j$ denotes the complex conjugate of μ_j, for every $j \in \{1, \dots r\}$, then the **Factors( )** command should display the corresponding factorization in form

$$\begin{pmatrix} x-\lambda_1 & k_1 \\ \vdots & \vdots \\ x-\lambda_m & k_m \\ x^2+p_1x+q_1 & s_1 \\ \vdots & \vdots \\ x^2+p_rx+q_r & s_r \end{pmatrix}, \quad (2)$$

where $x^2+p_1x+q_1 = (x-\mu_j)(x-\bar{\mu}_j)$, $j \in \{1, \dots r\}$. Otherwise the applet will not work properly.

Applets based on the other GeoGebra's commands for location of the roots of the characteristic polynomial and that have a different approach to the solution of the corresponding matrix equation $A\mathbf{x} = \lambda\mathbf{x}$, are available through the links in [11] – [13].

The applet is built via the GeoGebra's CAS view. The list of inputs in the CAS cells is given in Table 1, Table 2, and Table 3.

In each table, at the end of some of the inputs a semicolon is placed. This will suppress the output in the corresponding CAS cell and declutter the applet. If necessary, any such semicolon can be omitted.

The matrix must be defined at the beginning. Once the applet is built, it can be used for a different matrix by simply changing the input in the first CAS cell.

Table 1:
List of the inputs in the CAS cells $1 through $11

CAS cell	Input
$1	A:= (*definition of*)
$2	n:=Length(A)
$3	I:=Identity(n)
$4	λ*I-A
$5	P(λ):=Determinant(λ*I-A)
$6	CFactor(P(x))
$7	U:=Factors(P(x))
$8	U_1:Sequence(If(Degree(Element(U,k,1))=1, Element(U, k),Sequence({x - Element(CSolutions(Element(U,k,1)), m), Element(U,k,2)}, m, 1, 2)), k, 1, Length(U));
$9	L:=Flatten(U_1);
$10	AM:Sequence(Flatten({RightSide(CSolve(Element(L, 2*p-1))), Element(L,2*p)}), p, 1, (1/2) Length(L))
$11	GM:=Sequence({Element(AM, q, 1), Length(A)-MatrixRank(A-Element(AM, q, 1)*I)}, q, 1, (1/2) Length(L))

The outputs in CAS cell $10 and CAS cell $11 are in a form of matrices, each with two columns and the same number of rows. The first column is the same for both matrices and it consists of the roots of the characteristic polynomial. The second column in the output of CAS cell $10 displays the *algebraic multiplicity* of the corresponding roots. If all the roots of the characteristic polynomial are real, this matrix will be equal to the one in CAS cell $7. The second column in the output of CAS cell $11 displays the *geometric multiplicity* of the corresponding roots. If this matrix equals to the one in the output of the CAS cell $10, then the matrix A is diagonalizable. Otherwise, the matrix is defective.

Table 2:
List of the inputs in the CAS cells $12 through $24

CAS cell	Input
$12	ListVar:{x1,x2,x3,x4,x5,x6,x7,x8,x9,x10};
$13	X:=Transpose(Take(ListVar, 1, n));
$14	B:=Sequence(Sequence({Element(I, s, t)}, s, 1, Length(I)), t,1,Length(I));
$15	B_1:=Sequence(Flatten(Solve(I* X-Element(B, q), X)),q,1,n);
$16	SYS:=Sequence({"for","λ=", Element(AM, q, 1),":", (A-Element(AM, q, 1)*I),"*", X,"=",0*X}, q, 1, (1/2) Length(L))
$17	SYS1:=Sequence({"for","λ=", Element(AM, q, 1),":", (A-Element(AM, q, 1)*I)* X,"=",0*X}, q, 1, (1/2) Length(L));
$18	SOL1:=Sequence({"for","λ=", Element(AM, q, 1),":",Flatten(Solve((A-Element(AM, q, 1)*I)* X, X))}, q, 1, (1/2) Length(L));
$19	SOL:=Sequence({"for","λ=", Element(AM, q, 1),":",Transpose(RightSide(Element(SOL1, q, 5)))}, q, 1, (1/2) Length(L))
$20	Eigenspaces:=Sequence({"E(", Element(AM, q, 1),")=","span", Sequence(Substitute(Element(SOL, q, 5),Element(B_1, p)), p, 1,n)\{0*X}}, q, 1, (1/2) Length(L))
$21	M_1:=Join(Sequence(Element(Eigenspaces, q, 5), q, 1, (1/2) Length(L)));
$22	M_2:=Sequence(Transpose(Element(M_1, r)), r, 1, Length(M_1));
$23	M:=Transpose(Join(M_2))
$24	If(Element(Dimension(M), 1)==Element(Dimension(M), 2), {{"SpectralMatrix=",Expand(M^(-1)*A*M)}}, nondiagonalizable)

For matrices with dimensions larger than 10×10 the list of unknows in CAS cell $12 should be extended with x11, x12, x13… , as needed.

The eigenspaces will be displayed in the output of the CAS cell $20 in a form of a matrix where, for each different eigenvalue, the corresponding eigenspace is given by its basis.

If the matrix is diagonalizable then its modal matrix will be displayed in CAS cell $23, while its spectral matrix will be displayed in CAS cell $24. If this is the case, for verification, the inputs listed in the next table can be added.

Table 3:
List of the inputs in the CAS cells $25 through $27 (optional)

CAS cell	Input
$25	M*(Expand(M^(-1)*A*M))*M^(-1)
$26	Expand(Eigenvalues(A))
$27	Expand(JordanDiagonalization(A))

The output in the CAS cell $25 should be equal to the matrix defined in CAS cell $1. The elements in the output list in the CAS cell $26 are the eigenvalues of the matrix. Each value is listed as many times as its algebraic multiplicity. The second matrix in the output of the CAS cell $27, *up to a permutation of its diagonal elements*, should equals to one in the output in the CAS cell $25.

Usually, the commands in CAS cell $26 and CAS cell $27 work quite well if the expressions are replaced with Eigenvalues(A) and JordanDiagonalization(A), respectively, but for some matrices we've observed unusually coined expressions in the outputs, which were overcome by adding the external command **Expand( )**.

While the Jordan canonical form of a given square matrix is unique, up to a permutation of its Jordan blocks, the similarity matrix or, in a case of a diagonalizable matrix its modal matrix, is not. This can happen even for the same permutation of the Jordan blocks. This is due to the fact that for each eigenvalue, regardless of whether the matrix diagonalizable or defective, neither the corresponding eigenvectors, nor the basis for corresponding eigenspaces is unique. Hence, the modal matrix in the

output of CAS cell \$27, or any other application with similar capacities, may differ from the modal matrix obtained in the output of the CAS cell \$27. This does not affect the results for the eigenspaces themselves. They are just described differently i.e., with a different set of vectors as their basis.

3. Alternative GeoGebra applet for finding the eigenspaces

In some cases, the GeoGebra's command **Factors( )** does not factorize the characteristic polynomial in form (2). For example, instead of trinomials, the matrix in the CAS cell \$7 in the previous applet may contain polynomials of larger degrees. In these cases, one alternative is to use other GeoGebra's commands to locate the roots of the characteristic polynomial. The other one is to use GeoGebra's command **JordanDiagonalization( )** and build the applet with the inputs in the CAS cells as listed in Table 4. Again, the matrix must be defined at the beginning i.e., in the CAS cell \$1.

Table 4:
List of the inputs in the CAS cells for the alternative applet

CAS cell	Input
\$1	A:= (*definition of*)
\$2	I:=Identity(Length(A))
\$3	o:=Transpose(Sequence(0, k, 1, Length(A)))
\$4	E:=Eigenvalues(A)
\$5	JD:=JordanDiagonalization(A)
\$6	S:=Element(JD, 1)
\$7	SS:=Sequence(Sequence({Element(S, p, q)}, p, 1, Length(S)), q,1,Length(S))
\$8	EE:=Unique(E)
\$9	Eigenspaces:=Sequence({"E(", Element(EE, q), ")=", "span", RemoveUndefined(Sequence(If(Expand((A - Element(EE, q) I) Element(SS, k)) == o, Element(SS, k)), k, 1, Length(A)))}, q, 1, Length(EE))

4. Examples

Below are the results obtained with each applet for few matrices. Both applets work for real and complex matrices and matrices with undefined entries (parameters).

Example 4.1. The eigenspaces for the matrix

$$A = \begin{bmatrix} a & 1 & 0 & 0 \\ 0 & b & 1 & 0 \\ 0 & 0 & a & 1 \\ 0 & 0 & 0 & b \end{bmatrix},$$

obtained with the applets are given in Figure 1 and 2.

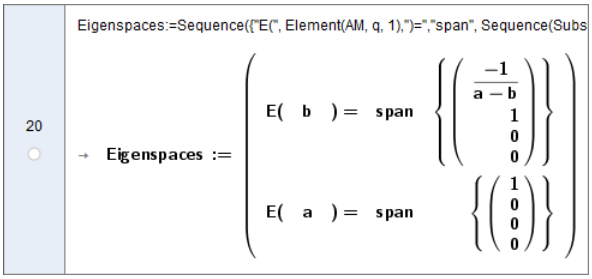


Figure 1: The result for the matrix from the Example 4.1 obtained with the first applet

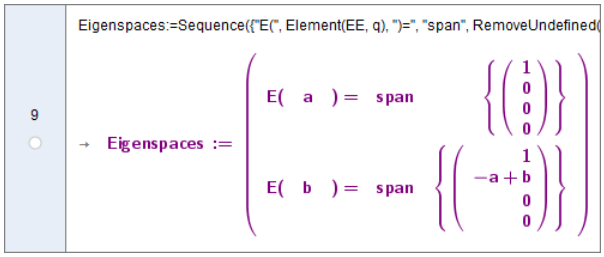


Figure 2: The result for the matrix from the Example 4.1 obtained with the second applet

Example 4.2. The eigenspaces for the matrix

$$A = \begin{bmatrix} 5 & 3 & 5 & 1 & 7 \\ 2 & 3 & 4 & 2 & 4 \\ 0 & -5 & 4 & 7 & 3 \\ 3 & 1 & 6 & 6 & 7 \\ -3 & 2 & -7 & -8 & -7 \end{bmatrix},$$

obtained with the applets are given in Figure 3 and 4.

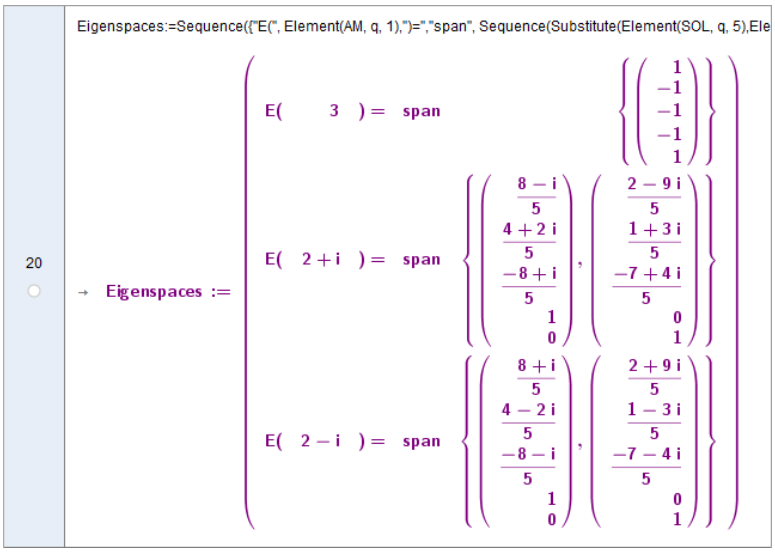


Figure 3: The result for the matrix from the Example 4.2 obtained with the first applet

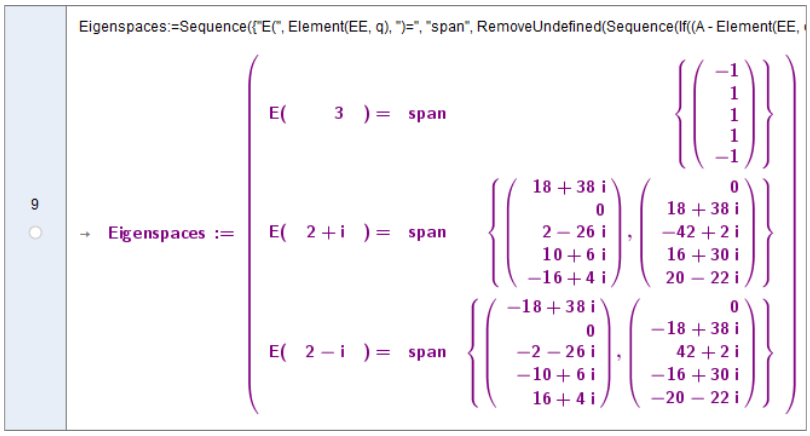


Figure 4: The result for the matrix from the Example 4.2 obtained with the second applet

Example 4.3. The eigenspaces for the matrix

$$A = \begin{bmatrix} 2 & 1 & 3 & 0 & -3 & -2 & -1 & 0 & -2 \\ 0 & 4 & 0 & -1 & 0 & -1 & -2 & 1 & 0 \\ -2 & -13 & 7 & -2 & 0 & -3 & 8 & 1 & -1 \\ -5 & -2 & 5 & 8 & -6 & 0 & 3 & 0 & -3 \\ -2 & -13 & 2 & -2 & 5 & -1 & 8 & 1 & 1 \\ 5 & -1 & -5 & -1 & 7 & 3 & -1 & 0 & -1 \\ 1 & 2 & -1 & -1 & 1 & -1 & 0 & 1 & 0 \\ -2 & -4 & -2 & 5 & 3 & 1 & 2 & 2 & -4 \\ -5 & 1 & 5 & 1 & -7 & 0 & 1 & 0 & 4 \end{bmatrix},$$

obtained with the applets are given in Figure 5, where (a.1) – (a.6) are the results from the first applet, (b.1) – (b.6) are the results from the second applet. The order in which the eigenspaces are displayed in the applets differs from one applet to other.

(a.1)

$$E(\ 5\ ) = \text{span}\left\{\begin{pmatrix}1\\0\\1\\0\\0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\\-1\\0\\0\\1\end{pmatrix}\right\}$$

(b.1)

$$E(\ 5\ ) = \text{span}\left\{\begin{pmatrix}1\\0\\1\\0\\0\\0\\0\\0\\0\end{pmatrix}, \begin{pmatrix}0\\0\\0\\1\\0\\-1\\0\\0\\1\end{pmatrix}\right\}$$

(a.2)

$$E(\ 3\ ) = \text{span}\left\{\begin{pmatrix}1\\0\\1\\0\\0\\1\\0\\1\\0\end{pmatrix}\right\}$$

(b.2)

$$E(\ 3\ ) = \text{span}\left\{\begin{pmatrix}1\\0\\1\\0\\0\\1\\0\\1\\0\end{pmatrix}\right\}$$

(a.3)

$$E(\ 7+i\ ) = \text{span}\left\{\begin{pmatrix}0\\0\\-1\\1+i\\-1\\-1\\0\\i\\1\end{pmatrix}\right\}$$

(b.3)

$$E(\ 7+i\ ) = \text{span}\left\{\begin{pmatrix}0\\0\\1\\-1-i\\1\\1\\0\\-i\\-1\end{pmatrix}\right\}$$

(a.4)

$$E(\ 7-i\ ) = \text{span}\left\{\begin{pmatrix}0\\0\\-1\\1-i\\-1\\-1\\0\\-i\\1\end{pmatrix}\right\}$$

(b.4)

$$E(\ 7-i\ ) = \text{span}\left\{\begin{pmatrix}0\\0\\-i\\1+i\\-i\\-i\\0\\1\\i\end{pmatrix}\right\}$$

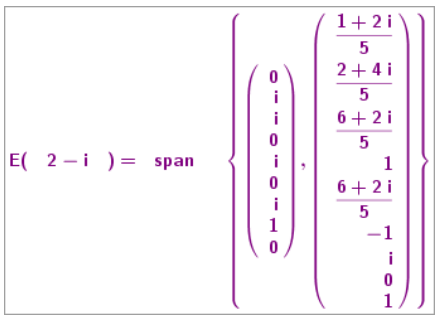
(a.5)

$$E(\ 2+i\ ) = \text{span}\left\{\begin{pmatrix}0\\-i\\-i\\0\\-i\\0\\-i\\1\\0\end{pmatrix}, \begin{pmatrix}\frac{1-2i}{5}\\\frac{2-4i}{5}\\\frac{6-2i}{5}\\1\\\frac{6-2i}{5}\\-1\\-i\\0\\1\end{pmatrix}\right\}$$

(b.5)

$$E(\ 2+i\ ) = \text{span}\left\{\begin{pmatrix}-1+2i\\0\\-4-2i\\-5\\-4-2i\\5\\2+i\\4+2i\\-5\end{pmatrix}, \begin{pmatrix}0\\-1+2i\\-1+2i\\0\\-1+2i\\0\\-1+2i\\-2-i\\0\end{pmatrix}\right\}$$

(a.6)



(b.6)

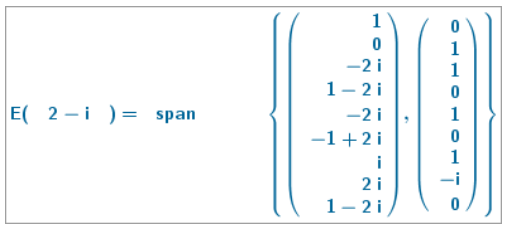


Figure 5: Comparison of the results obtained from the applets for the matrix from the Example 4.1

Example 4.4. The matrix

$$A = \begin{bmatrix} 31 & -7 & 49 & 24 & -5 & 27 & -50 & -33 \\ 4 & -2 & 6 & 3 & -1 & 3 & -6 & -4 \\ -57 & 12 & -87 & -42 & 9 & -47 & 87 & 58 \\ -13 & 0 & -18 & -10 & 3 & -10 & 18 & 13 \\ 16 & -6 & 24 & 11 & -3 & 12 & -24 & -16 \\ 8 & 2 & 11 & 6 & -2 & 6 & -11 & -8 \\ -60 & 15 & -90 & -43 & 9 & -48 & 90 & 61 \\ 31 & -9 & 47 & 22 & -4 & 25 & -48 & -33 \end{bmatrix},$$

has only one eigenvalue and one eigenspace with a dimension equal to 3. Its basis obtained with the applets are shown in Figure 6 and 7.

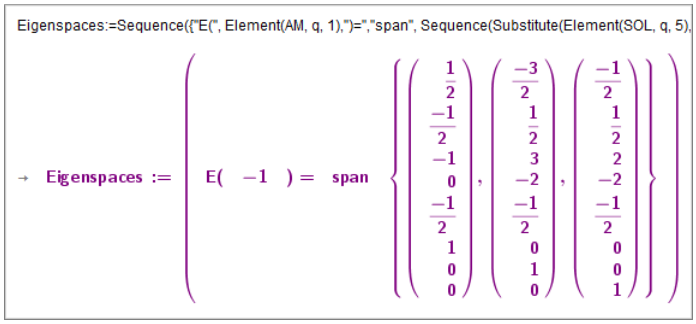


Figure 6: The result for the matrix from the Example 4.3 obtained with the first applet

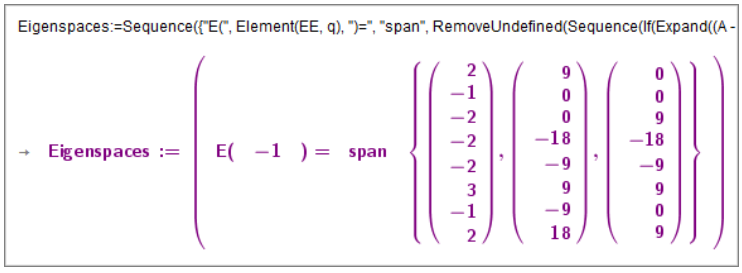


Figure 7: The result for the matrix from the Example 4.3 obtained with the second applet

5. Limitations of the applets

Both applets were tested on various type of matrices, including the ones with undefined entries (like the one from Example 4.1), randomly generated entries, and the ones that were designed with the intention that the output in specific CAS cells can be easily verified with the results obtained if the calculations were performed manually. In some cases, the results were quite impressive. In others, due to the complexity of the algorithms and the extensiveness of the eternal calculation, one or both applets, failed to display the answers. This happened even in the cases where the inputs in some of the CAS

cells were simplified so that can be performed step by step across additional CAS cells (number of which depends on the size of the matrix) and the proper results would be obtained quite quickly.

In the cases where the intermediate results contain some sort of rounding, it was more likely that applets will not produce a proper output in the next cells.

Sometimes problems were overcome if the objects were recomputed (via View → Recompute all objects) or reentering the input in one or more cell by simply positioning the cursor at the end of the expression and pressing Enter. In many cases, saving the changes in the applet after modification of the matrix in the CAS cell $1, closing the application and reopening it, resolved the incorrect display in the outputs of the cells. Increasing the CAS Timeout in GeoGebra Classic 5 from the default 5 seconds to the maximum of 60 seconds, was also helpful in many cases.

6. Conclusion

Depending on the matrix, each applet has a capacity to quickly display the eigenvalues, eigenvectors, eigenspaces and additional intermediate results about the matrix. Each one can be quite useful tool in the undergraduate courses in matrix theory. It can used by teachers for demonstration during the lectures and by students when solving problems that involve finding the eigenvalues and corresponding eigenvectors and eigenspaces of a given square matrix. The applets incorporate a proper interpretation of intermediate results and the outputs of the GeoGebra's commands. This can help students to fully grasp the concept of the Jordan decomposition of a matrix and understand the information that this decomposition contains about a given matrix.

References:

[1] K. M. Abadir, J. R. Magnus, Matrix Algebra, Cambridge University Press, 2005.
[2] R. Bronson, G. B. Costa, Matrix Methods: Applied Linear Algebra and Sabermetrics (4th ed.), Academic Press, 2020.
[3] R. Bronson, Matrix Methods: An Introduction, Academic Press, 1969.
[4] J. B. Fraleigh, R. A. Beauregard, Linear Algebra (3rd ed.), Addison-Wesley, 1995.
[5] S. R. Garcia, R. A. Horn, A Second Course in Linear Algebra, Cambridge University Press, 2017.
[6] R. A. Horn, C. R. Johnson, Matrix Analysis, Cambridge University Press (2nd ed.), 2013.
[7] S. Kurepa, Konačno dimenzionalni vektorski prostori i primjene, Tehnička kniga - Zagreb, 1967.
[8] R. Martínez-Guerra, O. Martínez-Fuentes, J. J. Montesinos-García, Algebraic and Differential Methods for Nonlinear Control Theory: Elements of Commutative Algebra and Algebraic Geometry, Springer Int. Pub. 2019.
[9] M. Taboga, "Generalized eigenvector", Lectures on matrix algebra (2021). URL: https://www.statlect.com/matrix-algebra/generalized-eigenvector.
[10] M. Taboga, "Jordan chain", Lectures on matrix algebra (2021). URL: https://www.statlect.com/matrix-algebra/Jordan-chain.
[11] GeoGebra Community Resources, Diagonalmatrix und Eigenvektoren Rn (2018). URL: https://www.geogebra.org/material/show/id/upUZg79r.
[12] GeoGebra Community Resources, Finding Eigenvectors/Eigenvalues (2015). URL:https://www.geogebra.org/classic/fjxu5cYU.
[13] GeoGebra Community Resources, Jordan-Normalform Hauptvektoren 1.+2.Stufe R5x5 (2019). URL: https://www.geogebra.org/m/cbrraju7.

Exploring the Impact of Pair Programming on Student Achievement: A Comparative Analysis

Ilker Ali[1], Aybeyan Selim[1], Blagoj Ristevski[2] and Sonja Mančevska[2]

[1]*Faculty of Engineering and Architecture, International Vision University, Gostivar, North Macedonia*

[2]*Faculty of Information and Communication Technologies – Bitola, University "St. Kliment Ohridski" – Bitola, Republic of Macedonia*

ilker@vision.edu.mk, ilker.ali@uklo.edu.mk; aybeyan@vision.edu.mk; blagoj.ristevski@uklo.edu.mk; sonja.manchevska@uklo.edu.mk

Abstract:
This study investigates the impact of the pair programming technique on student success in a primary school setting. The research was conducted at Mustafa Kemal Ataturk Primary School in Gostivar, Macedonia, using data from the e-Dnevnik[1] database. This study involved two groups: Group 1, where pair programming was applied, and Group 2, acting as a control without this intervention. Both groups underwent pre- and post-intervention tests to evaluate their programming proficiency. The results revealed that the initial programming proficiency levels were comparable across the groups. The acceptance of the null hypothesis in the pre-intervention phase indicated the similarity of the groups before the pair programming technique was introduced. In the post-intervention analysis, while the null hypothesis showed comparable results, a deeper exploration revealed a significant difference in success rates between the two groups post-implementation. These findings suggest that the pair programming technique positively influences student success. The study contributes insights into innovative teaching methods and their potential impact on student outcomes.

Keywords:
pair programming, educational technology, statistical data analysis, innovative teaching methods

1. Introduction

In the modern landscape of education, the pursuit of effective teaching methodologies is a paramount endeavor [1]. As classrooms evolve to accommodate digital advancements, innovative techniques gain prominence for their potential to enhance student engagement and achievement. One such technique that has gained attention is pair programming, a collaborative approach wherein two students jointly work on coding tasks [2]. Through shared problem-solving and real-time collaboration, pair programming is believed to deepen understanding and foster critical thinking [3] [4].

This study explores the impact of pair programming on student achievement within the primary school context. By assessing its influence on programming proficiency, problem-solving skills, and learning outcomes, we aim to provide insights into the efficacy of this pedagogical strategy [5]. The comparative analysis involves two distinct groups: one engaged in pair programming and the other following traditional individual programming methods.

[1] Web-based electronic register software for primary and secondary school education organizations in North Macedonia

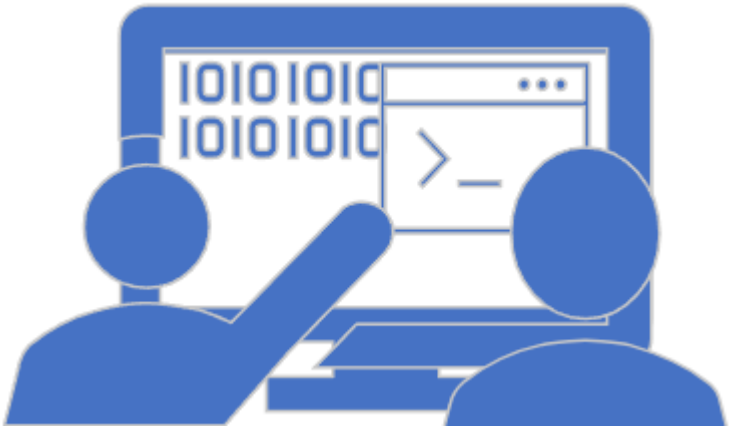


Figure 1 - Pair programming - Students act as drivers and navigators.

The research extends from the premise that innovative teaching methods have the potential to revolutionize classroom dynamics, thereby shaping the success trajectories of students [6]. Pair programming aligns with this philosophy by capitalizing on collaborative learning and the interplay of diverse perspectives to enhance programming proficiency [7].

This paper sets out to explore the effects of pair programming on student achievement by analyzing the results of our comparative study. We delve into the implications of our findings for education, shedding light on the potential benefits of this approach and contributing to the ongoing dialogue on effective teaching methodologies.

2. Application of Pair Programming in Primary Education

Pair programming, a collaborative technique commonly utilized in software development, has been increasingly recognized for its potential to enrich primary education environments [8]. This innovative approach involves two students working together on a single computer, one assuming the role of the "driver," responsible for writing code, and the other as the "observer," offering insights, suggestions, and reviewing the code in real-time, as shown on Fig. 1. While pair programming has been extensively applied in professional contexts, its adoption in primary education presents a promising avenue to enhance learning outcomes, foster teamwork, and cultivate foundational coding skills.

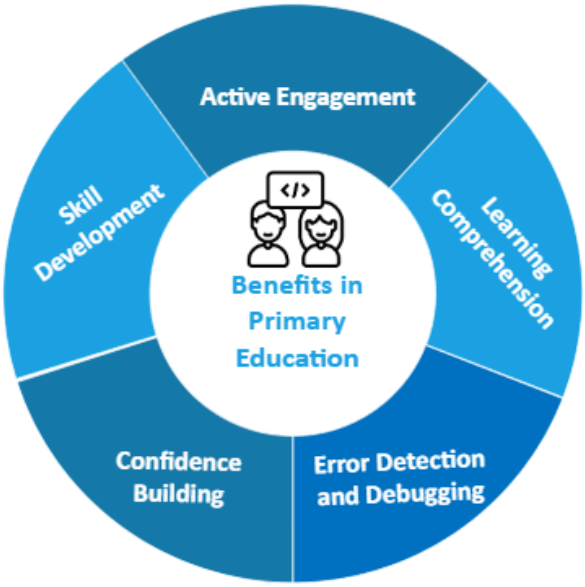


Figure 2 - Benefits in Primary Education.

The main benefits in primary education are shown on Fig. 2:

1. **Active Engagement**: Pair programming captures students' attention through interactive and hands-on activities. The collaborative nature of the approach encourages participation, as students work together to solve problems, share ideas, and develop solutions [9].
2. **Skill Development**: Pair programming promotes the development not only of the coding skills but also crucial soft skills such as communication, teamwork, and problem-solving. Students learn to articulate their thought processes, discuss ideas, and provide constructive feedback to their partners.
3. **Learning Comprehension**: Through real-time discussions, students gain a deeper understanding of programming concepts. The process of explaining code to a partner fosters a more thorough grasp of the material, reinforcing learning outcomes [10].
4. **Confidence Building**: Collaborative problem-solving can boost students' self-assurance. As they work alongside a partner, students can rely on one another's strengths, which in turn cultivates a sense of accomplishment and resilience [11].
5. **Error Detection and Debugging**: The collaborative nature of pair programming enables immediate error detection. Partners can identify mistakes or misunderstandings early, leading to quicker resolution and a more efficient learning process.

During the implementation of the pair programming, the following should be considered:

1. **Pair Formation**: Pairs can be formed based on varying skill levels, promoting peer learning. Mixing students with different proficiencies encourages a supportive environment where both participants contribute to the partnership [12].
2. **Rotation**: Periodically rotating roles between the driver and observer ensures that students experience both aspects of coding and teamwork. This practice enhances empathy and a well-rounded understanding of the process [13].
3. **Clear Guidelines**: Establish clear guidelines on communication, respect, and collaborative problem-solving. Encourage students to explain their thought processes and exchange ideas freely.
4. **Adaptable Challenges**: Tailor coding challenges to suit the pair programming approach. Assign tasks that require joint problem-solving and allow for creativity in solution development [14].
5. **Feedback and Reflection**: Incorporate feedback sessions where pairs discuss their experiences, challenges they've faced, and lessons they've learned. This reflective practice enhances the learning process and encourages continuous improvement.

Incorporating pair programming into primary education aligns with modern pedagogical trends that emphasize active learning, collaboration, and technology integration. By fostering a collaborative spirit and nurturing coding proficiency from an early age, educators set the stage for students to embrace technology with confidence and enthusiasm in an increasingly digital world [15].

3. Research methodology

The data was sourced from the e-Dnevnik database of Mustafa Kemal Ataturk Primary School in Gostivar. To evaluate the influence of the pair programming technique on success, both the independent programming and pair programming groups underwent the same test. The test results data was analyzed using appropriate methods, and the outcomes for both groups were examined.

3.1. Material

The study incorporated two distinct tests, which were administered to all participant groups. These tests covered diverse task categories. In the first category, participants were required to convert decimal numbers to binary and answer theoretical questions about variables in the C++ programming language. In the second category, students were tasked with constructing and coding solutions on a computer. Within this problem set, three question types were presented:

1. Multiple-choice questions.
2. Questions requiring written answers.
3. Questions involving coding in the C++ programming language.

3.2. Procedure for data collection

The process of data collection involved two distinct phases: the pre-pair programming technique assessment, and the assessment of challenges encountered during pair programming. The execution of each stage extended over a twelve-week period in this research. The pre-pair programming technique test aimed to proactively identify potential issues that might emerge during the subsequent pair programming tests [16]. The assessment encompassed a total of 85 participants.

4. Results

Commencing the data analysis endeavor, the primary stride encompassed the calculation of the grade averages and corresponding standard deviations for each class within the sixth grade, preluding the application of the pair programming technique. The encapsulated outcomes are showcased in Table 1, where a harmonious alignment of mean values and standard deviations is evident in the context of the pre-pair programming stage. This visual representation substantiates the notion that the levels of programming acumen exhibited minimal discrepancies among the various classes.

Table 1:
Group statistics of classes before the pair programming technique

	N	Average Grade	Standard Deviation of Grades
VI-1	14	3.60	2.61
VI-2	22	2.83	2.40
VI-3	15	3.33	2.54
VI-4	23	2.99	2.46
VI-5	11	3.13	2.50
Total	85		

To assess the influence of pair programming on student success within a simplified context, two distinct groups were established: Group 1 and Group 2. In Group 1, the pair programming technique was applied, while in Group 2, this technique was not incorporated.

Following this stage, our focus shifted to selecting the subsequent procedure for analyzing the outcomes of both the pair programming (Group 1) and individual programming (Group 2) groups. In determining the suitable approach, we evaluated the distribution of the results, which exhibited normal distribution characteristics. Consequently, we opted for the Independent-Samples t-test as the analytical technique, considering its compatibility with data conforming to a normal distribution.

To ascertain the comparability of the groups, we administered the Levene test for equality of variances in the gathered data [17]. In this context, we formulated the subsequent null and alternative hypotheses[2] for the test's results:

$$H_0: \sigma^2_{Group\ 1} = \sigma^2_{Group\ 2}$$
$$H_a: \sigma^2_{Group\ 1} \neq \sigma^2_{Group\ 2}$$

Table 2:
Levene's test for equality of variances before pair programming technique

	F	Sig.
Equal variances assumed	.150	.699
Equal variances not assumed		

[2] σ^2 represent the variance of the groups.
H_0: The variances of two groups are equal, indicating that there is no variation in the response variable among different treatment groups.
H_a: The variances of the groups are not equal, suggesting that there is variation in the response variable among different treatment groups.

Analyzing the findings presented in Table 2, it becomes evident that the calculated significance level stands at 0.699, which is greater than the conventional threshold of 0.05. Consequently, the null hypothesis (H_0) is accepted. This implies that the differences in outcomes before the implementation of the pair programming technique between Group 1 and Group 2 were comparable. In simpler terms, the initial results of both groups exhibited similarity prior to the utilization of the pair programming technique.

We utilized the t-test for equality of means as a means to compare the average success levels between the two groups, and the resulting hypotheses[3] are as follows.

$$H_0: \bar{X}_{Group\ 1} = \bar{X}_{Group\ 2}$$
$$H_a: \bar{X}_{Group\ 1} \neq \bar{X}_{Group\ 2}$$

Extracting insights from the data provided in Table 3, we see that the computed significance value stands at 0.834, surpassing the conventional threshold of 0.05. Consequently, the null hypothesis (H_0) finds acceptance in this scenario. This pivotal outcome underscores that a notable discrepancy in success rates between the two groups is absent. In simpler terms, the pre-existing state of the groups was closely akin before the initiation of the pair programming technique.

The affirmative acceptance of the null hypothesis implies that any observed disparities in success rates post-implementation can be attributed to the introduced technique's influence rather than inherent discrepancies. The alignment of the groups' success metrics is a critical foundation for further examining and assessing the technique's efficacy. As the experiment unfolds, these insights lay the groundwork for a comprehensive evaluation of the technique's impact on the success rates of the involved groups.

Table 3:
t-test for equality of means before the pair programming technique

T	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
.210	83	.834	.116	.553

This analysis underscores the notion that the two groups were nearly indistinguishable regarding their baseline characteristics before integrating the pair programming approach. It is worth noting that the engagement and alignment of the two groups' performance metrics provide a solid foundation for meaningful comparisons post-intervention. Group 1 underwent a three-month training phase involving the pair programming technique.

After training, we apply the same test with the same hypotheses to check the impact of pair programming on students' success. The testing results are given in Table 4 and Table 5.

Table 4:
Levene's test for equality of variances after pair programming technique

	F	Sig.
Equal variances assumed Equal variances not assumed	3.856	.053

The findings presented in Table 4 indicate that the significance value observed in the posttest was 0.053, which surpasses the threshold of 0.05. As a result, we accept the null hypothesis (H_0). This outcome suggests that the two groups can be comparable in their post-testing results, as the variances in their respective test outcomes exhibited parity. Consequently, we deduce that the efficacy of implementing the pair programming technique applies to assessing the results of these two groups.

[3] $\bar{X}$ represent the mean of the groups.
H_0 : There is no significant difference among the means of the groups. In other words, both groups means are equal.
H_1: There is a significant difference among the means of the groups. At least one group mean is different from the others. There is a significant difference among the means of the groups. At least one group mean is different from the others.

Table 5:
t-test for equality of means after the pair programming technique

T	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
3.959	83	.000	2.015	.509
4.015	79.082	.000	2.015	.502

According to the results in Table 5, the meaning value is 0.00 <0.05, so in this case, Ha is accepted. This means there is a significant difference in the success of Group 1 and Group 2 after the application of pair programming. Consequently, we conclude that sufficient evidence suggests that the pair programming technique is an effective teaching method and increases the student's success. The means of success after pair programming is given in Table 6.

Table 6:
Group statistics after pair programming technique

	N	Mean	Std. Deviation	Std. Error Mean
Pair Programming	36	3.75	2.196	.366
Individual Programming	49	2.63	2.405	.344

5. Discussion

While this study contributes valuable insights into the impact of the pair programming technique on student success, it is important to acknowledge certain limitations that may have influenced the results and interpretations.

- **Sample Size and Generalizability:** The study was conducted with a relatively small sample size of 85 participants. This limited sample size might restrict the generalizability of the findings to a broader student population. A larger and more diverse sample could provide a more comprehensive understanding of the technique's effectiveness across different demographic and educational contexts.
- **Duration of Intervention:** The study implemented the pair programming technique within a three-month period. This duration might not fully capture the long-term effects of the technique on student success. Longer-term studies could reveal how the impact of pair programming evolves over extended periods of time.
- **Contextual Factors:** The study was conducted within the specific context of Mustafa Kemal Ataturk Primary School in Gostivar. Different schools or institutions with varying resources, teaching methods, and student populations might yield different outcomes. Cultural, socio-economic, and institutional factors could influence the effectiveness of the pair programming technique.
- **Test Selection and Measurement:** The study utilized specific tests to measure student success. The choice of tests and the specific skills they measure could limit the scope of the findings. A more comprehensive range of assessments could provide a more holistic understanding of the technique's impact on diverse aspects of student success.
- **External Variables:** The study might not have accounted for all potential external variables that could influence student success. Factors such as individual learning styles, prior programming experience, and the level of student engagement with the technique could impact the observed outcomes.
- **Experimental Design:** The study employed a quasi-experimental design with Group 1 receiving the pair programming intervention and Group 2 acting as the control. While efforts were made to ensure comparability between the groups, inherent differences between the groups could still influence the results.
- **Learning Curve and Adaptation:** Students in Group 1 might have experienced a learning curve when adapting to the pair programming technique. The initial phases of implementation might

not accurately reflect the technique's long-term effects, potentially influencing the observed success rates.

- **Participant Motivation:** The study did not extensively explore participant motivation, interest, or attitude toward pair programming. These psychological factors can play a significant role in the effectiveness of any educational technique.
- **Teacher Influence:** The role of teachers in facilitating pair programming and creating a conducive learning environment could impact the outcomes. Variations in teacher expertise and teaching styles might introduce confounding variables.
- **Time Constraints:** The study's three-month timeline might have imposed time constraints on the effectiveness of the pair programming technique. A more flexible and extended implementation period might yield different results.

6. Conclusion

This study delved into the impact of the pair programming technique on student success within the context of Mustafa Kemal Ataturk Primary School in Gostivar. Through a comparative analysis of Group 1 (pair programming applied) and Group 2 (individual programming), valuable insights were gained into the efficacy of this innovative teaching method.

The study's findings unveiled several significant insights. First, the preliminary analysis revealed that the two groups displayed comparable baseline characteristics regarding programming understanding, as indicated by the harmonious alignment of mean values and standard deviations. This served as a solid foundation for meaningful comparisons throughout the study.

The application of the pair programming technique demonstrated its potential to enhance student success. Accepting the null hypothesis in the pre-intervention phase highlighted the comparability of the two groups, setting the stage for assessing the technique's impact. Moreover, the null hypothesis's acceptance in the post-intervention step implied a lack of significant success rate disparities between the groups post-implementation.

However, a closer examination of the same post-intervention phase results revealed a significant difference in success rates between Group 1 and Group 2, confirming the positive influence of the pair programming technique. This outcome underscored the technique's effectiveness as a teaching approach, suggesting collaborative programming enhances student success.

Despite the valuable contributions, this study is not without limitations. The relatively small sample size, contextual factors, and specific measurement tools employed all pose potential constraints on the generalizability and scope of the findings.

This study contributes to the ongoing exploration of innovative teaching methodologies and their impact on student success. The evidence gathered supports the notion that the pair programming technique has the potential to improve student success rates significantly. As education evolves, further research is warranted to explore the technique's effectiveness across diverse settings, demographics, and learning outcomes. The findings of this study, while shedding light on the positive effects of pair programming, invite continued investigation into optimizing its implementation and understanding the nuances of its impact on student success.

References

[1] C. R. Rodgers and M. B. Raider, "Presence in teaching," *Teachers and Teaching: theory and practice*, vol. 12, no. 3, 2006.

[2] A. Selim and I. Ali, "Dijital Okuryazarlık Becerilerinin Kuzey Makedonya'da Yükseköğretim Üzerindeki Etkisi," in *ICEB - 6 - International Congress of Economics and Business*, Gostivar North Macedonia, 2022.

[3] I. Ali, B. Ristevski and I. Jolevski, "Pair Programming In Primary Education - A Macedonian Case Study," in *12th International conference on Applied Internet and Information Technologies (AIIT2022), October 14th 2022, Zrenjanin, SerbiaAt: Zrenjanin, Serbia*, Zrenjanin, 2022.

[4] C. Lewis, "Is pair programming more effective than other forms of collaboration for young students?," *Computer Science Education,* vol. 21, pp. 105-134, 2011.

[5] F. Jakoš and D. Verber, "Learning Basic Programing Skills With Educational Games," *Journal of Educational Computing Research,* 2016.

[6] H. Christothea, S. Mike, G. Mark and K.-H. Agnes, "Innovative Pedagogies of the Future: An Evidence-Based Selection," *Frontiers in education,* vol. 4, 2019.

[7] N. Pellas and S. Vosinakis, "The effect of simulation games on learning computer programming: A comparative study on high school students' learning performance by assessing computational problem-solving strategies," *Educ Inf Technol,* vol. 23, p. 2423–2452, 2018.

[8] F. Rodríguez, K. Price and K. Boyer, "Exploring the Pair Programming Process: Characteristics of Effective Collaboration," pp. 507-512, 2017.

[9] F. Xu and A. Correia, "Adopting distributed pair programming as an effective team learning activity: a systematic review," *Journal of Computing in Higher Education,* pp. 1867-1233, 15 02 2023.

[10] M. Seyam, D. S. McCrickard, S. Niu, A. Esakia and W. Kim, "Teaching mobile application development through lectures, interactive tutorials, and Pair Programming," *IEEE,* 2016.

[11] M. Schmitz and H. Winskel, "Towards effective partnerships in a collaborative problem-solving task," *The British journal of educational psychology,* vol. 78, pp. 581-96, 2008.

[12] C. McDowell, L. Werner, H. E. Bullock and J. Fernald, "Pair programming improves student retention, confidence, and program quality," *Communications of the ACM,* vol. 49, no. 8, pp. 90-95, 2006.

[13] A. Selimi and A. Useini, "Yenilikçi Eğitim ile Dijital Yetkinlik ve Girişimcilik Becerilerinin Geliştirilmesi - Kuzey Makedonya Örneği," in *ICEB'19 – International Congress of Economics and Business*, 2019.

[14] F. Skender and I. Ali, "Big Data in Health and the Importance of Data Visualization Tools," *Journal of Intelligent Systems with Applications,* 2022.

[15] M. U. Bers, "Coding as another language: a pedagogical approach for teaching computer science in early childhood," *Journal of Computers in Education,* vol. 6, p. 499–528, 2019.

[16] A. Kirksekiz and S. Kol, "Bibliometric Analysis of Early Childhood Education Studies on The Theme of Robotic Coding From A Developmental Perspective," *International Journal of Educational Research Review,* 2023.

[17] M. B. Brown and A. B. Forsythe, "Robust Tests for the Equality of Variances," *Journal of the American Statistical Association,* vol. 69, no. 346, pp. 364-367, 1974.

Graph-Based Task Management Parameterized by Linguistic Path Attributes

Dalibor Šeljmeši[1], Vladimir Brtka[1], Edit Boral[2] and Ivana Berković[1]

[1]*University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*
[2] *ASA College/Department of Computer Technology, New York, USA*

dalibor.seljmesi@tfzr.rs; vladimir.brtka@tfzr.rs; eboral@asa.edu; ivana.berkovic@tfzr.rs

Abstract:
The paper deals with task management using graph representation. The method of setting management conditions using linguistic terms, that is, spoken language words, is presented. This is achieved by using the principles of fuzzy logic. A modified Dijkstra algorithm was used to calculate the optimal path under restrictive managing conditions. The presented approach enables the reduction of large graphs, that is, the rejection of nodes that do not meet the linguistic management conditions. The results of experiments that were undertaken with the aim of gaining a deeper insight into the possibilities of this type of task management are also presented.

Keywords:
task management, linguistic variables, fuzzy logic, graph search

1. Introduction

Task management is the process of planning, organizing and monitoring tasks in order to achieve the desired outcome. It is a key skill for individuals and teams to be successful in their work. Traditional approaches to task management often rely on lists or spreadsheets, which can be difficult to use and maintain, especially for large projects. Task management is an activity that occurs frequently in various application domains and at various levels of complexity. Task control can be very complex and demanding, but the importance of good management is undeniable. The term task refers to a part of a process or procedure that is complex in nature. Complexity implies that the process contains many tasks that are logically connected. So, in this case, the task is an integral part of the process. The procedure for managing the execution of tasks is more difficult to implement if the process to be carried out does not require the execution of all tasks it consists of but only certain ones. This means that there are multiple ways of executing the process. Hence, we deal with dynamic processes that can have several initial states (beginning of the process) and several final states (end of the process), which means that the process can be performed in several ways. A dynamic process is a collection of tasks without a predetermined sequence of execution, which can cause many problems when implementing the process. Some examples of dynamic processes are:

- Project management because the project's requirements and scope can change over time,
- Software development because the requirements of the software can change over time,
- Market research because the market can change over time,
- Customer service because the needs of customers can change over time.

There are many challenges related to this, such as: unattainable deadlines, ineffective communication, poor risk management, unclear goals and intentions, etc. This paper deals with Graph-based task management, which has a number of advantages over traditional task management approaches. It can help to:

- Visualize relationships between tasks,
- Identify dependencies between tasks,
- Improve planning and scheduling,
- Improve the accuracy of task estimates,
- Improve overall project performance.

This approach to task scheduling is based on graph theory, which later extends to the use of linguistic variables. So, in this paper, the approach to the execution of dynamic processes is achieved using a graph representation, where the nodes of the graph represent states and edges represent logical dependencies between states (these are tasks). Each graph edge is characterized by several numerical attributes that describe the characteristics of the task. The assumption is that the initial state of the process to be executed is known, as is the final (terminal) state of the dynamic process; it is necessary to determine the path between the initial and final states. The determination of the path depends on the characteristics of the connections between states. The essence of this work is the way in which it defines the conditions that the path must satisfy in order to execute the dynamic process. It is convenient to use linguistic terms, i.e., words of spoken language, to define the conditions that the path should satisfy. Examples are diverse and can be: find a path with maximum priority OR minimum duration; find a path with maximum cost AND average complexity, etc. Obviously, logical operators such as AND and OR are used, and linguistic expressions can be defined as fuzzy propositions. This approach enables a natural way of defining the conditions for executing a dynamic process. For path calculation, a variation of Dijkstra's algorithm is used.

The paper is organized as follows: Section 2 contains a brief insight to previous work in the field of task management. In Section 3 the main principles of graph-based task management parameterized with linguistic path attributes are explained on a simple example. Section 4 contains some experimental results and discussion, while conclusions and future work are given in Section 5.

2. Previous work

Some insights into task management are given in [1]. This paper recognizes the importance of process management and emphasizes the increased effectiveness of good process management in the case of complex, ambiguous, and time-sensitive processes. In this case, it is important to choose the right representation of the dynamic process.

The effectiveness of task management depends on several factors, such as the type of tasks, the structure of the dynamic process, and its execution. As it was said in [2], a good representation of the project is the key component in solving tasks in dynamic processes.

Each task management method has its own weaknesses and strengths. For example, Kan-ban boards were good for visualizing work and tracking progress. The findings represented in [3] can be used by individuals, organizations, and researchers to choose the right task management technique for their needs.

In the realization of each project, task management has a large and significant application. It is necessary to choose the best method for a certain type of project. Also, in addition to the selected method, it is important to prioritize the tasks so that the important tasks are performed first and then the others that are less important [4].

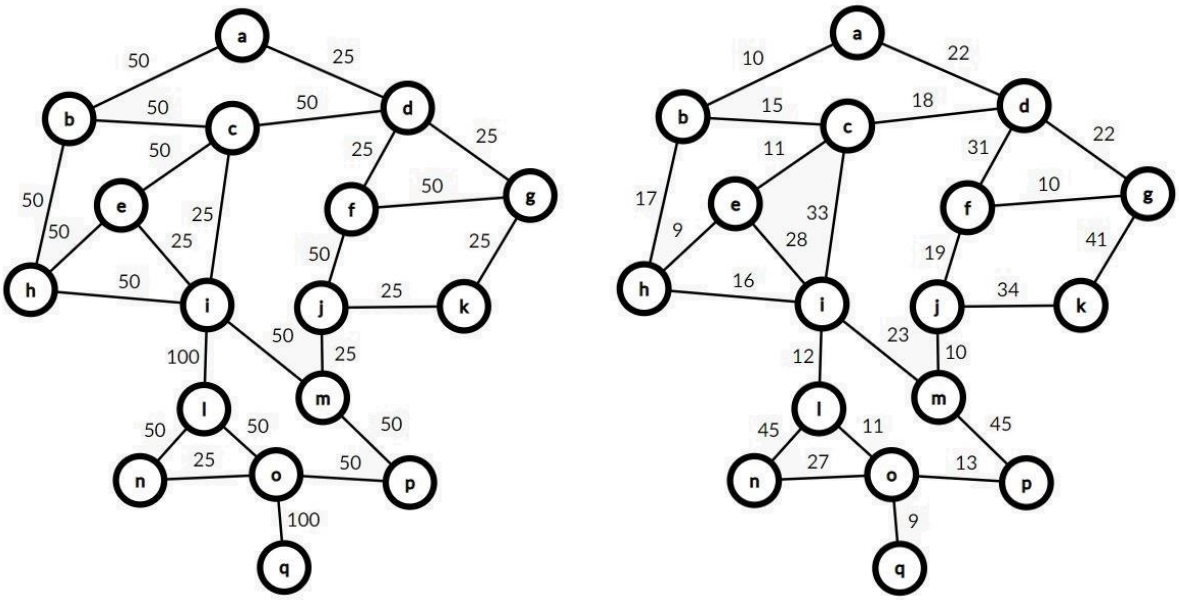
The best task management technique for a particular situation will depend on a number of factors, such as the type of task, the size of the team, and the project's goals. The research presented in [5] also found that the effectiveness of task management tools varied. Some tools were more effective for certain types of tasks than others. For example, Kan-ban boards were more effective for tasks that were complex and time-sensitive, while to-do lists were more effective for tasks that were simple and straightforward.

Based on the review of previous research, it can be concluded that a suitable representation of the process is very important, especially when the processes are of a dynamic nature. For this purpose, graphs are often used in many domains [6] [7], and it is convenient to have the possibility of a simple and natural definition of the conditions required for the execution of the process.

3. Linguistic path attributes

Traditional task management techniques use charts, diagrams, and graphs to organize and track tasks. Here, managing action is executed on a graph representation. As there are multiple values related to each task, edges of the graph are associated with multiple values represented by an array of scalars or a single vector. This might be a problematic situation for path calculation algorithms such

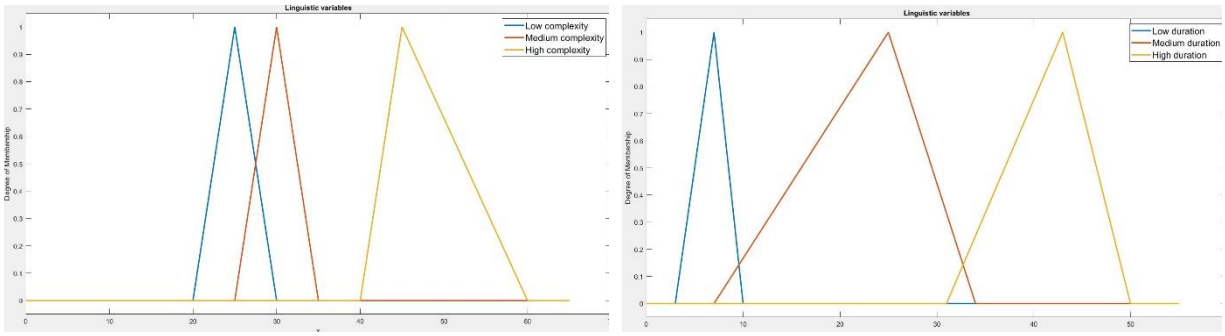
as Dijkstra or Floyd-Warshall to execute, since edge costs are not a scalar values. Here, edge costs are calculated as the aggregated value of multiple scalars, which leads to numerous possibilities for total cost calculations. In order to describe the problem, Figure 1 shows an example of a simple graph that represents some arbitrary dynamic process that needs to be realized. We can write this graph as a set G = {V, E}, where V is the set of nodes and E the set of branches [8]. We have: $V = \{a, b, c, d, e, f, g, h, i, j, k, l, m, n, o, p, q\}$ and $E = \{\{a, b\}, \{a, d\}, \{b, h\}, \{b, c\}, \{c, e\}, \{c, i\}, \{c, d\}, \{d, f\}, \{d, g\}, \{e, h\}, \{e, i\}, \{f, j\}, \{f, g\}, \{h, i\}, \{i, l\}, \{i, m\}, \{j, m\}, \{j, k\}, \{l, n\}, \{l, o\}, \{m, p\}, \{n, o\}, \{o, p\}, \{o, q\}\}$. In this simple example, each graph edge is associated with two values: let them be task complexity and task duration.



a) b)

Figure 1: Simple graph: a) task complexity b) task duration

As there are two values associated with each edge, there are two topologically equivalent graphs but with different edge costs representing task complexity and task duration. Both values are scalars representing percent. Before applying any graph search algorithm, it is necessary to aggregate these two values in order to form a single cost value for each edge. The aggregation function is implemented as a fuzzy logical operator, meaning that scalar values must be converted to [0,1] interval by fuzzification procedure. In order to implement the procedure, task complexity and task duration must be defined as linguist variables [10], so that the values of task complexity are: low complexity, medium complexity and high complexity, while the values of task duration are: short duration, medium duration and long duration. These fuzzy values are defined by membership functions, as shown in Figure 2.



a) b)

Figure 2: Fuzzy variables: task complexity and task duration

Fuzzification procedure is done by a triangular membership function (1), although different membership functions can be used (trapezoidal, Gaussian, etc. [9]).

$$trimf(x) = \begin{cases} 0, & x \leq a \\ \dfrac{x-a}{b-a}, & a \leq x \leq b \\ \dfrac{c-x}{c-b}, & b \leq x \leq c \\ 0, & c \leq x \end{cases} \quad (1)$$

By varying the parameters, a, b, and c in (1), different linguistic terms are defined. For a given start node and end node, it is possible to define different path attributes (conditions) that the path between the start and end node should satisfy. Path attributes are defined by fuzzy propositions aggregated by fuzzy operators [11]. For two fuzzy variables, path attributes are expressed by (2).

$$A_1 = a_1 \oplus A_2 = a_2 \quad (2)$$

In (2), task complexity is represented as A_1, while task duration is represented as A_2. Also, a_1 and a_2 represent linguistic values of A_1 and A_2 respectively. Operator $\oplus$ usually represents AND logical operator (implemented as *min*) or OR logical operator (implemented as *max*). There is also the *Min-Max* compensatory operator (3).

$$\textit{Min-Max}: (1-\lambda) \min_{i=1..n}(\mu_i) + \lambda \max_{i=1..n}(\mu_i) \quad (3)$$

In (3) *min* function corresponds to the linguistic AND operator, while the *max* function corresponds to the linguistic OR operator. By varying $\lambda \in [0,1]$ which is user-defined, one can choose between „pure AND“ or „pure OR“ operators, or some operator „between“ these two. Operator (3) for $\lambda \neq 1, \lambda \neq 0$ is used when there is no clear indication wheter to use AND operator or OR operator ($\lambda = 0.5$ is default value). If F is a fuzzy proposition and $\oplus_\lambda$ is fuzzy compensatory operator parameterized by λ, then the path parameterized by n propositions is given by (4).

$$\oplus_\lambda(F_i): i = 1, \ldots, n \quad (4)$$

In order to investigate application possibilities, the following section contains descriptions of a few experiments under various restrictions.

4. Experiments and Discussion

In order to execute experiments, the modified Dijkstra algorithm was used. Modification involves finding the path with the highest cost. Such a modification of the algorithm is necessary, bearing in mind that edges costs represent the measures with which linguistic path attributes (linguistic conditions) are satisfied.

The first experiment varies the starting node, as well as the final node, while the fuzzy operator is AND (*min*). The path attribute (linguistic condition) is **low complexity AND medium duration**. The results are shown in Table 1.

Table 1:
Results of experiment 1

Start node	Path	Final node	Total cost
g	g-d-a-b-h-i-l-o-n	n	5.900
m	m-j-k-g-d-a	a	5.000
p	p-o-l-i-c-b	b	3.875

l	l-i-c-d-g	g	3.025
e	e-c-d-f-j-m	m	3.100

Figure 3 shows edge costs after applying the fore-mentioned linguistic condition.

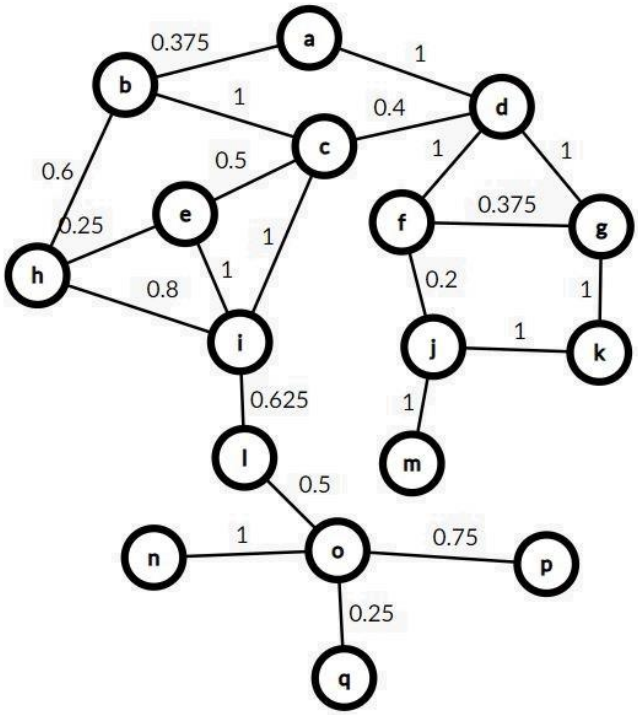


Figure 3: Fuzzified edge costs

The total cost represents the algebraic sum of all measures by which this linguistic condition is satisfied. Therefore, the maximal total cost is the best, so that the best starting position is node *g* and the best end position is node *n*, so that the optimal path is: *g-d-a-b-h-i-l-o-n*. The optimality of the path can also be affected by the number of included nodes, which depends on the specific application.

The pre-conditions of the second experiment are: the start node is *a*, the final node is *p*, while the value of $\lambda = 0.5$. The essence of this experiment is to investigate the influence of varying path attributes on the generated path. Table 2 shows the results of the second experiment.

Table 2:
Results of experiment 2

Linguistic variable for task complexity	Linguistic variable for task duration	Path	Total cost
High complexity	Short duration	a-b-h-i-m-p	1.66667
Low complexity	Short duration	/	0
High complexity	Medium duration	a-b-h-i-m-p	2.47917
Medium complexity	Medium duration	a-b-h-i-l-o-p	1.75000
Low complexity	High duration	a-d-g-k-j-m-p	3.37500

The maximum value of total cost is 3.375, which indicates that attributes low complexity, high duration give the best result. An interesting result was produced under the condition of low complexity, short duration. Namely, in this case, it happens that there are edges with a cost of 0, and therefore there are non-connected nodes, so they are discarded. Figure 4 illustrates this case.

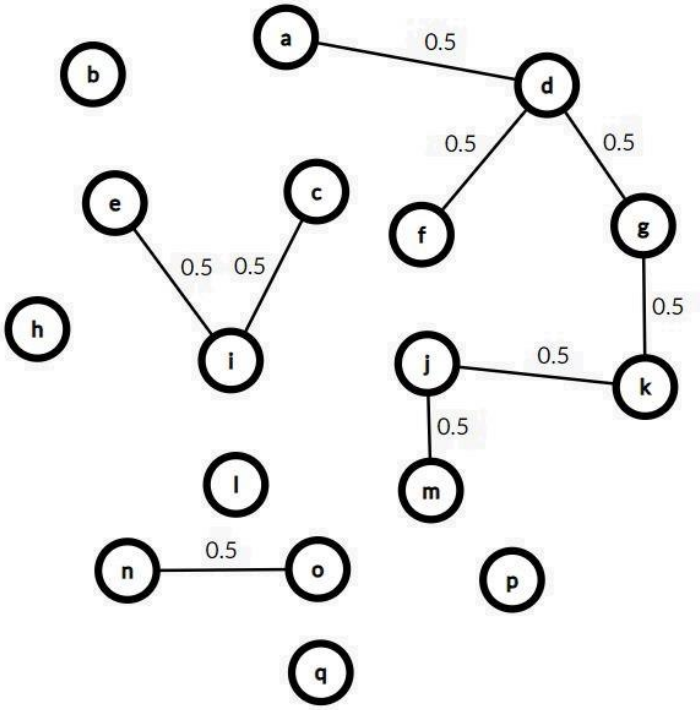
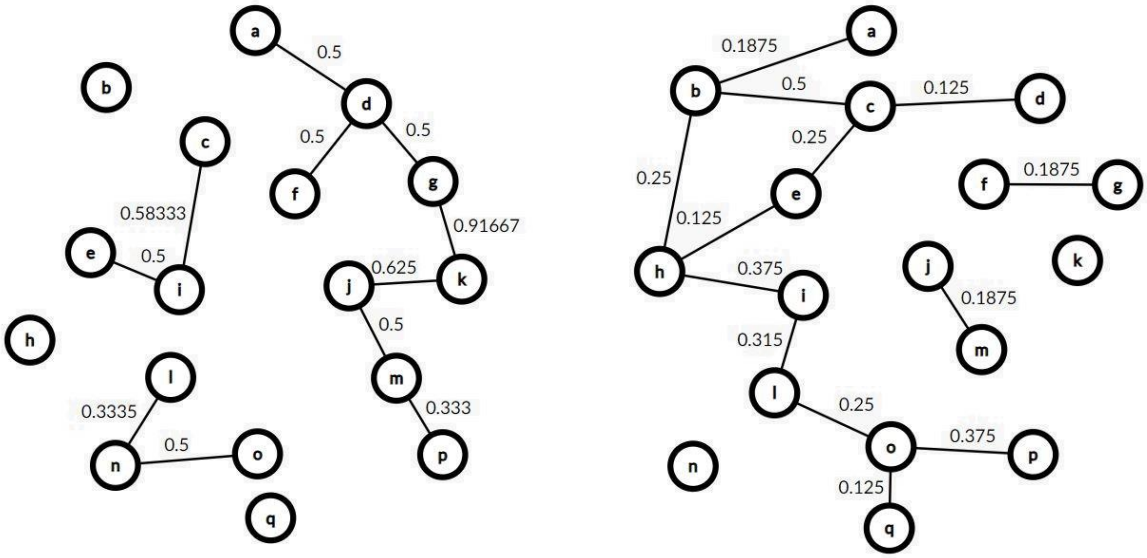


Figure 4: Discarded nodes

This means that there are cases when the linguistic conditions cannot be met, i.e., it is not possible to determine the path that connects the start and final nodes under the given conditions. Figure 5 illustrates an example of how different linguistic conditions affect path generation.



a) b)

Figure 5: a) low complexity, high duration b) medium complexity, medium duration

It is possible to see that the differences between graphs a) and b) from Figure 5 are significant in terms of edge costs and the number of discarded nodes. This can be a convenient visual tool that allows spotting suitable paths regarding given linguistic conditions.

5. Conclusions

Task management is certainly very important when monitoring and implementing various processes. The approach to task management using graph representation certainly enables automation and facilitates the management process itself. The approach presented in this paper shows that it is possible to set management conditions using linguistic terms, i.e. words of the spoken language, which greatly facilitates the process. The following can be concluded:

- It is possible to set management conditions linguistically and enable rapid prototyping of different management alternatives.
- In the case of extensive graphs, this procedure enables discarding nodes, which reduces the volume of the graph, i.e. discarding nodes that absolutely do not meet the pre-conditions.
- It is possible to single out sub-graphs that can be considered separately, using different linguistic management conditions.
- The presented solution can be the basis for a visual representation of the task management process.

This approach enables the reduction of large graphs, which can significantly improve the efficiency of task management. The results of experiments show that the proposed approach is effective in reducing the size of graphs and improving the efficiency of task management.

The presented way of task management can be applied in various domains where such management methods are otherwise applied. In addition to the above-mentioned applications, applications in industrial processes are especially interesting, where the objects being processed can be exposed to different processing procedures.

Future work will include experiments on of larger graphs in terms of the number of nodes and edges. Also, it is interesting cases with varying parameters and different fuzzy logic operators.

Acknowledgment:

Ministry of Science and Technological Development, Republic of Serbia financially supported this research, under the project number TR32044: ”The development of software tools for business process analysis and improvement”.

References:

[1] Chaudhry, A., Khan, S. U., Anjum, M.S., Pervaiz, I. Javed, M. “The Impact of Task Management on Individual Performance”. Information System Frontiers 2019.
[2] Nagel, K., Scwade, J. “The Effectiveness of Task Management Tools in Virtual Teams”. Computer in Human Behavior 2020.
[3] Dingsøyr, T., Gravingholt, B., Moe, N.B. “A Comparison of Task Management Techniques”. Information and Software Technology 2012.
[4] Shen, Y., Shen, Q., Li, J. “The Role of Task Management in Project Success”. Journal of Project Management 2019.
[5] Shen, Y., Shen, Q., Li, J. “A Study of Task Management in Agile Software Development”. Information and Software Technology 2012.
[6] Nan Deng, Qin Zhang. “The Application of Dynamic Uncertain Causality Graph Based Diagnosis and Treatment Unification Model in the Intelligent Diagnosis and Treatment of Hepatitis B”. Mathematical Modeling and Computation Methods in Science and Engineering III 2021.
[7] Bin Feng, Qing Zhu, Mingwei Liu, Yun Li, Junxiao Zhang, Xia Fu, Yan Zhou, Maosu Li, Huaguai He, Weijun Yang. “An Efficient Graph-Based Spatio-Temporal Indexing Method for Task-Oriented Multi-Modal Scene Data Organization” ISPRS International Journal of Geo-Information 2018.
[8] Voloshin, V.I. “Introduction to Graph Theory”. Nova Science Publishers New York, NY, USA, 2009.

[9] Arpit Jain, Abhinav Sharma. “Membership function formulation methods for fuzzy logic systems: a comprehensive review”. Journal of Critical Reviews 2020.
[10] Zadeh, L.A. “The concept of a linguistic variable and its application to approximate reasoning”. Information Sciences 1975.
[11] “An Introduction to Fuzzy Logic”, link: https://www.mql5.com/en/articles/1991 accessed on 04.09.2023.

Data Mining and Big Data Analytics Using Accelerate Data

Valmir Sinani[1], Blagoj Ristevski[2]

[1,2]*University "St. Kliment Ohridski" – Bitola, Faculty of Information and Communication Technologies - Bitola, Partizanska bb., 7000, Bitola, Republic of Macedonia*

1 sinani.valmir@uklo.edu.mk; 2 blagoj.ristevski@uklo.edu.mk

Abstract:
The use of smart phones has become increasingly prevalent, with nearly every individual carrying either in their hand or pocket, making it an integral part of their daily lives. Furthermore, the quantity of sensors integrated within the smart phone is constantly expanding. Through this integration within the Smartphone, it is presented the opportunity to utilize the Smartphone for more than just communication purposes presented. One of the sensors incorporated within the Smartphone is also an accelerometer sensor, which has found significant application in various industries, greatly enhancing the quality of life. In this paper, we will present an implementation of a mobile application for collecting data utilizing using accelerometer sensor. Here is an overview of the architecture and technology used in the development of this application. Subsequently, it presents the delve into the data analysis process and utilizes the R programming language for the collected data that was collected. Diverse functions and algorithms have been employed in the data analysis such as dbScan, Extrascan, and other custom functions for step counting of walking, running, counting floors while going upstairs/downstairs.

Keywords:
Accelerometer sensor, MongoDB, dbScan, kNN, clustering algorithms

1. Introduction

The number of Smartphone users is increasingly prevalent. In the previous year, there were 6.4 billion worldwide Smartphone users' number of Smartphone users was 6.4 billion, and it is predicted that the number of Smartphone users is going to be increased to 7.7 billion which is an increase of 21 percent from 2022 [1].
In addition to various features already integrated into smart phones, sensors such as the GPS sensor, audio sensor, light sensors, directions sensor and accelerometer sensor and commonly found in these devices.

Our primary focus in this paper will revolve around the utilization of the accelerometer sensor. We will collect data from this sensor and conduct an in-depth analysis of the collected data. The application of this sensor has made significant contributions across various industries, greatly facilitating and improving the quality of life. This sensor can play a vital role in identifying the issue with roads, poetically reducing the cost of maintaining the service, while simultaneously enhancing the pedestrian and passengers' safety [2]. In relation to other research, collecting data from this sensor employed to promote health and well-being by generating various reports on activities such as daily, monthly, and weekly walking, running and upstairs and stairs activities [3]. Data generated from accelerometer sensor has been utility to analyses the characteristic of young individuals' alcohol consumption behavior [4]. As technology continues to advance, it reshapes the way it works, and this transformation in work patterns has given rise to modern challenges, including stress. To detect and measure the level of stress in the work environments, is employed an accelerometer sensor from a smart phone [5]. During the aging process, individuals may start to exhibit specific illnesses. Data generated from accelerometer sensors has been utilized to address issues related to failed detections in the elderly population. [6].

The data presented in this paper was collected from the accelerometer sensors of two distinct Android smart phones. Users generated this data while engaging in their daily activities within the living

room. They created an Android application using Xamarin Forms which generate data and then send data in real-time. To enable communication in real time, it is build an API application using NodeJS, which received information through web socket from mobile app, and then saved data in MongoDB. Then, for saving this generated data is created a NoSQL database using MongoDB. For more details about this process it will be shown in next section below.

Following the data collected and its storage in MongoDB database, the next phase involves the analysis of this data. Data analysis has been done using R programming languages, by importing them from excel. In this process, various algorithms and custom functions are employed for analyzing the collected data. Initially, we will introduce the six basic functions. Subsequently, we utility algorithms such as: dbScan, ExtraScan, HDScan, knn, hullPlot and fuzzy logic. Additionally, other functions are applied to analyze count walking steps, running, going up/downstairs.

This work is presented in section such as: In section 2 it is presented Related work which utilized the acceleration smartphone sensor in different industry. In section 3 it is presented an application implementation to track accelerometer data from Mobile App in which is presented also the technology and minor for collection data. In section 4 it is presented the part of analyzing data using different function and algorithms with R programming language.

2. Related works

The author in [7], has presented the monitoring of structural integrity and safety under extreme and normal loads using big data products generated by sensors. The building structure and bridge health is presented while earthquakes happen. The seismic scale for frequency of amplitude is measured. There is a sinusoidal wave monitor table for smartphone testing. Also, there is a test comparing the accuracy of three different smartphones compared with reference data. For the test, there were used low- and high-vibration scales.

The objective of this research [2] is to detect the quality of asphalt and cobblestone while driving the vehicle without extra effort. The problem identification is separated into three levels, such as the level of identification of pavements, the level of classification of pavement is cobalt street or dirty road and the last is to detect if classification if pavement is speed bumps, vertical patches or raised markers and other types of asphalt-obstacles. The data has been collected from two different smartphone acceleration sensors, which were installed inside the vehicle using a flexible suction holder near the dashboard. The smartphone generates data with three axes and GPS while using an application called 'Asfoult'.

The data was classified into two classes, such as: Regular and Deteriorate. The next classification is Asphalt pavement type: Cobblestone Street the Dirt Road. And the last one is to detect different types of obstacles in the street: (i) speed bump, (ii) vertical patch, (iii) raised pavement markers, and (iv) raised crosswalk. For classification is used 1NN algorithms and for measuring distance is used DTW, LCSS, DDDTW and DTDDTW.

In this research [3], the author has presented activity recognition using cell phone acceleration. The objective of this research is to detect six daily activities from users such as: walking, sitting, standing, ruing, upper stairs, and downstairs within a specific period. The data has been collected from 29 users from smartphone acceleration sensors, which carry cell phones in their pockets. By analyzing this data collection, the duty is to send weekly mail to users, to inform them their health could be a good basin on performance.

Here are some features such as: Average, Standard Deviation, Average Absolute Difference, Average Resultant Acceleration, Time Between Peaks and Binned Distribution. In order to analyze data mining there are three techniques: decision trees, logistic regression and multilayer neural networks to predict daily activities. Based on the results that are generated, the Multilayer Perceptron shows more correct data than other techniques. Based on errors that were generated, the more difficult daily activities to predict are the ones going upstairs and downstairs.

Accelerometer sensor data generation has many applications in different fields. Accelerometer data collected can help alcohol consumption to classify the drinking behavior of young adults in an urban and ecologically valid nightlife setting, as the author described in this [4] research.

In this research [5], the author has used accelerometer data to automate stress level detection in working environments from smartphones.

In this research [6], the author has proposed a system to support physicians determining an accurate elderly frail diagnosis, by collecting accelerometer data from physical activity.

3. Application Implementation to track accelerometer data from Mobile app

In this section, we will demonstrate the real-time implementation of accelerometer data from a mobile device to web charts. The architecture of the implementation of managing data accelerometers from mobile apps to web charts is presented in Fig. 1.

The mobile application is built using Xamarin forms, the back end is built using NodeJS and Database management system is built in NoSQL with MongoDB. Each part of this architecture is described in the following sessions.

3.1. NoSQL with MongoDB Database

It is used by a NoSQL database to save accelerated data. MongoDB, as NoSQL, is used to create a collection. The name of the collection for save data is called 'magnetomererdatas', as shown in Figure 1.


```
{  "_id": {    "$oid": "646ff1091a3fa33fb5edd97c"  },
  "x": -5.45649004,
  "y": 5.38566017,
  "z": -6.12025023,
  "did": "2f985c765f344777",
  "dt": "2023-05-25T23:36:40.959993Z"
}
```


Figure 1. Example of acceleration document data in MongoDB

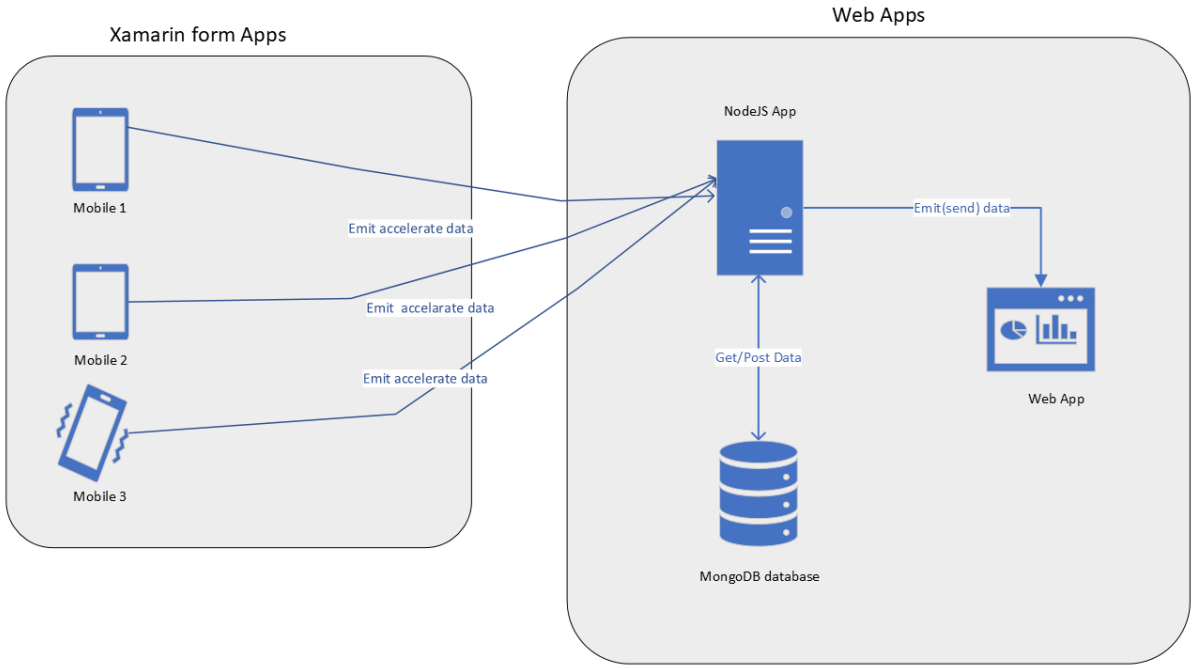


Figure 2. Mobile app and web app communication architecture

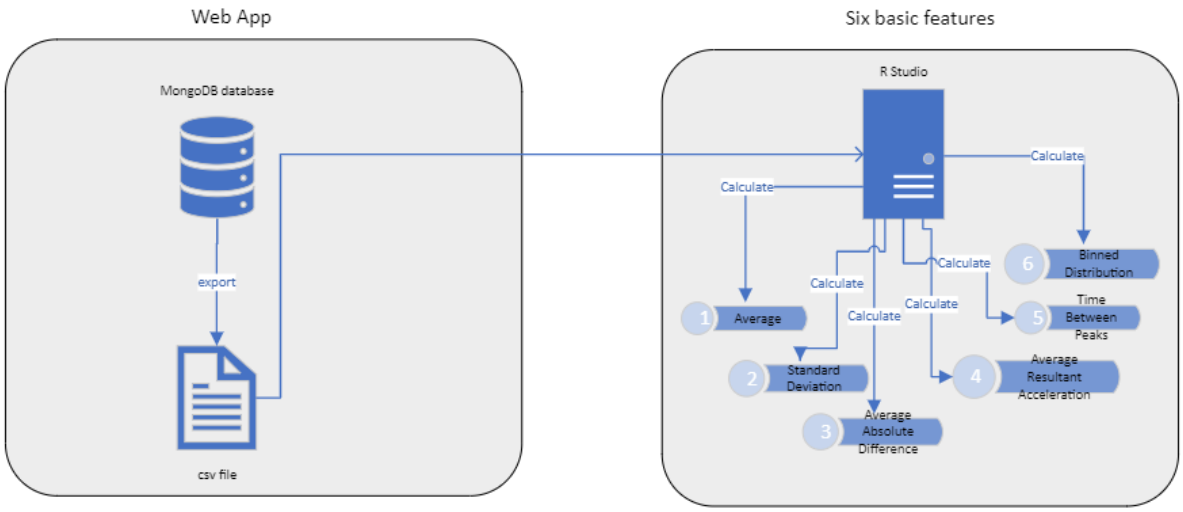


Figure 3. Web app communication and R application for generate six features

In Figure 1, there is an example document with data. The column "_id" represents a unique key within the collection, while column "x", "y" and "z" presents the values of the accelerometer sensor data that could be generated by mobile. The column "did" represents the id of the mobile device from which accelerated data was generated, and "dt" column represents the datetime in which data was generated.

Figure 2 displays the architecture of communication between MongoDB applications and R applications. The data from the MongoDB database was exported to a .csv file with data. The csv file contains the same columns that were defined in the MongoDB database schema. Then, the csv file was read late before the execution of each part of code, while analyzing data.

4. Analyses and results using R Programming

Here we have used different algorithms and function in order to analyze acceleration data and plot the results. Here is presented six basic features(Figure 3), which is mentioned below, dbScan, extraScan, HDBScan, knn, hullPlot, fuzy logic c-means, k-means and it uses a custom function to count steps that was displayed in table 1.

- The dbScan is a popular algorithm for clustering density-based clustering, which could be used from different tools such as ELKIT, scikit-learn, R, Weka and many others. dbScan is used to measures minimum density based on predefined minimum number of point known as minPts, within specific radius ε (which is fixed distance). A point which couldn't be achieved the minPts object numbers are considered as Noise point [8].
- The Knn (K-nears Neighbors) is part of lazy algorithms. knn uses the Euclidean distance to connect 2 points. Knn algorithm is non-parametric function for classification [9].
- The k-means algorithm is considered the most popular algorithm for clustering data with k groups (with similar attributes) [10].
- The fuzzy c-means, which is showing in Figure 7, is a soft clustering algorithm with each data in a not defined clustering group, each data could be part of more than two groups by percentage (0-1) [11].

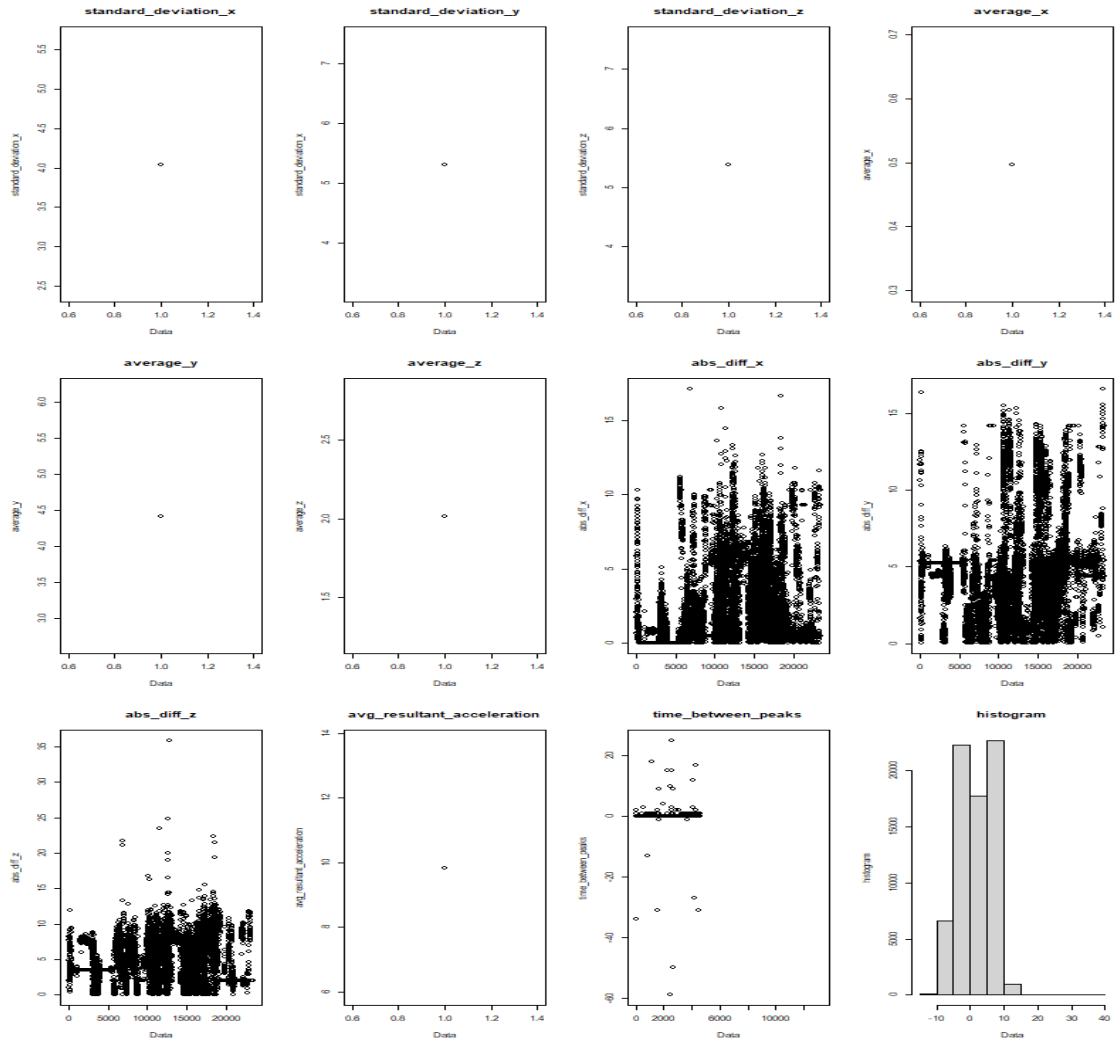


Figure 4. Shows the result of plotting six basic features (1.Average, 2.Standard Deviation, 3.Average Absolute Difference, 4.Average Resultant Acceleration, 5.Time Between Peaks, 6.Binned Distribution).

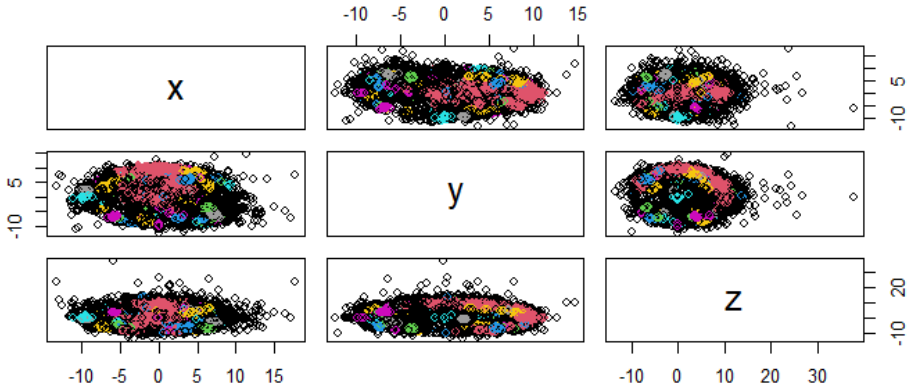


Figure 5 Result of clustering with dbScan

Figure 4 shows six basic features, such as Average, Standard Deviation, Average Absolute Difference, Average Resultant Acceleration, Time Between Peaks, and Binned Distribution.

The dbScan is shown in Figure 4. By acronym the dbScan means **D**ensity-**b**ased **S**patial **C**lustering of **A**pplication with **N**oise.

The algorithm dbscan find the clustering within data through density-base expression points [12]. This Figure 5 is displayed by execution the code below:

```
R >rdbscan_model<- dbscan(accel_features, eps = 0.5, minPts = 5)
R >rpairs(accel_features, col = dbscan_model$cluster+ 1L)
```

In Figure 6 is displayed the ExtrectdbScan, hdbscan, knn with k=5, kNNdistplot, fpccluster , by reading accelerometer data as was described above.

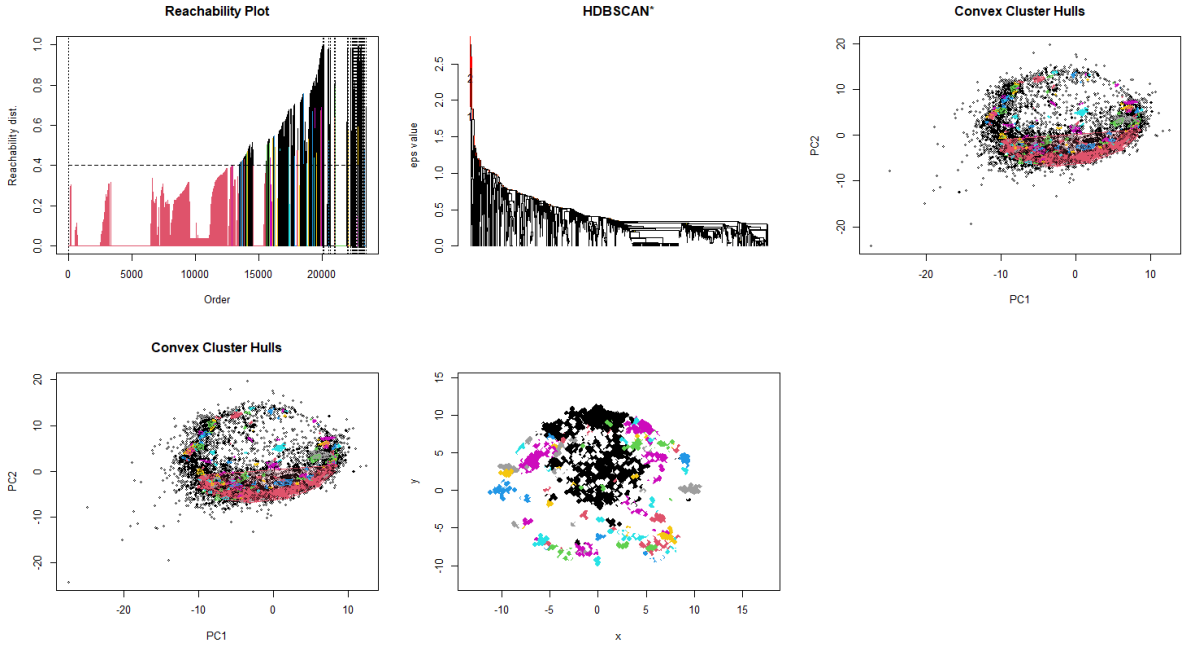
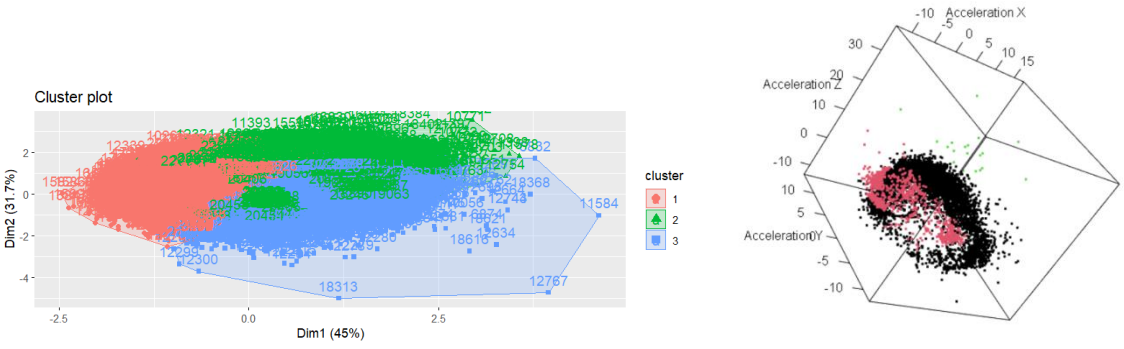


Figure 6 Result of plot: extractDBSCAN, HDBSCAN, knn k=5, hullplot.

Figure 5 shows some algorithms from the package 'dbscan' family. The extract DBSCAN Algorithms extract the cluster like DBSCAN, the letter difference is that this algorithm generates litter noise [13]. The hdbscan, which is seen in Figure 6, is a hierarchical dbscan, which is an improved algorithm of dbscanner[14]. The hullplot algorithms classify the data into two convex hull groups [15]. In order to display the result in Figure 5 and Figure 6, it is needed to access the part of code from [16] using R application:



a) b)

Figure 7. Result of fuzzy logic c-means algorithm with silhouette coefficient(a) and Visualize the clusters in a 3D scatter plot (b).

Table 1

Result of calculating daily activity

Walking steps	Running steps	Stairs steps
607	230	0

In order to count the walking steps, running steps and stairs steps, as could be seen in Table 1, there are shown the two functions below, which use accelerometer data from daily activity. The walking and running steps use some function with different threshold parameters. The code for calculate daily activities is:

```
R > countSteps < - function(pthreshold){
R > step_count < - 0 #init value
R > previous_state < - “below” #init value
R > current_state < - “” #init value
R > for (i in 2:length(acceleration_magnitude)) {
if (acceleration_magnitude[i] > pthreshold) { current_state < - “above” }
else { current_state < - “below” }
if (previous_state == “below” && current_state == “above”){step_count < - step_count + 1 }
previous_state < - current_state
return(step_count)
}
R > walkedsteps< -countSteps(threshold)
R > runsteps< -countSteps(threshold_run)
R > step_count <- 0
R > for (i in 2:length(acceleration_magnitude)) {
if (acceleration_magnitude[i] > threshold) { current_state < - “above” }
else { current_state <- “below”}
if (previous_state == “below” && current_state == “above”) {
if (acceleration_magnitude[i-1] - acceleration_magnitude[i] > threshold)
{step_count step_count + 1}
}
previous_state <- current_state
}
R > c <- c(walkedspaps,runspaps,step_count)
R > df <- data.frame(work=c[1],run=c[2],stairs=c[3])
```

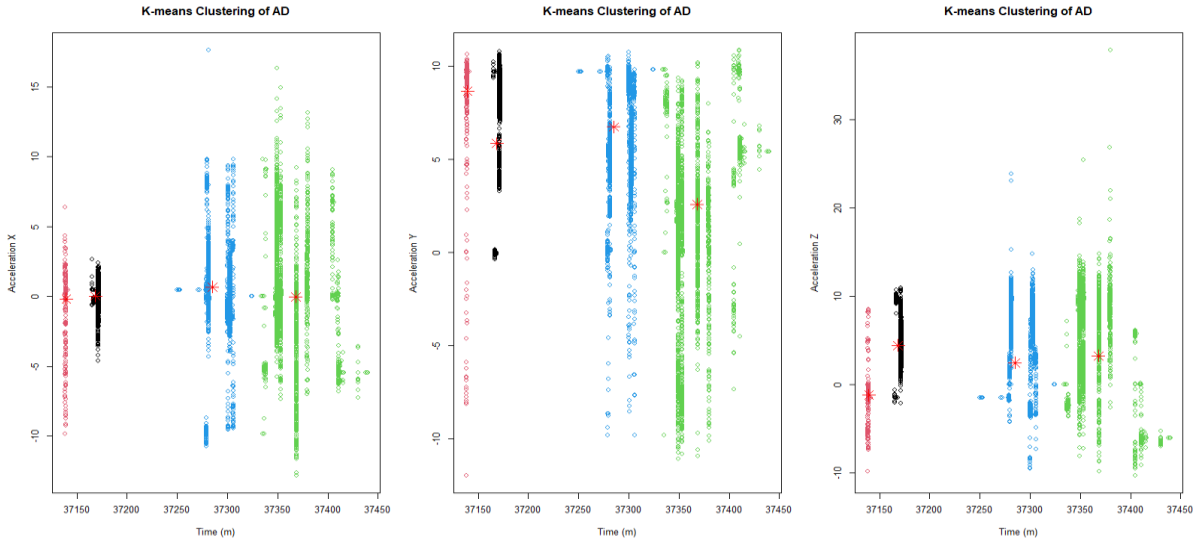


Figure 8 Result of displaying k-means using acceleration data with three axes.

In order to display the k-means (Figure 8) in three axes using acceleration data by executing the code below: the sign with red color is presented by the clusters.

5. Conclusion and future work

The extensive uses of smartphones and integration of numerous sensors within smartphones that generate valuable data, have opened up vast opportunities for researchers in the field of data mining and many industries.

In this presentation is shown the related work that involves the utilization of accelerator sensor data. It is a demonstration of a wide range of implementation across various industries, continuing to the enhancement, assisting and facilitating of various aspects of human life.

This presentation highlights a smartphone application for the collection of data using accelerometer sensor in the smartphone. It illustrates the process of saving and processing this collection data. The creation of this application involves the careful selection of appropriate technology efficient process and storing the collected data. Then, to analyze efficiently data collected, were employed a various of information and functions. These tools aid in data visualization, clarification, and extraction valuable insight.

In this presentation, the analysis primarily focused on accelerometer sensor data in isolation. However, it's worth noting that there are significant potential benefits in combining this data with information from other sensors for comprehensive understanding and broader applications.

The combination of accelerometer sensor data could involve various aspects, such as:

- The combination of this data with high-quality air data can be instrumental in extracting valuable insights. For instance, it can provide a deeper understanding of activities like walking and running and their associated health benefits.
- Integration of this data sensor with heart rate monitoring holds significant potential for health and fitness tracking. This correlation between physical activity and heart rate changes can serve as a preventive measure against various heart diseases.
- Combining smartphone sensor data with other acceleration sensor data can be valuable for various purposes, including combining data from large objects with multiple floors to aid in orientation and navigation.
- Information reports could be automated by dependent based on configurations derived from other predefined sensors. Such automation has the potential to safeguard human lives and assets, save valuable time and resources, and enhance service efficiency.

References:

[1] R. Ablamowicz, B. Fauser, Clifford: a maple 11 package for Clifford algebra computations, version 11, 2007. URL: http://math.tntech.edu/rafal/cli11/index.html.

[2] Vinicius M.A. Souza, Asphalt pavement classification using smartphone accelerometer and Complexity Invariant Distance, Engineering Applications of Artificial Intelligence, Volume 74, 2018, Pages 198-211, ISSN 0952-1976, https://doi.org/10.1016/j.engappai.2018.06.003.

[3] Kwapisz, Jennifer R., Gary M. Weiss, and Samuel A. Moore. "Activity recognition using cell phone accelerometers." ACM SigKDD Explorations Newsletter 12.2 (2011): 74-82.

[4] Santani, Darshan, et al. "DrinkSense: Characterizing youth drinking behavior using smartphones." IEEE Transactions on Mobile Computing 17.10 (2018): 2279-2292.

[5] Garcia-Ceja, Enrique, Venet Osmani, and Oscar Mayora. "Automatic stress detection in working environments from smartphones' accelerometer data: a first step." IEEE journal of biomedical and health informatics 20.4 (2015): 1053-1060.

[6] Fontecha, Jesús, et al. "Elderly frailty detection by using accelerometer-enabled smartphones and clinical information records." Personal and ubiquitous computing 17 (2013): 1073-1083.

[7] Feng, M.; Fukuda, Y.; Mizuta, M.; Ozer, E. Citizen Sensors for SHM: Use of Accelerometer Data from Smartphones. Sensors 2015, 15, 2980-2998. https://doi.org/10.3390/s150202980
[8] Schubert, E., Sander, J., Ester, M., Kriegel, H. P., & Xu, X. (2017). DBSCAN revisited, revisited: why and how you should (still) use DBSCAN. *ACM Transactions on Database Systems (TODS)*, *42*(3), 1-21.
[9] Pandey, A., & Jain, A. (2017). Comparative analysis of KNN algorithm using various normalization techniques. *International Journal of Computer Network and Information Security*, *11*(11), 36.
[10] "K-Means Clustering in R Tutorial - DataCamp." https://www.datacamp.com/tutorial/k-means-clustering-r. Accessed 20 Jul. 2023.
[11] Partitioning Cluster Analysis Using Fuzzy C-Means" 20 Nov. 2017, https://cran.r-roject.org/web/packages/ppclust/vignettes/fcm.html. Accessed 19 Jul. 2023.
[12] Fontecha, Jesús, et al. "Elderly frailty detection by using accelerometer-enabled smartphones and clinical information records." Personal and ubiquitous computing 17 (2013): 1073-1083.
[13] Hahsler, Michael, Matthew Piekenbrock, and Derek Doran. "dbscan: Fast density-based clustering with R." Journal of Statistical Software 91 (2019): 1-30.
[14] Campello, Ricardo JGB, Davoud Moulavi, and Jörg Sander. "Density-based clustering based on hierarchical density estimates." Pacific-Asia conference on knowledge discovery and data mining. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013.
[15] "hullplot: Plot Convex Hulls of Clusters – Rdocumentation" https://www.rdocumentation.org/packages/dbscan/versions/1.1-11/topics/hullplot. Accessed 19Jul. 2023. "Density-Based Clustering Based on Hierarchical Density Estimates." https://link.springer.com/chapter/10.1007/978-3-642-37456-2_14. Accessed 19 Jul. 2023.
[16] "Number of smartphone mobile network subscriptions worldwide from 2016 to 2022" https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/. Accessed 19 Jul. 2023.

NoSQL Database Support to Big Data Storage Systems and Specific Use Cases: a Review

Valmir Sinani[1]

[1]*University "St. Kliment Ohridski" – Bitola, Faculty of Information and Communication Technologies - Bitola, Partisans bb., 7000, Bitola, Republic of Macedonia*

1 sinani.valmir@uklo.edu.mk;

Abstract:
The amount of data being generated every day is increasing exponentially, and systems for managing this increase are adopting to keep up. The systems for managing this growth are known as big data storage systems and NoSQL databases.
In this paper, our focus is on a class of systems developed to manage large amounts of data. We will also show how big companies faced challenges in managing large amounts of data and how they found solutions through the use of NoSQL systems. Using NoSQL systems, we will show characteristics of NoSQL systems, including those related to distributed databases, models, and query languages. We categorized NoSQL systems into four main types: Document-based NoSQL systems, NoSQL key-value store, Column based or wide column NoSQL systems, Graph-based NoSQL Systems, are other large systems that are not easy to categorize into one of four main systems. In the third part, we will show more details about Document-Based NoSQL Systems and MongoDB as a popular document-based NoSQL database system. In third part of MongoDB, we will show an example of MongoDB Modeling, crud operation and replication.
In the last part, it will be provided a thorough overview of a review paper that discusses suitable database scenarios for managing big data storage.

Keywords:
NOSQL Systems, Big Data, MongoDB, Replication

Introduction

In this seminar, we will focus on a class of systems developed to manage large amounts of data, such as social media applications, Facebook, twitter, and email.

The term NOSQL is an acronym for "Not only SQL" and is used to refer to systems that were designed to manage large amounts of data. These systems are often used to convert applications which were developed for more traditional data management.

Most NOSQL systems are designed to work with distributed data storage, high availability, data replication and scalability even though they lack the consistency, data structure and powerful query language provided by traditional databases.

Emergence of NOSQL Systems

Many companies are faced with applications that generate large amounts of data, such as an email application which was developed by companies like Google and Yahoo, which has millions of users, with each user properly having thousands of messages. Therefore, other systems are necessary to store and manage these large amounts of data. There are several reasons why SQL standards are not suitable for handling large amounts of data:

- SQL systems offer many services such as powerful query and concurrency control, which may not be needed for these types of applications that generate large amounts of data.

- The Sql-based systems offer structured data models, like traditional models, which may be too restrictive for those applications that generate large amounts of semi-structured data.

An example of this can be seen in applications such as Facebook, which has a large number of users who post a variety of content including photos, videos, and other types of media. These posts are then displayed on the walls of other uses using social media relationships. They manage this large data collection and high availability, while traditional relational systems are not suitable for this type of system [1].

Some companies that are faced with huge amount of data coping with relational databases decide to develop their own systems [5]:

- **Google** has developed its own NOSQL system known as BigTable, which was used by Google to develop applications that needed vast amounts of data storage, such as GMail, Google Maps, Google Earth, Web indexing, Google finance and dozens of others[6]. Google innovated a new NOSQL system category known as Column-based or wide-column for storing data [1].
- **Amazon** has also developed its own NOSQL systems known as DynamoDB which could be used by amazon cloud service. DynamoDB is being used by Amazon developers for many systems, such as Alexa, Amazon.com. Amazon innovated a new category of NOSQL systems known as key-value data storage [7].
- **Facebook** developed a NOSQL system called Cassandra, which is an open source and is known as Apache Cassandra. Cassandra is a distributed storage system used for maintaining large amount of data across many servers with no single point of failure. This NOSQL System utilizes key-value and column-based system types. Cassandra run in cheap hardware resources with high speed without having effect in reading.[8]

Another category of NOSQL systems are graph-based NoSQL systems, examples of which include Neo4j and GraphBase. NoSql systems such as OrientDB, combine several of the concepts mentioned above [1].

2.1. Characteristics of NOSQL Systems

This chapter will describe 2 types of characteristics of NoSQL systems: those related to distributed databases and systems, and those related to data models and query languages.

NoSQL Characteristics relate to distributed databases and distributed systems.

NoSQL systems are known for their high availability, scalability, and high performance.

Here are more details about these characteristics:

1. **Scalability**: there are two types of scalabilities in distributed systems: horizontal and vertical scalability. Horizontal scalability is more commonly used because it allows for more data growing by adding additional nodes to the system. To achieve higher level of available and scalability, systems often need to scarify the consistency. On the other hand, vertical scalability is focused on improving the computational power of existing nodes. Horizontal scalability can be achieved while the system is running, and there is a necessary function to synchronize data between nodes [9][1].
2. **Availability**, Replication and Eventual Consistency: Most applications that used NoSQL systems require continuous availability and to achieve this, distributed nodes need to be synchronized. In case one node fails, another node can take over this task. Replication improves availability and read performance by allowing reads from replicated data nodes. Writing nodes will become more complex as changes made to one node will need to be replicated in other nodes. Replication is achieved by adding new copies of nodes, thereby enhancing availability, and ensuring synchronization among at least 80% of nodes for consistency. This is accomplished through the use of node "versioning" to maintain consistency.[10].
3. **Replication Models**: There are two main types of model replication: Master-slave replication and master-master replication. In **master-slave replication**, one copy is designated as master copy and other copy, known as slave copies, are replicated from the master. The replicated process is called eventual control. There are two ways to configure read operations in master-slave replication:

distributed concurrency control, where all reads are performed on master copy, and non-guaranteed reed, where reads can be performed on slave copy, so this method does not guarantee that the values are latest updated. **Master-master replication** allows read and write operations to be performed from any copy but does not guarantee that all copies will have the same data. A recognition method should be applied to synchronize different copy nodes. A Lorq algorithm based on consistency it is proposed with the aim of striking a balance between the quantity of service (QoS) and quality of data (QoD). A Lorq algorithm is used to provide strong consistency and eventually consistency in data replication system to obtain high performance and availability.[1][11].

2.2. NoSQL characteristics related to data models and query languages.

Here are some of these characteristics:

1. **Not Requiring a Schema**: some NoSQL systems do not require a strict data structure schema and allow saving data as semi-structure schema. Users could specify a partial schema, but most NoSQL systems such as key-value and document-store data model do not require it. Some programming languages, such as JSON and XML, are often used to present semi-structure data. [1][11]
2. **Less Powerful Query Languages**: Due to variations in the features supported by different data stores, a common query language is not currently available. Most NoSQL systems do not require advanced query language level like traditional SQL, because searching is often done by directly retrieving specific objects from specific files based on unique keys. NoSQL systems support API applications for read/write operation by offering CRUD or SCRUD operations. Most NoSQL systems do not offer JOIN operations.[1][11]
3. **Versioning**: Algorithm is used to cope with versioning and to resolve conflicting values in different stages. Some NoSQL systems support multiple versioning, by adding a timestamp, which shows when the data was created. [12].

2.3. Categories of NoSQL Systems

NoSQL systems are generally categorized into four main types [4]:

1. **Document-based NOSQL systems**: Those systems store data in format like JSON. Accessibility to documents is achieved through unique Document Id and Indexing. Saved data as collection of documents and simple query mechanism is supported. [4][9]
2. **NoSQL key-value store**: These systems are designed to provide fast access data using key value pairs. The value could be a record, an object, a document, or more complex data structure. The key is string and value is a data, it is similar to dictionary [4].
3. **Column based or wide column NoSQL systems**: Those NoSQL systems have a simple model for storing data in separate tables, with columns, each column contains its own file. This method is also referred as data compression within the distributed column store column is stored on the different disk. Data is stored using versioning. Those Systems are well-suitable for large- scale database and data intensity application, such as data warehousing and business intelligent.[13] [1]
4. **Graph-based NoSQL Systems** - In data representation, Graph NoSQL systems store data as graphs. Nodes can be located by traversing edges using path expression.

Additionally, there are other large systems that are not easy to categorize into one of four main systems. Here are some such systems:

5. **Hybrid NoSQL systems**: Those systems are part of two or more categories that we mentioned above.
6. **Object Databases**.
7. **XML Databases**.

Key-based search engines store vast amounts of data with fast access. This type of data store can be classified as large NoSQL systems big data storage.[1]

Table 1

One possible database for manage in/out employee relational database schema.

b) Employee Status

Id	Description
1	Aktiv
2	Resignation
3	Termination

g) Teams

Id	Code	Description
1	1	Team
1	2	Tema 2
3	3	Microsoft

c) Event

Id	Descriptions	InUse
1	Entry	1
4	Entry 2	0
3	Exit	1

a) Departments

Id	Code	Description
1	MNG	Manage.
4	MSC	Board
101	OPS	Programs.

f) Team Employees

Id	TeamId	EmployeeId
1	1	15
2	1	17
3	1	101

d) Employees

Id	TeamId	EmployeeId
1	1	15
2	1	17
3	1	101

***e)* In Out Activities**

Id	Code	Name	Bdate	Salary	DepId	EventID
1	HR-d00001	Augustuschul.	1/9/1965	30000	2	102
2	HR-s00002	ElishaHilpert	5/2/1975	45000	2	102
3	HR-s00012	AlyceLehner	8/22/1985	32000	3	101

Document-Based NOSQL Systems and MongoDB

In NOSQL Systems, Document-Based databases store data as a collection of similar documents and are known as document stores. Document base database systems store data as key-value pairs in JSON forma. Each document within has a special key, “Id”, which ensures uniqueness. [14] MongoDB is an open-source document-oriented database system initially developed by DoubleClick in 2007 and is currently owned by Google.[4]. Document-based systems differ from other types of systems, such as XML, because document-based systems do not specify a schema and the document structure is flexible and self-describing. The documents within a document can remain the same, and adding new documents with different structures does not affect the structure of the documents below.[2]. The user can specify the indexing in the system for some elements in the document. Document within the collection is simple and can contain different attributes, and documents between one other could not contain the same attributes. The popular language to specify documents in the NOSQL system is JSON (JavaScript Object Notation).[1].

There are many document-based NoSQL Systems, such as MongoDB, CouchDB, and others, systems that have different models and languages.

To illustrate the advantages of MongoDB in comparison to other relational databases,a detailed comparison will be presented in the following paragraph. Also, a comparative analysis will be conducted between MongoDB and MySQL, utilizing various cloud instances and heterogeneous data collected from air sensors based on specific criteria. In this experiment is shown that MySQL in small data and high server performance isn’t any different in performance. While the dataset increases, it is observed that MongoDB exhibits superior performance. [2] Based on results returned by exhibiting the

comparison between MongoDB and MsSQL in different dataset based on specific criteria shows better performance of MongoDB. The criteria utilized for this comparison include Bulk insert, simple query and aggregation query and different data set sizes of 1M, 5M and 10M datasets [3].

3.1. MongoDB Data Model

MongoDB saves data in BSON format, which is more efficient than JSON and has more data types.

During the presentation of an example data model in MongoDB, we will base ourselves on Figure 1. Table 1 presents a data model database schema for a system for managing the entries and exits of workers of an institution, company, or organization.

The Command for 'createCollection' is used to define a collection. As an example, here is a creation of a collection called Team.

```
db.createCollection("teams", { capped: true, size:5242880, max:2000 } ))
```

The first parameter "teams" is the name of the collection. Other other parameters, such as capped, are used to set the maximum size of collection, and max is used to set the max number of documents.

Here is another example where we will create a document collection called employees to hold information about employees which are part of one team.

```
db.createCollection("employees", { capped:true, size:5242880, max:2000 } )
```

Each document in the collection has a unique field ObjectId(_Id) which can be generated by systems formats (16 bytes, which is a combination of a 4-byte timestamp, a 3-byte node ID, a 2-byte process ID, and a 3-byte counter) or user could specify the format.

A collection does not have a schema. Users can decide to follow a normalized design structure form or denormalized design.

In Figure 1 a), a connection between team and employee is shown using a denormalized structure form. The field _id, which is defined by the user, starts with "T" in the "team" collection, and starts with "E" in the "employee" collection.

Based on Figure 1 a), it is not necessary to create a new collection to hold employees, as they are embedded within team collections in an array format.

Another form of structure design is shown in Figure 1 b), where the reference of the employee's _id is used in the collection of teams, represented as an array within brackets. Figure 2 c) is used as the first form of normalization, as the reference of the team is within the documents of the employee.

Figure 2 c) does not represent a full form of normalization, as in cases where one employee is a part of multiple teams, it would be necessary to create another collection, "employee_team" to manage this scenario.

a) Team Document with an array of Employee documents within

```
{   _id: "T1", Code: 1, Description: "Team 1", InUse: true, Employees: [
                                {Name: "Augustus Schulist},
                                {Name: "Elisha Lehner", Salary: 28000}
                            ]
    }
```

b) Employee id documents that are presented as an array within the Team document

```
{_id: "T1", Code: 1, Description: "Team 1", InUse: true, Employees:
["E1", "E2"]   }
{_id: "E1", Name: "Augustus Schulist"}
{_id: "E2", Name: "Elisha Lehner", Salary: 28000}
```

c) Normalized teams and employee documents in first normalized design form.

```
{_id: "T1", Code: 1, Description: "Team 1"}
{_id: "E1", Name: "Augustus Schulist", TeamId:"T1"}
{_id: "E2", Name: "Elisha Lehner", Salary: 28000}
```

Figure 1 Example of simple documents in MongoDB. **a**) Denormalized form design with subdocuments within, **b**) Reference of _id documents that are presented as an array within the document, **c)** Normalized documents.

d) inserting of documents from in the collection "employees" and "teams".

```
db.teams.insert({_id: "T1", Code: 1, Description: "Team 1", InUse: true})
db.employees.insert([
        {_id: "E1", Name: "Augustus Schulist", TeamId : "T1"  }
        {_id: "E2", Name: "Elisha Lehner", Salary: 28000}
            ])
```

Figure 2 (**d**) inserting of documents in the collection employees and teams.

MongoDB CRUD Operations

Please MongoDB has several operations for CRUD (Create, Read, Update and Delete). Documents in MongoDB can be created and inserted using the 'create' and 'insert' operations, whose format is:

db.<colection_name>.insertOne(<document>), or
db.<colection_name>.insertMany([<documents>])

An example of an insert operation is presented in Figure 2 d). In this example, a multi-array document is inserted. The delete operation called deleteOne (previously called remove) or deleteMany, and is a format:

db.<colection_name>.deleteOne(<document>);
db.<colection_name>.deleteMany(<documents>)

The example of the remove operation is used below. Executing the query means that it will delete the single employee with _id equal with "E2".

```
db.employees.deleteOne({ "_id" : "E2" })
```

The update operation is called updateOne/updateMany, and the format is:

db.<colection_name>.updateOne({<filter>}, {$set:{<update_data>}})

The example of the update operation is used below. Executing the query means that it will update the code of the single employee with code equal to "HR-s00012". MongoDB also supports the update. Many operations for updating multiple documents based on filters.

```
db.employees.updateOne({Id:3},{$set:{Code:"HR-s00012"}})
```

The read could be done using update operation, and the format is:
db.<colection_name>.find({<filter>})

Executing the query means that it will find a single employee with _id equal with "E2".
db.employees.find({_id:"E2"})

Other operations, such as count, limit, and sort, can also be used to perform specific scenarios. These can be found in the MongoDB documentation. Also the snipped code for initial database which is used in this paper in resented on online repository [21]

4.1. Replication in MongoDB

The concept of a replica set is used to create a copy of the same data in different nodes in a distributed database, using a variant of the master-slave model.

Here we will show an example of creation replication in MongoDB using docker. First, we will pull a MongoDB image from Docker Hub using 'docker pull mongo' and then start a MongoDB container with the '--replSet' options:

```
docker run -d --name mongo-node-1 -p 27017:27017 mongo --replSet "rs0"
docker run -d --name mongo-node-2 -p 27018:27017 mongo --replSet "rs0"
docker run -d --name mongo-node-3 -p 27019:27017 mongo --replSet "rs0"
```

By execution the example above, we will create a 3 docker images and then we can add to a replica set by executing the code below :

```
rs.initiate()
rs.add("mongo-node-2:27017")
rs.add("mongo-node-3:27017")
```

And then, when we change our primary node (mongo-node-1), it means that the other nodes (mongo-node-2, mongo-node-3) will be replicated from primary nodes.

A replica set will have one primary copy of the collection employee stored in node called 'mongo-node-1' and two other secondary copies of the employee collection stored in nodes called 'mongo-node-2' and 'mongo-node-3'.

Review and discussion

The author in this paper [15] has presented a design as a solution for saving and managing big data, which is known as IOTMDB System Architecture. Before explaining this architecture, the author explains some features of IoT Data, such as: Muti-source and heterogene, temporal spatial correlation, interoperability and multi-dimensional. Here are also some NoSQL databases and their specific, and comparison between Relational databases and Non-relational Databases systems. The IoTMDB consists of four parts: Master Node, Standby Node, Data Replication Node and Slave Node, where each of them has specific responsibilities and interacts with each other. One of the advantages of this architecture, which enables fast performance, is the segregation of data in two categories: lightweight and multimedia. Additionally, the architecture employs a single point of failure, and the process involves slave nodes. This architecture also encompasses different IoT application quarry types, including Historical Data retrieval and Tracking quarries.

The objective of this research [16] is to find an effective way for saving and managing heterogeneous big data, comparing MySQL and NoSQL Database Systems. For comparison, there are used three criteria of evaluation: Evaluation of manipulation of heterogeneous data, evaluation of increasing workload and evaluation in level of cloud. The test data utilized in this research comprises air pollutant reading obtained from sensors located in various places. That data was stored in three tables within both

MySQL and NoSQL databases. The corporation was done using three different instances on cloud servers. In scenarios with a small number of sensors and high performance, the MongoDB and MySQL exhibit similar performance levels. In case the number of sensors is increased, it seems that MongoDB appears to deliver superior performance compared to MySQL. Based on the mentioned criteria, MongoDB outperforms MySQL in terms of low resource costs.

The main duty of this research [17] is to find a suitable NoSQL database management system in an IoT application. The author has used five criteria while comparing NoSQL database, including: Scalability, Realtime processing, Security, Aggregation and Spatial Data Handling. And some second criteria include: Data distributions, quarry languages and applicable domain. The NoSQL Databases that were compared in this paper are: Redis, Cassandra, Couchbase, MongoDB and Neo4J. The author has described each database based on criteria that was mentioned here. Based on five criteria, it seems that CouchBase considers a database more suitable for IoT Application. CouchBase demonstrates superior performance in spatial data handling, scalability, and real-time processing compared to MongoDB, while Neo4J is well-suitable for managing spatial data.

In this research [18] the author claims that MongoDB is more suitable for TSDMS compared to Relational databases. Here is how MongoDB is a solution for some projects that need scalability, high availability and that manipulate with big data. Relation database is not designed and is not suitable workload that is needed for Telco Subscriber database management. The current solution was developed using RDBMS and Oracle. The new proposed solution suggests storing data in MongoDB, while keeping the component unchanged, still utilizing PostgressSQL. The test here was done in three groups of data: 1M, 5M and 10M data. The evaluation criteria is done using four types: Bulk Insert, ETL (Extract, Transfer, Load), Simple Query, and Aggregation Queries. Based on the result that was generated, it was shown that Bulk Insert, Simple Query and ETI MongoDB perform better than Relational databases, while RDBMS perform better in aggregation queries. In conclusion, the author said that MongoDB is more suitable than RDMNS because MongoDB shows better performance in quarries that are used more often compared to aggregation quarries.

In the research [19] the author presented data management IoT. Here are some characteristics of data management in IoT. In the context of IoT applications, RDBMS is often considered as unsuitable, while NoSQL databases are considered more appropriate. Also, here it is a description of Time Series Databases and their characteristics.

The author in this research [20] has described and forecast data increasing based on IDC forecast until 2025. As a result of this data generation growing, the science community has recognized the potential of using this data to harness the valuable data, which can help companies, healthcare and security to increase their services and qualities. The primary objective of this research is to use the CQRS pattern for adapting complex requirements from the endpoint and managing enormous increase of data generated from IoT. While evaluating Time Series databases, there were the top five NoSQL Databases: InfluxDB, Kdb+, Graphite, Prometheus and RRDtool and their characteristics. InfluxDB based on features it has is a more suitable database for IoT time series database management. Big Companies like Facebook and ebay use Time series databases to oversee some specific data devices. CQRS is easily applied in IoT applications because it separates the model of writing and reading in databases. By using these advantages, MongoDB is used as a read database and PostgresSQL is used for writing data. RabbitMQ is used as a synchronization component, which transfers data between the written database (PostgresSQL) and the read database (MongoDB) as distinct intervals. Apache JMeter is utilized to conduct performance testing by retrieving data from 24 million data rows using two separate threads (100, 20 and 300). Based on the obtained results that were generated it appears the CQRS performs 1200 times better than simple architecture with a single database.

This research [21] is proposed as an effective design for processing of the data in vehicle-to-everything (V2X) using the CQRS pattern for saving and processing data in different formats. The number of applications and industries was increased, resulting in the generation of vast amounts of data in various formats, the associated costs of data management have also been escalated. RDBMS is incompatible to address the issue of enormous growing and format of data and as a solution for this the author has used the CQRS pattern. The author has implemented the CQRS pattern, utilizing a distinct database structure. Specifically, they have employed NoSQL database System for writing Purposes, while employing MSSQL database system for reading purposes. The data transfer from NoSQL to RDBMS was performed in an aggregated format due to the requirement of a fixed schema by RDBMS.

The test was done using the Azure cloud as infrastructure. Based on the result, it seems that 10000 messages don't have any advantages using CQRS, but in case when the number of messages has grown to 100000 the CQRS shows a better performance.

Conclusions

The generation of Data is growing every day from different sources, requiring advanced storage systems to manage this growth effectively. The big tech companies such as Google, Amazon and Facebook while developing, faced challenges in managing large amounts of data as they grew. They found solutions through the use of NoSQL systems, some of which are open-source and available for use worldwide. Characteristics of NoSQL systems have attributes which are compatible to support large amounts of unstructured and semi structured data. NoSQL systems provide high performance, scalability and availability.

In NOSQL Systems, Document-Based databases store data as a collection of similar documents. Document-based systems do not specify a schema and the document structure is flexible and self-describing.

There is no specific NoSQL database solution for all cases of big data storage and processing. Every NoSQL database has specific advantages, disadvantages and their specifics. Therefore, depending on the specific use cases, it can be recommended appropriate NoSQL database with their respective advantages tailored to those scenarios.

Conditioning specific requirement in big data storage and processing, include substantial amount of unstructured and semi-structured data, relational database system may not be well-suited for big data storage due to limitations in scalability and their rigs data model. In certain situations, combining relation database with NoSQL could be a suitable approach.

MongoDB is the most popular document-based NoSQL database system for managing big data storages. MongoDB is document-based systems is schema-free, and its document structure is self-describing. By adding new documents with different structures does not affect the structure of the documents.

References:

[1] Elmasri, Ramez. "Fundamentals of database systems seventh edition." (2021). Chapter 24, pages 883 – 910.

[2] Eyada, Mahmoud Moustafa, et al. "Performance evaluation of IoT data management using MongoDB versus MySQL databases in different cloud environments." IEEE access 8 (2020): 110656-110668.

[3] Nichie, Aaron, and Heung-Seo Koo. "A Comparison of Performance Between MSSQL Server and MongoDB for Telco Subscriber Data Management." 65.3 (2016): 469-476.

[4] Amghar, Souad, Safae Cherdal, and Salma Mouline. "Which NoSQL database for IoT applications?." 2018 international conference on selected topics in mobile and wireless networking (mownet). IEEE, 2018.

[5] Moniruzzaman, A. B. M., and Syed Akhter Hossain. "Nosql database: New era of databases for big data analytics-classification, characteristics and comparison." arXiv preprint arXiv:1307.0191 (2013).

[6] Chang, Fay, et al. "Bigtable: A distributed storage system for structured data." ACM Transactions on Computer Systems (TOCS) 26.2 (2008): 1-26.

[7] Elhemali, Mostafa, et al. "Amazon {DynamoDB}: A Scalable, Predictably Performant, and Fully Managed {NoSQL} Database Service." 2022 USENIX Annual Technical Conference (USENIX ATC 22). 2022.

[8] Avinash Lakshman and Prashant Malik. 2010. Cassandra: a decentralized structured storage system. SIGOPS Oper. Syst. Rev. 44, 2 (April 2010), 35–40. https://doi.org/10.1145/1773912.1773922.

[9] Cattell, Rick. "Scalable SQL and NoSQL data stores." Acm Sigmod Record 39.4 (2011): 12-27..

[10] Chandra, Deka Ganesh. "BASE analysis of NoSQL database." Future Generation Computer Systems 52 (2015): 13-21.
[11] Pankowski, Tadeusz. "Consistency and availability of Data in replicated NoSQL databases." 2015 International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE). IEEE, 2015.
[12] Pankowski, Tadeusz. "Consistency and availability of Data in replicated NoSQL databases." 2015 International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE). IEEE, 2015.
[13] Daniel J. Abadi, Peter A. Boncz, and Stavros Harizopoulos. 2009. Column-oriented database systems. Proc. VLDB Endow. 2, 2 (August 2009), 1664–1665. https://doi.org/10.14778/1687553.1687625.
[14] Hecht, Robin, and Stefan Jablonski. "NoSQL evaluation: A use case-oriented survey." 2011 International Conference on Cloud and Service Computing. IEEE, 2011.
[15] Li, Tingli, et al. "A storage solution for massive iot data based on nosql." 2012 IEEE International conference on green computing and communications. IEEE, 2012.
[16] Eyada, Mahmoud Moustafa, et al. "Performance evaluation of IoT data management using MongoDB versus MySQL databases in different cloud environments." IEEE access 8 (2020): 110656-110668.
[17] Amghar, Souad, Safae Cherdal, and Salma Mouline. "Which NoSQL database for IoT applications?." 2018 international conference on selected topics in mobile and wireless networking (mownet). IEEE, 2018.
[18] Nichie, Aaron, and Heung-Seo Koo. "A Comparison of Performance Between MSSQL Server and MongoDB for Telco Subscriber Data Management." 전기학회논문지 65.3 (2016): 469-476.
[19] Nasar, Mohammad, and Mohammad Abu Kausar. "Suitability of influxdb database for iot applications." International Journal of Innovative Technology and Exploring Engineering 8.10 (2019): 1850-1857.
[20] "CQRS and Event Sourcing for the IoT - Sense Tecnic." 5 Sep. 2017, http://sensetecnic.com/cqrs-and-event-sourcing-for-the-iot/. Accessed 12 Jun. 2023.
[21] Choi, Sangkon Han1and Jung-In. "Data Processing System Using CQRS Pattern and NoSQL in V2X Environment." (2020).
[22] V. Sinani. 2023. Repository of InOut Database product case. https://github.com/valmirsinani/CheckInOutDBInMongoDB. Accessed 5 Apr. 2023.

New Possibilities of Applying Digital Marketing in Business

Duda Balje [1]

[1]*University Educons, Novi Sad, R. Serbia*

duda.balje@gmail.com

Abstract:

In this scientific-research work, the possible aspects of the influence of information communication technologies and the process of human resources management in the application of modern methods of internet marketing and branding are examined theoretically and empirically with the aim of helping small and medium-sized enterprises in this area to improve their business and market more easily. its products and services. Business conditions in the modern world pose enormous business challenges to companies, how to maintain their business in the conditions of an increasingly open market and global competition, keep business clients and improve business by reducing the total costs of business systems. Information and communication technologies are being imposed as one of the solutions to solve the business challenges, however, their application in business has not yet been investigated to the necessary and sufficient extent. The latest research in practice indicates that a huge percentage of projects introducing information and communication technologies for business purposes ended in failure precisely because of the lack of an adequate introduction methodology. The aim of this paper is to investigate the influence of human resources, the Internet and social networks as an effective means of online marketing and increased visibility; and some approaches as a way to attract new customers and support sales at all stages of the buying process. The initial thesis of this research work is that small and medium-sized enterprises are viewed as business systems that sell-provide their products/services via the Internet, using social networks such as Facebook, Instagram, etc. which provide wide opportunities for business success. In this sense, relevant domestic and foreign literature, as well as the latest research in the world in this field, were studiously analyzed for the purposes of the work. Based on the obtained results, it is evident that information and communication technologies and specifically the Internet network offer numerous challenges and business opportunities, thereby realizing a greater number of potential users. Empirical research consists of interviews and questionnaires, which were conducted in certain companies. In this direction, information obtained from conversations with key users was used for qualitative analysis of the impact of Internet marketing and human resources on business, while information obtained from questionnaire results was used for quantitative analysis of business. At the end of the paper, concluding considerations with key findings are given, and new opportunities for the influence of information and communication technologies and human resource management in Internet marketing, as well as limitations in the implementation of research, are presented.

Keywords:

Internet, web marketing, human resources management,

1. Introduction

The Internet is one of the fastest growing mediums of today's time, which provides great opportunities in the field of business, banking, information exchange and much more by bringing many innovations in learning about new technologies and developing new skills and facing new challenges. In the early 1990s, the Internet shifted to commercial purposes, as political and technical experts saw its potential for use by the broad social masses. The number of Internet users in the world and in Serbia is

growing rapidly. A similar case is also in our environment, where in recent years social networks have become a key factor in the digital marketing strategy in Kosovo, and activity on them brings measurable results, whether it is about sales, strengthening the brand itself, or promoting political strategies. The development of e-business, and especially the development of the Internet and the introduction of Internet commerce, undoubtedly significantly contributed to changes in consumer habits, and thus to changes in advertising patterns. For the purpose of promotion on the Internet, in the relevant literature it can be found that the Internet and the ability to search have changed this activity forever, so that now fers are available directly through the Internet and their websites [8]. The main reason for the emergence and rapid development of social media advertising is its highly relevant targeting techniques, which are based on users' personal data [10]. In this sense, some authors point out that companies that appear on the Internet are more and better seen by customers, better directed, informed, more refined and better technologically adapted to the younger population [16]. In this sense, the academic literature takes a rather holistic approach: Dehghani and Tumer (2015) propose a theoretical framework that suggests that social media advertising boosts sales through brand image and market equity [4]. The traditional way of advertising and marketing is an extremely expensive way of advertising for companies. In this sense, the emergence of internet marketing is considered exceptional. Today, businesses, shops, companies, small entrepreneurs, but also multinational companies and corporations are switching to the online variant, i.e. following the logic that everything can be obtained and found out via the Internet! The arrival of digital media has fundamentally changed the way businesses communicate with customers, the rules of the marketing game have changed forever, and it all just started with the introduction of the Internet in 1989, and since then, the number of different types of digital media channels has increased every year [2]. However, for the realization of quality marketing on the Internet, trained human resources and adequate IT infrastructure are necessary. In this regard, Kosovo is lucky to have good conditions and the fast internet that it offers to its citizens, which implies the daily use of the Internet and therefore, social networks. To succeed in the digital economy, businesses must be digitized and digital at the same time, but here the first important differences in the meaning of the words appear [11]. In this sense, Ross et al. (2017) add that digitization implies standardized business processes and mainly refers to lower costs and operational efficiency. Of course, digitization is important for enabling and appropriate use. That is, digital technology can also support operational efficiency, but it cannot automatically turn all these companies into digital organizations [11].

2. Marketing and social media

In the relevant literature, it is pointed out that today's social media have changed the way buyers and sellers communicate, and have enabled users to co-create the value and impact of a marketing strategy, which allows them to interact and exchange information with customers and non-customers and the ability to influence each other and their decisions [6]. What is extremely important is that, unlike traditional marketing, in digital marketing the exact number of views, inquiries, users, customers is known, which is of crucial importance, and here we work on addressing a potential customer or user, where we always know what needs to be corrected or repaired in order to increase the impact on consumers.

It is characteristic that the world of social media is extremely dynamic and constantly changing, and that social networks have become the most widespread media, because they provide direct contact with the audience. Therefore, a new marketing practice has emerged with guidelines for the use of digital media that includes customers and is the only way to achieve marketing goals [12]. Interaction, which is intertwined with individualization, enables the provider to choose information that is significant for him, and advertising establishes two-way communication with the consumer [13].

Some authors emphasize that in Internet marketing, a method of communication is increasingly being applied which is no longer under the control of the company, but the customers take over the power [5]. However, it should be borne in mind that advertising on the Internet can be easy, only if we use the good and weak sides of that medium correctly. Internet advertising allows reaching more people on a global scale, better targeting, lower costs, easier comparison of results and adoption of changes [17]. These are activities that are constantly developing and changing, since the Internet has existed and since the number

of Internet users has been growing. For the development of the aforementioned activities, among other things, and the promotion of Internet marketing, there are significant supports, one of which is the digital Europe program and cooperation with the countries of the Western Balkans, which will enable legal entities of Serbia to participate in competitions for the award of grants in the areas of high-performance computing, artificial intelligence, advanced digital skills, development and the best application of digital capacities and interoperability. It is clear that in today's time when everyone is using the internet, and digital marketing strategies are progressing exponentially. In this sense, the authors Hanson and Kaliani (2007) emphasize that marketers from all over the world have started using a new tool, with which they are able to create marketing material that had, and still has, global coverage at very low costs [7]. This is the reason why, in the world, small and medium-sized businesses are increasingly competing with larger ones. The fact that business systems and customers are connected in communication and learn from each other has become much more interesting. Today, if someone is not physically present, the conversation will happen even without them, or if they decide to join, companies have to directly communicate and respond to customers one-on-one [14].

3. Internet marketing

According to Chaffey (2009), the term internet marketing can be defined as: the realization of marketing goals through the use of digital technology [3]. In addition to the above definition, Digital Marketing or Internet marketing can be defined as any activity of an organization whose goal is to attract the attention of a specific target group through certain communication on the Internet in order to make them interested in its product or service. Also, the definition of digital marketing is the sale and promotion of products and services through online strategies. Strauss and Frost (2012, p. 28) defined e-marketing as the use of information technologies in marketing activities and processes that communicate, deliver and exchange offers, useful for customers, clients, associates and wider society [15]. In Figure 1, we can see the characteristics of six main groups of online marketing channels[2].

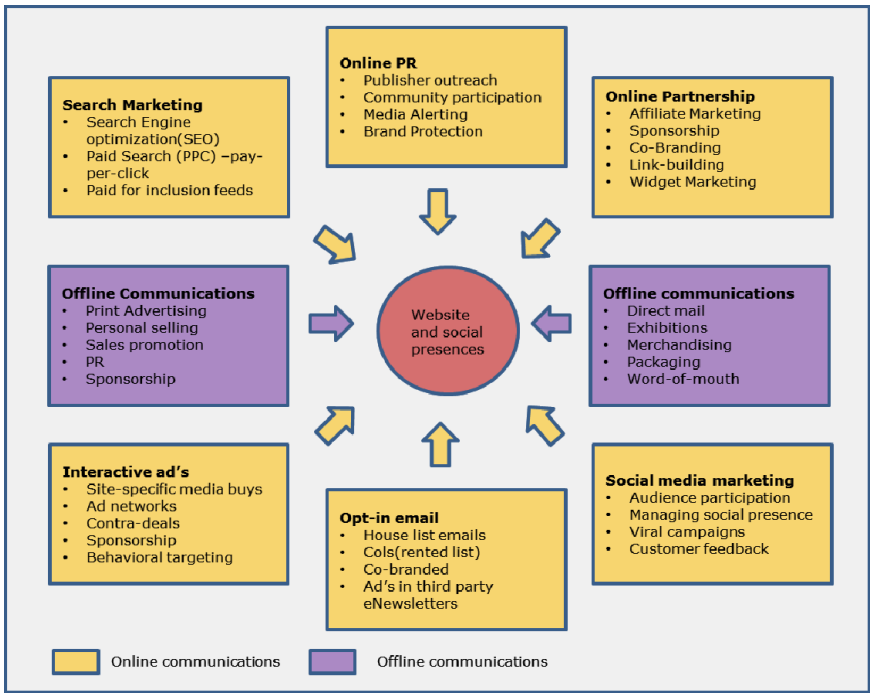


Figure 1. The main groups of online marketing channels
Source: D. Chaffey & F. Ellis-Chadwick, *Digital Marketing: Strategy, Implementation and Practice*, 2012.

A simple definition of e-marketing is that it is the result of information technology that is linked to traditional media. Digital marketing, according to Chaffey (2009), has a similar meaning to e-marketing. Both describe the management and execution of marketing using digital media such as the web, e-mail, interactive television, and wireless media linked to digital data containing consumer characteristics and behavior [3].

Well-known authors point out that it is very important for online marketing companies to understand how this method is built and works, because research shows that the recommendations of friends, family or even other web users are an important factor when choosing products and suppliers [2]. The application of Internet marketing in the function of retail in most industrially developed countries has reached an enviable level. It is important to note that the Internet has a great impact on the citizens of many countries (Strauss & Frost, 2012). Some experts predict higher bandwidth, faster connection speeds, artificial intelligence, an integrated social network or modular web applications that reduce the need for computer software. Forrester Research believes that interactive media will overtake traditional media (Strauss & Frost, 2012), and that this realization will open the door to many new opportunities for products that offer added value by differentiating customers in the future [15]. In this sense, Chaffey and Smith, state that, in general, e-marketing can be used to support the following goals: recording - the Internet can be used for market research to determine the existing needs and desires of consumers; expectations - the Internet provides additional channels through which users can shop and access data; satisfaction - the key success factor is to achieve customer satisfaction (easy to use, proper application of customer service standards, product delivery). [3]. Digital marketing is the use of online and digital channels, such as social media, email, search engines and websites, to promote a brand, product or service and interact with potential and existing customers[19]. Since today's customer is searching for brands online and learning about products from other customers on social media, companies want to join them and influence negotiations - or initiate them, when introducing a new product [15]. Some authors emphasize that choosing the right way of advertising is a difficult task for advertisers because they have to choose between many different forms of ads and with a noticeable lack of knowledge about their effect on consumers [1]. There are several types of digital marketing in the literature, including social media marketing, search engine optimization (SEO), email marketing, content marketing, pay-per-click (PPC) advertising, video marketing, influencer marketing, and affiliate marketing.

Each type of digital marketing has its advantages, strategies and tactics, and companies can choose and combine them based on their marketing goals and target audience. To support marketing, the Internet and other digital media are used, which according to Chaffey et al. (2009), defined as communications [2].

Table 1. Summary of different digital media channels

media channel	Different communications techniques
Marketing through web browsers	• The Search Engine Optimization (SEO) list involves achieving the highest possible position of companies and their products on the life list, which is paid in accordance with the optimization of pages and established links. • Pay-per-click (advertising) sponsored listings, for example, using Google AdWords • Products or a list of keywords can be embedded in search results (Yahoo! Search Submit)

Online public relations	• Summarized content (eg press releases), getting positive references, reputation management on third-party sites, especially on forums and social networks • Using blogs and online resources
Web partnerships, including affiliate marketing	• Fee-based affiliate marketing • Creating long-term partnerships, such as sponsorships, tie-ups or publishing
Interactive display of advertising	• Focus on media leasing • Use of advertising networks • Behavioral advertising
Opt-ine mail marketing	• Business email acquisition, including co-branded campaign list rental, e-newsletter advertising • Maintaining activity and growth, such as domestic lists for e-newsletters and e-mail campaigns for customers • Automatically launch campaign activities from e-mail
Viral and electronic word of mouth marketing	• Creation of "viral agents" or attractive interactive content • Promotion through viral messages • Use of client representation influence • Marketing through online tools (eng. Widgetmarketing)

Source: D. Chaffey, F. Ellis-Chadwick, F. Mayer & K. Johnston, Internet marketing: strategy, implementation and practice, 2009, str. 503-504.

In Internet marketing, care must be taken about how to expose the consumer to an ad, which includes a certain part of the content that the consumer searched for online or saw on a previously visited e-commerce site [18], Figure 2 shows the rating and importance of different organizational goals that content marketing supports, based on studies of marketing content and trends (Content marketing institute, 2014). As can be seen here, the brand gave the most important results with 84%.

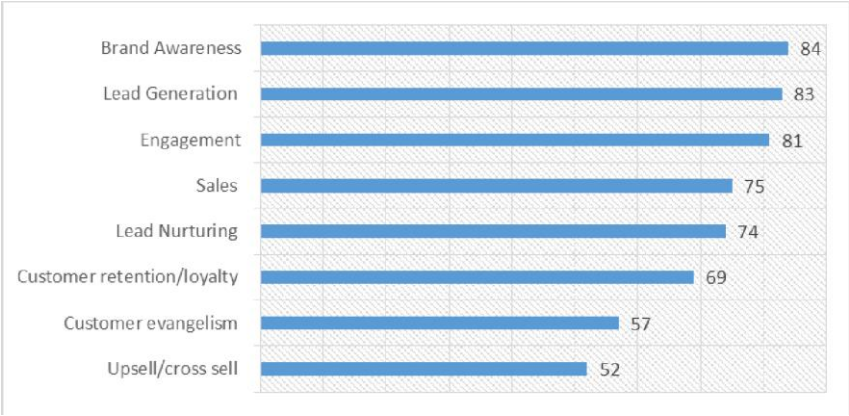


Figure 2. Content of marketing goals in %
Source: J. Pulizzi, *New B2B Content Marketing Research: Focus on Documenting Your Strategy*, 2014.

Based on data obtained from secondary sources and empirical studies carried out in practice, it can be concluded that social networks such as Facebook can contribute to a faster and better visibility of a small business, despite the fierce competition in the market, because it is such a large network of people to give them different information that quickly circulates through the network where companies have a business opportunity to present themselves to potential customers. In addition, they do not require large financial investments, opening an account on the social network Facebook is completely free[20]. Facebook also enables a personal relationship between customers and businesses in the form of a community, where they exchange opinions and answers with each other, which increases customer trust in the business and helps the company discover customer preferences and check trends. In this way, the company attracts new customers to visit the site and therefore decide to buy with different marketing methods. The following figure 3 shows the percentage of advertisement visibility in the analyzed company. As can be seen, the vast majority of 85% of the total number of 232 respondents confirmed the visibility of ads on the social network Facebook.

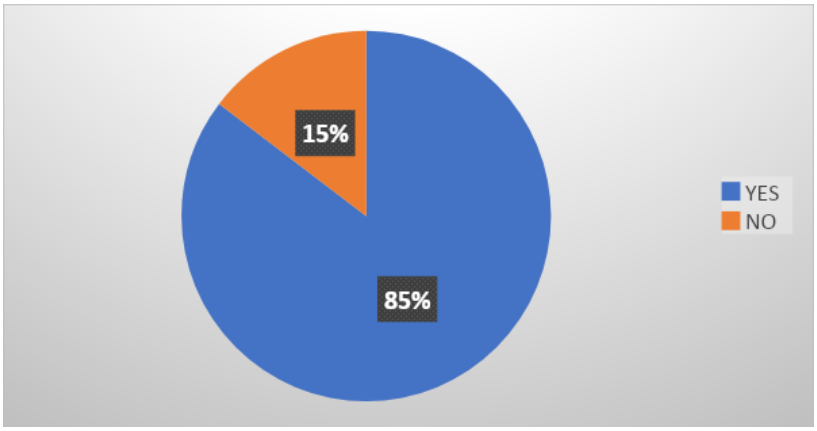


Figure 3: Visibility of ads on Facebook in % (n = 232)

4. Conclusions

Today, the world of marketing has changed drastically, and it is now a new era for marketers and companies that they must accept, adopt and thus tackle new challenges and opportunities. Existing research and research results in the world in the analysis of users' attitudes towards internet marketing differ in general. In this sense, there are research studies that have shown that users have a positive attitude towards Internet advertising, while others show negative attitudes. As one of the solutions to the issue of marketing products/services, digital marketing represents the answer. However, it requires qualified human resources as well as effort from companies both in terms of investments and in terms of the competencies they must build to respect the complexity that digital marketing brings. Due to the digital world we live in today, traditional marketing communications can only have a limited impact in creating awareness and generation with limited investment costs. On the contrary they will have very high investment costs. The research results of this paper show that video ads are the most popular among users in the area.

The goal of the work was to examine the development of the Internet and social networks as an effective tool for online marketing in selected companies and, consequently, greater visibility, by engaging human resources in the application of new marketing methods as a way to attract potential clients. The empirical part of the paper includes the application of interviews and online questionnaires, with the aim of researching the possibility of improving marketing, especially expanding visibility using social networks by engaging human resources in the company. The research results indicate that the most useful social networks for online advertising are Facebook and YouTube. In addition, even using search engine marketing, pay per click, e-mail marketing with consent, viral marketing, etc., sponsorships with famous people in Serbia and personal sales. They are aware that there are many other tools that can be used, but in

the future they want to entertain and educate customers and provide them with more content. The key outcome of this research project is the insight that the customer has the will to buy, and in this regard, the application of digital marketing is recommended. Although the use of these sophisticated technologies and tools for advertising is difficult for the owners of small and medium-sized businesses in the area, largely due to technical reasons, it is still necessary to include the principles and methods of internet marketing in order to solve the problem of product placement and sales.

When it comes to the opinion of respondents about digital marketing and digital communication channels, the respondents emphasized the efficiency and speed of communication, accessibility in the sense that digital advertising does not cause high costs, availability of information, greater attendance by the younger population, etc.

References:

[1] Burns, K. S., & Lutz, R. J. (2006). The Function of Format: Consumer Responses to Six On-line Advertising Formats. *Journal of Advertising, 35*(1),53–63.

[2] Chaffey, D., & Ellis-Chadwick, F. (2012). *Digital Marketing: Strategy, Implementation and Practice* (5th ed.). Harlow: Pearson Education Limited.

[3] Chaffey, D., Ellis-Chadwick, F., Mayer, R., & Johnston, K. (2009). *Internet marketing: strategy, implementation and practice* (4th ed.). Harlow: FT Prentice Hall.

[4] Dehghani, M. and Tumer, M. (2015) A Research on Effectiveness of Facebook Advertising on Enhancing Purchase Intention of Consumers. *Computers in Human Behaior*, *49*, 597-600. https://doi.org/10.1016/j.chb.2015.03.051

[5] Edelman, D., & Salsberg, B. (2010). Beyond paid media: Marketing's new vocabulary. *McKinsey Quarterly*. Retrieved November 24, 2012, from http://www.mckinsey.com/insights/marketing_sales/beyond_paid_media_marketings_new_vocabulary

[6] Gillin, P., & Schwartzman, E. (2011). *Social marketing to the business customer.*Hoboken, New Jersey: John Wiley & Sons.

[7] Hanson, W. A., & Kalyanam, K. (2007). *Internet Marketing and e-Commerce*. Mason: Thomson South-Western.

[8] Hutt, M.D. & Speh, T.W. (2013). Business marketing management B2B (11th ed), Mason: Thomson South-Western, Cengage Learning

[9] J. Pulizzi, *New B2B Content Marketing Research: Focus on Documenting Your Strategy*, 2014.

[10]. Jung, A. R. (2017). The influence of perceived ad relevance on social media advertising: An empirical examination of a mediating role of privacy concern. *Computers in Human Behavior*, *70*, 303-309.

[11] Ross, W.J., Beath M.C. & Sebastian, M.I. (2017). Digitized ≠ Digital. MIT Sloan Center for *Information Systems Reasearch*, *17*(10), 1-4.

[12] Ryan, D., & Jones, C. (2009). *Understanding digital marketing: marketing strategies for engaging the digital generation.* London: Kogan Page.

[13] Shimp, T. A. (2003). *Advertising, Promotion, and Supplemental Aspects of Integrated Marketing Communications* (6th ed.). Ohio: Thomson Learning

[14] Solis, B. (2010). *Engage.* Hoboken, New Jersey: John Wiley & Sons.

[15] Strauss. J. and Frost, R. (2001). E-Marketing. NJ, USA: Prentice Hall.

[16] Sutherland, M. and Sylvester, A.K. (2000) Advertising and the Mind of the Consumer, Alien and Unwin, NSW, Australia

[17] Tangelo. (2021). *Digital marketing strategy*. Tangelo Creative. https://tangelocreative.com.au/wp-content/uploads/2015/07/Tangelo-Digital-Marketing-Strategy.pdf

[18] Taylor, G. (2013). *Advertising in a Digital Age: Best Practices & Tips for Paid Search and Social Media Advertising*. B.k.: Global &Digital.

[19] Veleva, S. S., & Tsvetanova, A. I. (2020). Characteristics of the digital marketing advantages and disadvantages. *IOP Conference Series: Materials Science and Engineering*, *940*, 012065. https://doi.org/10.1088/1757-899X/940/1/012065

[20] Zarouali, B., Ponnet, K., Walrave, M., & Poels, K. (2017). Do you like cookies? Adolescents' skeptical processing of retargeted Facebook-ads and the moderating role of privacy concern and a textual debriefing. *Computers in Human Behavior*, *69*, 157–165. https://doi.org/10.1016/j.chb.2016.11.050

A Model of Problems Related to Scrum Team Communication in Meetings and their Impact on Job Satisfaction

[1]Maja Gaborov,[1]Željko Stojanov, [2]Srđan Popov

[1]*University of Novi Sad, Technical faculty "Mihajlo Pupin", Djure Djakovica bb,Zrenjanin, Serbia*
[2]*University of Novi Sad, Faculty of Technical Science, Trg Dositeja Obradovica 6, Novi Sad, Serbia*

maja.gaborov@tfzr.rs, zeljko.stojanov@uns.ac.rs, srdjanpopov@uns.ac.rs

Abstract: Advances in business have led to the emergence of agile approaches, primarily the Scrum framework for project management. Since Scrum is focused on people, great care should be taken to determine what problems can arise during communication between people, in this case between team members. The paper shows the problems that arise during communication between team members in connection with meetings, where the emphasis is on the fact that the satisfaction of employees is of great importance because the success of the project depends on them. The created model shows the problems that were found from literature sources from scientific databases. The results suggest that it is desirable to detect problems in time and prevent them or solve them if they occur because they greatly affect the team members on whom the success of projects depends.

Keywords: model, problems, communication, scrum, team meetings, job satisfaction

1. Introduction

In modern business there are new methods of project management where newer approaches are applied. Modern business conditions have led to the development of the agile approach, the Scrum framework and other methods, techniques and tools for project management [1]. Highsmith [2] emphasizes that agility is "the ability to both create and respond to changes with the aim of profit in a turbulent business environment" [2]. Agile approach implies a newer way of managing business and making decisions [3]. Agile approach implies a combination of elements of interactive and incremental development of project products [1, 4].Agile principles should be followed in order to establish an agile atmosphere during the project. There are a number of agile principles that emphasize the importance of direct communication [1, 5]. Schwaber and Sutherland [6] define Scrum as "a framework in which people can tackle complex problems while continuing to deliver products of the highest possible value productively and creatively". Scrum implies the use of various techniques and tools [1].In Scrum, a big focus is on people. If there is misunderstanding and conflict among team members, it can lead to poor performance, hurt feelings, and lack of motivation to continue working [7].

Sometimes it is difficult to provide the right amount of information to all the people involved. In software engineering, the impact of "proper" communication on project success is often debated. In software development teams, inaccurate or insufficiently communicated requirements often lead to unimplemented pieces of software. Insufficient communication can jeopardize the success of the project. Adequate communication can be crucial to the success of a project. Stapel [8] claims that the chosen communication channel is important for conveying the desired information. Cockburn [9] evaluated different types of communication channels according to their effectiveness and richness [9, 10, 11].

A prerequisite for the success of the project is the satisfaction of the employees. Whitman et al. [12] defined "unit-level job satisfaction" as "a common internal work unit a state that is expressed by the affective and cognitive evaluation of shared work experiences with others degree of favor or disfavour".

They emphasized the importance of togetherness as a critical prerequisite for formation of collective job satisfaction [13].

Based on all the above, the aim of this paper is to show communication problems related to meetings in Scrum teams where there is a high concentration on the satisfaction of team members. The identified problems are distributed in the developed model.

2. Scrum framework

Scrum is a framework in which people address problems by productively and creatively delivering quality products[14]. The primary function of the Scrum framework is to organize people and their relationships. After people are organized, the Scrum framework gives them the tools and procedures they use to manage their work [14].There are three main components of the Scrum framework: Scrum roles, Scrum artifacts and Scrum events [14] where we will focus on Scrum roles and events. These three components define the way of organizing work on solving complex problems. Roles define responsibility, clearly showing who is responsible for different types of decision-making. Artifacts provide transparency, providing information that supports the decision-making process. Events provide forums within which certain decisions are made [14, 15].

2.1. Scrum roles

Given that the big focus in Scrum is on people, it is necessary to organize people into teams that will work on the project. The total number of team members should not exceed ten, it is a sufficient number to maintain independence in the choice of tools, techniques and tasks as well as sharing knowledge and skills [6]. Scrum teams foster productivity, creativity and flexibility [1, 3, 6].The Product Owner is responsible for the quality of the project products made by the development tea. [1]. The product owner is responsible for defining project constraints such as time, costs and resources, but also for monitoring financial indicators and controlling budget execution [1, 14]. A scrum developer is a technical professional who participates in the team's product creation. Developers are organized into teams and collectively take responsibility for results to the customer [1]. Scrum development teams must have all the different skills needed to deliver the product. Members of the Scrum development team are equal. Shared accountability is one of the keys to the effectiveness of the Scrum framework. As problems arise, the development team must solve them themselves [41]. Therefore, the development team is a group of professionals who possess the knowledge necessary to create project products and whose work directly creates value in the project [1]. It is recommended that the Development Team has six to ten members in order to maintain effective communication and coordination of work [1]. The role of the Scrum Master in the Scrum Team is defined in terms of responsibilities as opposed to tasks. Scrum masters must ensure that the development team owns every part of the product feature delivery that is prioritized by the product owner [14]. Assists team members by organizing sprint refinement and sprint retrospective meetings, is responsible for avoiding barriers during the process and provides necessary resources for the team. It has a supporting role towards the product owner, the development team and the business [16]. Finally, he is also responsible for not adding any additional items during the sprint [15, 17].

2.2. Scrum events

The Scrum framework involves establishing Scrum events in the project to reduce time wastage. Also, it is planned to avoid organizing meetings that are not productive enough [1]. Scrum events have a limited duration that determines the dynamics of work [1]. Project products are developed in equal time iterations called Sprints. A sprint is defined as a time frame during which usable versions of project products are created. Essentially, during a Sprint, team members work on product development and building product

performance, thereby meeting stakeholder requirements [18].If certain tasks are not completed during a Sprint, they can be rescheduled for the next Sprint. At the end of the last Sprint and after delivering the final product to the customers, the project can be considered complete. The recommended sprint duration is from one to four weeks [1, 6].Each of the five Scrum events is an opportunity to check transparency and adjust. The five events are as follows [14, 15]:

1. *The Sprint*: During this event, the product is inspected and the Sprint Increment is created and adapted.
2. *Sprint Planning Meeting*: During this event, the product backlog is reviewed and the sprint backlog is created and adjusted. [14] All members of the Scrum team participate together in planning the implementation of the next Sprint. During the Sprint Planning event, the Product Backlog is refined and tasks are planned for the next period, which directly affects the time, costs and performance of project products [1].
3. *The Daily Scrum*: Daily Scrum is a fifteen-minute meeting of Development Team members [16] during which the implementation of tasks for the current day is planned. The responsibility for running the daily Scrum meeting rests with the Scrum Master, who deals with the time constraints and productivity of planning the daily tasks. A daily Scrum event involves asking three key questions: what was done during the previous day, what can be accomplished in the next meeting, and what are the obstacles the team is facing. By answering these questions, the work of the previous day is reviewed and compared with the plan for that day, then the tasks and activities whose implementation is planned for that day are defined, and finally the limitations and problems that slow down the execution. After the daily Scrum meeting, members of the Development Team can independently organize other meetings according to their needs [1].
4. *Sprint Review*: During this event, the new product version is reviewed and the product backlog is adjusted. [14] The Sprint Review is always held at the end of the Sprint and is limited to one hour for each week of the Sprint [19]. The presence of all project participants is mandatory because during the Sprint Review event, the tasks performed during the Sprint are reviewed, the degree of achievement of the Sprint Goals is assessed, and new requirements and changes are defined that are entered into the Product Backlog. The end result of the Sprint Review event is a revised Product Backlog that forms the basis for planning the next sprint [1].
5. *Sprint Retrospective*: A sprint retrospective is an event during which team members review their work, discuss previous sprints, and define measures for better work in the next sprint. The recommended duration of the Sprint Retrospective is 45 minutes [20] for each week of the Sprint. The Scrum Master has the greatest responsibility for holding the Sprint Retrospective event where he acts on an equal footing with the development team members with whom he discusses how to improve the Scrum framework application and other project management methods, techniques and tools [1, 15, 21].

3. Problems in team communication in meetings

The authors of Stray et al. [22], pointed out that sometimes the Scrum Master criticized the team member for not paying attention in the meeting which can lead to various problems. Quality assurance (QA) engineers, documentation writers and support staff did not receive enough attention in meetings compared to other team members. Not having everyone's attention in a meeting can lead to unbalanced contributions from team members in meetings.

In a more recent study by Stray et al [23], respondents stated that they did not have good experiences with leading a meeting, so there was a decrease in job satisfaction and trust among colleagues. The information exchanged was not considered relevant, especially due to the variety of roles, tasks and seniority. It happened that there were longer meetings if there were more people and therefore each member had less time to do their work tasks, which slowed down the work on the project. It was also

possible that the meetings were late, so the members spent more time waiting and participating in the meeting instead of doing some other activities.

In one of the studies also conducted by Stray et al [24], where software teams in companies in Malaysia, Norway, Poland and the United Kingdom were surveyed, information was shared with the team so that members did not have a bad opinion of the meeting. This leads to avoiding potential problems and dissatisfaction among members. An interesting fact is that one of the teams used video in many distributed daily standup meetings, which had a positive effect on participation and communication because people paid more attention than when they were on the phone. Team members appreciated seeing each other's faces. In several teams, a lot of informal communication reduced the necessity of a meeting. Much of the communication resulted in identifying problems that were solved or simply avoided because they communicated through other tools if they were not in the meeting in the same room.

This is followed by a more recent study by the mentioned authors [25] where in large teams not all the information shared during the meeting is relevant for individuals. It was also emphasized that the larger the team, the more dissatisfied the members are with the meeting because developers in larger teams see the meeting as less valuable than developers in smaller teams.

In an earlier study by Stray et al.[26], it was observed that in many meetings little time was spent on task coordination. There is not much communication about potential problems regarding project tasks so team members may feel that the meeting does not make sense to them. People were often late for meetings, so meetings rarely started on time, which only reduced the time for members to work on their other activities.

A study by Lalsing et al. [27], emphasizes that trust and communication are very important items and that they greatly affect software teams. When the project teams worked in a common workspace, it facilitated communication between them and they did not always have to hold formal meetings, but solved problems whenever a problem arose. When two teams were located in different geographical locations it was more difficult to communicate, so some teams communicated when they had an obligation through a meeting. Some of the project teams had better communication because they often went out for lunch and communicated verbally and this improved their familiarity. Team members sometimes forgot to convey information to other team members or felt that the information would not be relevant to others. Meetings sometimes ended with only a few people talking and the rest mostly listening.

Similar to the previous study, the study by Stadler et al. [28] also emphasizes the importance of building trust and interpersonal relationships between team members. In that study, it was observed that Daily Meetings were held with audio or video conferences. Therefore, some companies believe that it is possible to achieve good communication when not all team members are physically present at the meeting in one room. When it came to retrospective meetings, all team members were usually in one physical location.

The study by Guillot et al.[29] also emphasizes the importance of constant communication, where it is emphasized the importance of team members being motivated because this will affect success. Amorim et al.[30] point out that effective communication and motivation of members greatly influence software engineers, primarily their satisfaction.

This is followed by the study of Cruzes et al.[31], where asking questions and getting quick feedback was challenging due to the time difference and the unavailability of developers as needed. This is a serious problem that can arise in the meeting because it happens that not everyone is in the meeting so that some information does not reach the person who needs the information.

3.1. Development of model based on team communication problems in meetings

In this paper, the model was created based on a literature review in well-known scientific libraries. Problems related to communication are presented in the model in figure 1, where it is observed that these

problems affect the satisfaction of employees. All the individual problems related to communication are explained in chapter 3.

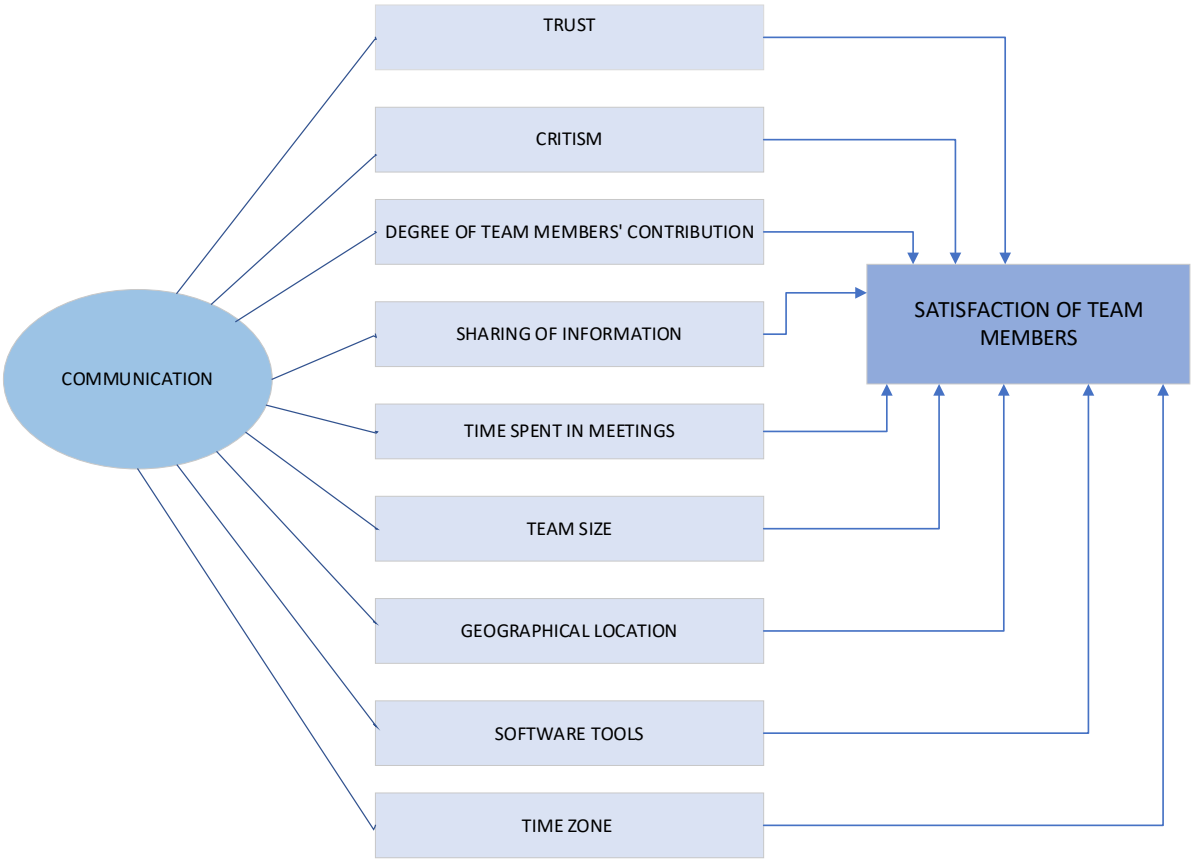


Figure 1. A model of team communication – related problems and relationship with satisfaction of team members

The model shows problems that are related to communication, and which have to do with the satisfaction of team members. These issues are: trust, criticism, degree of team members' contribution, sharing of information, time spent in meetings, team size, geographical location, software tools, and time zone.

4. Discussion of model based on team communication problems in meetings

Based on the review of the literature regarding the problems that arise during communication at meetings, a model was created and described with the relationships between the mentioned problems. It can be seen from the model that quality communication is important for everyone and that various problems can arise that can affect the satisfaction of team members. We believe that it is really important how the team members and employees feel in general because the finished product that is delivered to the customer depends on their work, on which the company's profit depends. One of the potential problems that can arise during communication is a lack of trust. During communication, some people have conflicts, or they haven't gotten to know each other enough, so they don't convey all the relevant information they need. Many things can affect the lack of trust between team members, so if team members are not in agreement

and do not have good communication, there can be great dissatisfaction and interruptions in work, and therefore the failure of the project, which is bad for the company. It is desirable that the members do not criticize each other but have a good relationship because it contributes to better work. A person who is criticized may not try to solve tasks because he will not be motivated to work. It may happen that the leader of the meeting does not allow everyone to express their opinion and thus contribute to the meeting to the right extent, so it may happen that the information is not conveyed to all members. Meetings can be long, so they can disrupt individuals' plans and increase meeting time. Meetings often last longer if there are larger teams or, for example, the teams are far away, so it is more difficult to contact these people because it often happens that not everyone is available at the right time due to the time zone. It may also happen that some team members are not happy with the communication software tool.

4.1. Research implications

This section will discuss the implications of the research for industry practitioners and academic researchers. This model can be used by industrial practitioners. It can help practitioners understand that communication between team members is also crucial for the success of the company, because it is very important that the meeting is carried out well, that all members are satisfied, that they have mutual trust, that information is shared between them, and so on. This model can be helpful to practitioners so that when they look at the model, they can see what could be automated in the meeting process and thus facilitate communication between team members which would increase employee satisfaction as certain problems would be solved. With the help of this model, researchers could see some of the problems during communication between team members and carry out an even more detailed analysis of the problem, because communication is a really important factor for the satisfaction of employees and the entire company. If the employees are not satisfied, it can affect the entire company and its profits, so communication and problems that may arise during meetings are not a naive thing, but worthy of attention for further work.

4.2. Validity of the research

This paper provides implications for the importance of communication, but the authors are aware that there are some limitations that affect the validity of the research. It is possible that we missed some case studies that were published in digital libraries that we did not search. We have searched the most influential libraries where papers are published after a rigorous review. Extracting data from the available papers was difficult because many studies did not explicitly mention and explain each of the problems we observed, requiring interpretation of the data, which includes personal bias. Some of the problems appeared in several works, which leaves the impression that there is a great emphasis on certain problems. All the authors of this paper participated in the discussion and development of the findings and finally agreed on the defined problems in the developed model. The applicability of this model in projects that apply other methodologies is questionable. In the IT industry, the most applicable Scrum methodology is with possible modifications, which increases the generalization of the presented model.

5. Conclusion

A Scrum team focuses on solving problems where respect and trust must first be built. The authors therefore believe that communication is a very important factor in the success of any company, so that the team members are satisfied, because the success of the project depends on their work and job satisfaction. As meetings are often held to communicate about specific topics and problems, the authors claim that it is therefore necessary to present specific problems at meetings. The presented model emphasizes the importance of communication, because with the help of good communication, potential problems can be prevented and thus reduce employee dissatisfaction. In the future, it could be investigated what all the team members are dissatisfied with because people are the key and they perform all the work tasks that should lead to the success of the project. Therefore, it is desirable to pay special attention to potential problems that may arise. Also, the automation of meetings is planned so that employees do not waste a lot of time, and have more time to perform their tasks, where communication would be reduced only to the transfer of essential information between team members.

References

[1] Grebić, B., & Stojanović, A. (2021). Application of the Scrum framework on projects in IT sector. Eur. Proj. Manag. J, 11(2), 37-46.
[2] Highsmith, J. (2009). Agile Project Management. Boston: Addison-Wesley.
[3] Grebić, B. (2019). Traditional vs Agile Project Management in the Service Sector. European Project Management Journal 9(2), pp. 55-63. https://doi.org/10.18485/epmj.2019.9.2.6
[4] Project Management Institute. (2017). Agile Practice Guide. Pennsylvania: Author.
[5] Agile Alliance. (2001). The Agile Manifesto. USA: Author.
[6] Schwaber, K., & Sutherland, J. (2017). The Scrum Guide. USA: Creative Commons.
[7] Brewer, E. C., & Holmes, T. L. (2016). Better communication= better teams: A communication exercise to improve team performance. IEEE Transactions on Professional Communication, 59(3), 288-298.
[8] Stapel, K.: Informationsflusstheorie der Softwareentwicklung. Dissertation, Gottfried Wilhelm Leibniz Universit¨at Hannover, M¨unchen (Apr 2012)
[9] Cockburn, A.: Agile software development: The cooperative game. Pearson Education (2006)
[10] Ambler, S.W., et al.: Agile modeling (2002)
[11] Klünder, J., Schneider, K., Kortum, F., Straube, J., Handke, L., & Kauffeld, S. (2016). Communication in teams-an expression of social conflicts. In Human-Centered and Error-Resilient Systems Development: IFIP WG 13.2/13.5 Joint Working Conference, 6th International Conference on Human-Centered Software Engineering, HCSE 2016, and 8th International Conference on Human Error, Safety, and System Development, HESSD 2016, Stockholm, Sweden, August 29-31, 2016, Proceedings 8 (pp. 111-129). Springer International Publishing.
[12] Whitman, D.S.; Van Rooy, D.L.; Viswesvaran, C. Satisfaction, citizenship behaviors, and performance in work units: A meta-analysis of collective construct relations. Pers. Psychol. 2010, 63, 41–81. [CrossRef]
[13] García-Buades, M. E., Peiró, J. M., Montañez-Juan, M. I., Kozusznik, M. W., & Ortiz-Bonnín, S. (2020). Happy-productive teams and work units: A systematic review of the 'happy-productive worker thesis'. International journal of environmental research and public health, 17(1), 69.

[14] Fowler, F. M., & Fowler, F. M. (2019). Scrum artifacts. Navigating Hybrid Scrum Environments: Understanding the Essentials, Avoiding the Pitfalls, 55-57.
[15] Baijens, J., Helms, R., & Iren, D. (2020, June). Applying scrum in data science projects. In 2020 IEEE 22nd conference on business informatics (CBI) (Vol. 1, pp. 30-38). IEEE.
[16] J. S. Saltz and A. Sutherland, "SKI : An Agile Framework for Data Science," in 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 3468–3476.
[17] J. S. Saltz and A. Sutherland, "SKI : A New Agile Framework that supports DevOps , Continuous Delivery , and Lean Hypothesis Testing," in Proceedings of the 53rd Hawaii International Conference on System Sciences., 2020.
[18] G. S. do Nascimento and A. A. de Oliveira, "An Agile Knowledge Discovery in Databases Software Process," in The Second International Conference on Advances in Information Mining and Management compliance, 2012, pp. 343–351.
[19] L. Williams, "Agile Software Development Methodologies and Practices," Adv. Comput., vol. 80, pp. 1–44, 2010
[20] N. W. Grady, J. A. Payne, and H. Parker, "Agile big data analytics: AnalyticsOps for data science," in Proceedings 2017 IEEE International Conference on Big Data, 2017, pp. 2331–2339.
[21] C. Dremel, M. M. Herterich, J. Wulf, J.-C. Waizmann, and W. Brenner, "How Audi AG established big data analytics in its digital transformation," MIS Q. Exec., vol. 16, no. 2, pp. 81– 100, 2017
[22] Stray, V. G., Lindsjørn, Y., & Sjøberg, D. I. (2013). Obstacles to efficient daily meetings in agile development projects: A case study. In 2013 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (pp. 95-102). IEEE.
[23] Stray, V., Moe, N. B., & Sjoberg, D. I. (2018). Daily stand-up meetings: start breaking the rules. IEEE Software, 37(3), 70-77
[24] Stray, V., Sjøberg, D. I., & Dybå, T. (2016). The daily stand-up meeting: A grounded theory study. Journal of Systems and Software, 114, 101-124.
[25] Stray, V., Moe, N. B., & Bergersen, G. R. (2017). Are daily stand-up meetings valuable? A survey of developers in software teams. In International Conference on Agile Software Development (pp. 274-281). Springer, Cham.
[26] Stray, V. G., Moe, N. B., & Aurum, A. (2012). Investigating daily team meetings in agile software projects. In 2012 38th Euromicro Conference on Software Engineering and Advanced Applications (pp. 274-281). IEEE
[27] Lalsing, V., Kishnah, S., & Pudaruth, S. (2012). People factors in agile software development and project management. International Journal of Software Engineering & Applications, 3(1), 117.
[28] Stadler, M., Vallon, R., Pazderka, M., & Grechenig, T. (2019). Agile distributed software development in nine central European teams: Challenges, benefits, and recommendations. International Journal of Computer Science & Information Technology (IJCSIT) Vol, 11.
[29] Guillot, I., Paulmani, G., Kumar, V., & Fraser, S. N. (2017, August). Case studies of industry-academia research collaborations for software development with agile. In CYTED-RITOS International Workshop on Groupware (pp. 196-212). Springer, Cham
[30] Amorim, L. F., Marinho, M., & Sampaio, S. (2020). How (un) happiness impacts on software engineers in Agile teams?. arXiv preprint arXiv:2006.03546.
[31] Cruzes, D. S., Moe, N. B., & Dybå, T. (2016, August). Communication between developers and testers in distributed continuous agile testing. In 2016 IEEE 11th International Conference on Global Software Engineering (ICGSE) (pp. 59-68). IEEE.

Navigating the Future: How Information Technologies are Shaping Entrepreneurship Ecosystems in Society 5.0

Dragan Ćoćkalo[1], Mihalj Bakator[2*], Sanja Stanisavljev[3], Melita Ćoćkalo-Hronjec[4] and Dragana Kovač[5]

[1] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*
[2] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*
[3] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*
[4] *High school "Laza Kostic", Laze Lazarevica 1, 21102, Novi Sad, Republic of Serbia5*
[5] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*

dragan.cockalo@tfzr.rs; mihalj.bakator@tfzr.rs; sanja.stanisavljev@tfzr.rs; dragana.kovac@tfzr.rs; melita.cockalo@gmail.com

Abstract:
Society 5.0 represents an ambitious vision for a future where economic growth and social well-being are balanced through the strategic integration of information technologies. Building on previous societal models, this concept leverages Big Data, cloud computing, and Artificial Intelligence (AI) for real-time decision-making, notably in smart cities. A human-centric approach in Society 5.0 aims to personalize services like healthcare and education, while also emphasizing sustainable development and social inclusion. In this modern business environment, employment dynamics change and focus is shifted from traditional manufacturing lines towards entrepreneurship. Thus, entrepreneurship ecosystems have an integral role in the concept of Society 5.0, and information technologies can be viewed as a catalyst that brings technological movement in entrepreneurship. The main goal of this paper is to develop a theoretical model for improving entrepreneurship ecosystems and competitiveness of enterprises. The paper contributes to the existing body of literature and provides a solid basis for future research in this domain.

Keywords:
Society 5.0, entrepreneurship ecosystem, competitiveness, information technology

1. Introduction

The concept of "Society 5.0," alternatively known as an "advanced intelligent society," projects a future where sophisticated technologies like artificial intelligence (AI), the Internet of Things (IoT), and robotics become an integral part of daily life [1]. Sometimes characterized as the "next phase of industrial evolution," Society 5.0 underscores the essential part that entrepreneurial endeavors play in technological progress and economic prosperity. Information technologies act as a link in this new societal structure, enabling real-time data analysis and fostering collaboration across sectors. Governments facilitate this entrepreneurship movement by offering both financial backing and other kinds of assistance to new businesses [2]. Through policy formulation in collaboration with the private sector, they also set the stage for innovation and economic growth. Information technologies are important here, helping to streamline bureaucratic processes and make public services more efficient and responsive to individual needs. Similarly, academic institutions are key players, serving as sources of the expertise and skillsets needed for both technological and business innovation. They contribute to the entrepreneurial ecosystem by launching incubator and accelerator programs, thus helping in the development of startups. Information technologies play a significant role by improving research capabilities and offering platforms for collaboration [3, 4, 5].

Established companies of all sizes contribute to economic growth by financially supporting and mentoring startups. These organizations often collaborate with governmental agencies and educational entities to promote advancements in business methodologies and technology. Through information

technologies, they can quickly adapt to market changes, explore new business models, and even develop new markets and industries. One of the main traits of Society 5.0's entrepreneurial environment is a high degree of inter-organizational collaboration [6]. This is particularly evident in technological spheres where various organizations, ranging from startups and established companies to academic bodies and governmental agencies, collaborate on groundbreaking technological and business initiatives. The role of information technologies is amplified in this context, acting as the conduit through which these diverse entities can collaborate effectively. The current body of literature doesn't address properly the dynamics of Society 5.0, entrepreneurship ecosystems, and information technologies. The goal of this paper is to close this knowledge gap by conducting an analysis of entrepreneurship ecosystems, Society 5.0, and informational technologies.

The body of this paper consists of three major sections, excluding the introduction and conclusion. First, information technologies and the concept of Society 5.0 are discussed. Following that, entrepreneurship ecosystems are addressed and the theoretical model is presented. Finally, suggestions and guidelines for improving entrepreneurship and competitiveness are noted.

2. Information technologies in Society 5.0

In the context of Society 5.0, the role of information technologies is crucial in linking and streamlining diverse societal sectors. The concept of Society 5.0, initially popularized by the Japanese government, aspires to balance economic growth with the resolution of social issues. The concept succeeds earlier societal models, including hunter-gatherer, agrarian, industrial, and information societies. One cornerstone in the context of Society 5.0 is the utilization of Big Data, cloud computing, and Artificial Intelligence (AI) for analytics [7]. These technologies enable the real-time gathering, processing, and analysis of data, which can lead to prompt and precise decision-making. For instance, in the framework of smart cities, these technologies could be harnessed to enhance energy efficiency, mitigate traffic congestion, and elevate public safety standards [8].

Another significant component is the Internet of Things (IoT). IoT plays a vital role in generating a fully interconnected society, incorporating everything from smart homes and connected vehicles to IoT-enabled industrial equipment. In such a landscape, data flows with minimal friction, enhancing the quality and efficiency of life by converting everyday objects into data points that can be scrutinized and acted upon [9]. AI stands as another key element, offering the capacity to process extensive datasets to yield actionable insights. The technology can automate routine tasks, contribute to various fields like healthcare and education, and even serve roles in companionship and personalized learning. Moreover, cybersecurity becomes increasingly essential in a society characterized by seamless data sharing and connectivity. To maintain data integrity and user privacy, advanced encryption algorithms, blockchain technologies, and AI-driven security measures are under development and implementation [10].

Additional interfaces for human interaction with technology come in the form of Augmented Reality (AR) and Virtual Reality (VR). These technologies are not only changing the way information is accessed but also how skills are acquired, offering new avenues for worker training, educational immersion, and social interaction. The aspiration for Society 5.0 is inherently human-centric, focusing on individual well-being. Personalized health monitoring systems, AI-empowered educational platforms, and customized public services aim to cater to the unique needs and preferences of each citizen. Alongside these goals is an emphasis on sustainable growth. Technologies like smart grids and AI-powered resource management systems aim to minimize waste and environmental impact. Finally, an inclusive approach towards social challenges is integrated into the fabric of Society 5.0. Machine learning, natural language processing, and data analytics are leveraged to tackle issues such as social inequality, healthcare disparities, and gaps in education. These technologies can facilitate a nuanced understanding of the root causes of social problems, paving the way for effective solutions.

In summary, information technologies in Society 5.0 seek to establish a system that is human-centric, sustainable, and inclusive. By leveraging the power of connectivity, data, and intelligence, this ambitious vision aims to address complex societal issues while elevating the overall quality of life.

3. Entrepreneurship ecosystems

Entrepreneurial ecosystems include a variety of components that make it easier to start and grow new businesses, such as government-enacted laws and regulations, financing opportunities, and the availability of skilled workers and other resources. These ecosystems are important for increasing economic growth and innovation because of the conditions they create. This is due to the fact that they create the conditions for entrepreneurs to start and grow new businesses [11]. A robust entrepreneurship ecosystem, can result in increased job creation, higher rates of innovation, and overall economic development within a region or country. Access to various forms of financial backing is one of the most important components of an entrepreneurial ecosystem. Capital is essential for entrepreneurs because it allows them to finance the expansion and development of their businesses. Venture capital, angel investments, and government grants and loans are all possible sources of funding [12].

Furthermore, the proliferation of alternative financing methods, such as crowdfunding and peer-to-peer lending, has increased the amount of capital available to entrepreneurs. Finally, a thriving investment community, as well as financial institutions that encourage entrepreneurial endeavors, both contribute to a thriving entrepreneurial ecosystem. Another important component of ecosystems that foster entrepreneurial activity is the availability of both human and physical capital, such as skilled workers and managerial expertise. Human capital includes skilled workers and managerial experience, whereas physical capital includes office space and equipment [13].

Universities and other types of research institutions can also help to increase the availability of talent and resources. These institutions are frequently used as sources of new ideas and cutting-edge technology. Furthermore, business incubators, accelerators, and co-working spaces can provide entrepreneurs with valuable resources, mentoring, and networking opportunities, all of which can help them grow their businesses [14]. The attitudes held within a society and its culture are also important factors that influence entrepreneurship ecosystems. A culture that encourages taking risks and being innovative, for example, can contribute to the creation of an environment that is more conducive to the establishment and expansion of businesses by individuals with entrepreneurial aspirations. This can be aided by highlighting successful entrepreneurs as role models, developing entrepreneurial communities, and launching programs that recognize entrepreneurs' accomplishments [15]. Coverage of entrepreneurial successes in the media and increased public awareness of those successes are also potential factors in the formation of an entrepreneurial culture.

Incorporating cutting-edge technologies can lead to the development of improved communication and collaboration tools, making it easier for entrepreneurs to connect with business associates, clients, and advisors. New technologies have the potential to enable remote work and distributed teams. This gives entrepreneurs access to talent and resources located all over the world, while also encouraging entrepreneurship in less developed regions with limited resources and talent. The increasing accessibility of various online education platforms and resources. Figure 1 depicts the previous section's theoretical model for enhancing entrepreneurial ecosystems and national competitiveness.

Finally, the presence of successful established businesses can strengthen an entrepreneurial ecosystem. New businesses can benefit from established businesses' mentorship, resources, and clientele, which helps to foster an innovative and entrepreneurial culture. Aspiring business owners can benefit from these successful companies as well by seizing opportunities to learn from them and benefit from their sharing of best practices and lessons learned from their experiences. The incorporation of cutting-edge technologies into Society 5.0 has the potential to significantly improve the ecosystems that support entrepreneurial ventures. Based on the analyzed literature and qualitative analysis of previous studies in the domain of entrepreneurship, competitiveness, Society 5.0, and information technologies, a theoretical model is developed. The model aims to concisely present the potential of improving entrepreneurship and competitiveness through information technologies in the context of Society 5.0.

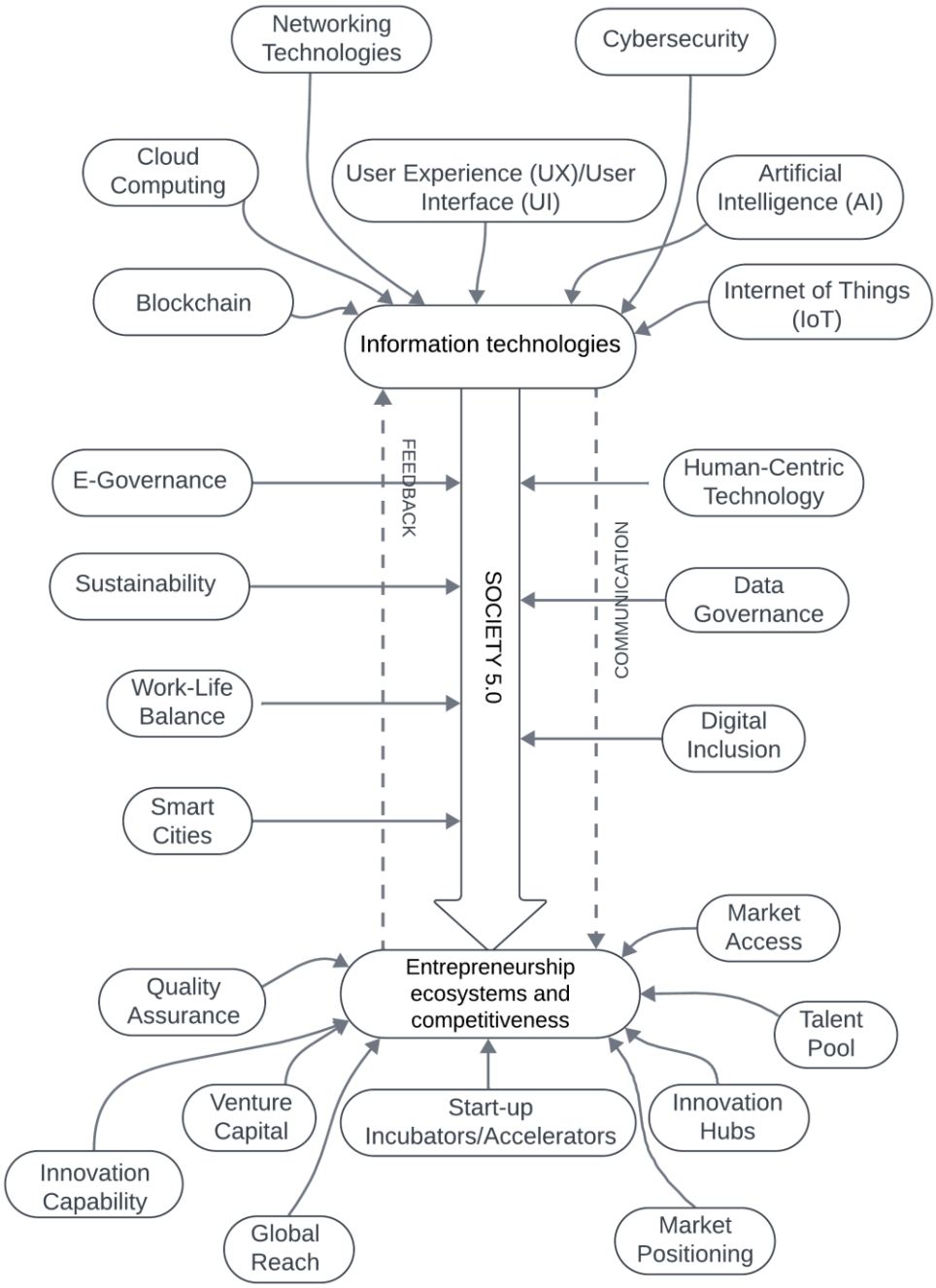


Figure 1: Theoretical model

4. Suggestions and guidelines

Based on the literature reviewed and the model developed, the following recommendations and guidelines for improving entrepreneurship ecosystems and competitiveness of enterprises are noted:

- Governments should prioritize the rollout of high-speed internet access to all corners of the nation. This would reduce the digital divide and provide an essential tool for startups and small businesses to participate in the global economy.

- To facilitate the applications of Internet of Things (IoT), Artificial Intelligence (AI), and Augmented Reality/Virtual Reality (AR/VR), governments should accelerate the deployment of 5G networks. This would create an ecosystem where businesses can leverage real-time data and automation effectively.
- Governments need to enforce robust data protection laws that balance individual privacy with business needs. Strong yet flexible regulations would build public trust in technology while not stifling innovation.
- Adopting and implementing international standards for emerging technologies can simplify the process of integration and scale. This will help domestic companies be globally competitive and facilitate cross-border partnerships.
- A public-private venture capital fund focused on information technology startups can spur innovation. Governments can act as an anchor investor, thereby reducing risks for other institutional investors.
- Tax breaks for investments in IT research and entrepreneurship can be a win-win strategy. Companies would be more willing to invest in long-term research, and governments would benefit from the downstream economic activities.
- Governments can offer direct financial support to startups specializing in Society 5.0 technologies through grants and subsidies. These funds can help these businesses pass the challenging early stages.
- Integrating elements of digital literacy and computational thinking into the school curriculum from an early age can lay the foundation for a digitally-savvy workforce.
- Hosting regular workshops and seminars can provide entrepreneurs and businesses with up-to-date knowledge on emerging technologies and best practices.
- Public support for technology incubators and accelerators can provide startups with the resources, mentoring, and networks they need to grow.
- A multi-stakeholder approach involving policy discussions among governments, academia, and industry can result in more informed and effective policy-making.
- Technology exchange programs with other countries can allow for the sharing of expertise, and help domestic companies understand global markets and standards.
- Companies should invest in internal research and development teams focused on Society 5.0 technologies. Dedicated R&D can lead to proprietary technologies that provide a competitive edge.
- Collaboration isn’t just for big enterprises. Open innovation allows companies to work with startups, educational institutions, and sometimes competitors to accelerate the development of new technologies.
- Organizations must invest in continuous employee training programs that keep the workforce updated on the latest technologies and methodologies.
- Digital platforms should be used to understand global market trends and consumer behaviors. This enables companies to expand beyond local markets with informed strategies.
- Advanced analytics and AI can be used to personalize customer experiences, leading to higher customer satisfaction and loyalty. Employing IoT sensors and analytics software in supply chain operations can lead to real-time insights, better inventory management, and overall operational efficiency. Companies should use smart technologies to minimize waste, optimize energy usage, and reduce their carbon footprint.
- As companies collect more data and rely heavily on digital platforms, robust cybersecurity measures are essential to protect assets and customer information. Regulatory compliance is increasingly complex in the digital age. Automated compliance tools can help businesses adapt to new regulations more efficiently.
- Summits, conferences, and other knowledge-sharing platforms can help both governments and enterprises stay abreast of global trends, share insights, and collaboratively tackle challenges.

By implementing these detailed strategies, both governments and enterprises can create a more supportive environment for entrepreneurship, improve technological capabilities, and enhance competitiveness in the context of Society 5.0.

5. Conclusion

For governments, the importance of infrastructure, particularly in terms of broadband and 5G, cannot be overstated. These foundational elements not only enable advanced applications but also democratize access to digital opportunities. Additionally, public funding mechanisms such as venture capital funds, tax incentives, and grants could act as catalysts that propel startups and SMEs into the limelight of Society 5.0. Education and skill development strategies aimed at future-proofing the workforce, along with the fostering of collaboration across sectors and international borders, complete the government's toolkit.

Enterprises, on the other hand, have a key role in driving innovation from the ground up. Investments in R&D, particularly in technologies central to Society 5.0, are indispensable for staying competitive. Talent management strategies focused on continuous learning and diversity can catalyze a culture of innovation. A customer-centric approach fueled by advanced analytics and AI provides a tangible way to measure the impact of Society 5.0 technologies on individual lives. Moreover, supply chain optimization and risk management encapsulate the operational resilience that enterprises must develop to succeed in a hyper-connected world.

The absence of any empirical research is the paper's most significant limitation. However, because the primary goal of this study was to develop a theoretical model, this limitation is not particularly severe. Individual aspects of improving entrepreneurship ecosystems and factors affecting competitiveness are suggested for future research.

Acknowledgement

This paper has been supported by the Provincial Secretariat for Higher Education and Scientific Research of the Autonomous Province of Vojvodina, number: 142-451-2963/2022-03.

References:

[1] S. Huang, B. Wang, X. Li, P. Zheng, D. Mourtzis, L. Wang, L., Industry 5.0 and Society 5.0 - Comparison, complementation and co-evolution, Journal of manufacturing systems, 64 (2022) 424-428. https://doi.org/10.1016/j.jmsy.2022.07.010

[2] M. I. Muttaqin, A. S. Rifa'i, A. N. Rosyidah, A. A. Prisma, S. Kanitatun, U. Maskanah, Internalization of Morals Form The Character Of Muslim Islamic Entrepreneurship in The Era of Society 5.0, Mudir: Jurnal Manajemen Pendidikan, 5 (2023), 314-327. https://doi.org/10.55352/mudir.v5i2.589

[3] F. J. Álvarez-Rodríguez, R. A. A. Vera, Assessment of digital graduation competences for programs degrees in computing and information technology under the society 5.0 paradigm, IEEE Revista Iberoamericana de Tecnologias del Aprendizaje, 17 (2022), 208-214. https://doi.org/10.1109/RITA.2022.3167006

[4] I. Shaffiyah, V. Dwiyanti, A. Masek, Smart city and society 5.0: Involvement of information technology in the development of public service systems in Indonesia, ASEAN Journal of Community Service and Education, 1 (2022), 31-42.

[5] D. D. Mirzaakbarov, Directions for the introduction of new information technologies in education. Экономика и социум, 4-1 (2021), 211-214.

[6] C,.Volkmann, K. Fichter, M. Klofsten, D. B. Audretsch, Sustainable entrepreneurial ecosystems: An emerging field of research, Small Business Economics, 56 (2021), 1047-55.

[7] K. Fukuda, Science, technology and innovation ecosystem transformation toward society 5.0. International Journal of Production Economics, 1 (2020), 107460.

[8] V. Potočan, M. Mulej, Z. Nedelko, Society 5.0: balancing of Industry 4.0, economic advancement and social problems, Kybernetes, 50 (2020), 794-811.

[9] S. Balaji, K. Nathani, R. Santhakumar, IoT technology, applications and challenges: a contemporary survey. Wireless Personal Communications, 108 (2019), 363-388. https://doi.org/10.1007/s11277-019-06407-w
[10] A. Hasija, T. L. Esper, In artificial intelligence (AI) we trust: A qualitative investigation of AI technology acceptance, Journal of Business Logistics, 43 (2022), 388-412. https://doi.org/10.1111/jbl.12301
[11] Z. Cao, X. Shi, A systematic literature review of entrepreneurial ecosystems in advanced and emerging economies. Small Business Economics, 57 (201), 75-110. https://doi.org/10.1007/s11187-020-00326-y
[12] B. Wurth, E. Stam, B. Spigel, Toward an entrepreneurial ecosystem research program, Entrepreneurship Theory and Practice, 46 (2022), 729-778. https://doi.org/10.1177/1042258721998948
[13] Q. Kang, H. Li, Y. Cheng, S. Kraus, Entrepreneurial ecosystems: analysing the status quo. Knowledge Management Research & Practice, 19 (2021), 8-20. https://doi.org/10.1080/14778238.2019.1701964
[14] A. Berman, M. Cano-Kollmann, R. Mudambi, Innovation and entrepreneurial ecosystems: fintech in the financial services industry, Review of Managerial Science, (2021), 1-20.
[15] E. Liguori, J. S. Bendickson, Rising to the challenge: Entrepreneurship ecosystems and SDG success, Journal of the International Council for Small Business, 1 (2020), 118-125. https://doi.org/10.1080/26437015.2020.1827900

On the Applicability of Bloom's Taxonomy and Teacher Digital Competencies for Learning how to Code in Primary Schools

Maja Videnovik[1], Ana Madevska Bogdanova[2], Elena Vlahu Gjorgjievska[3] and Vladimir Trajkovik[2]

[1] *Center for Innovations and Digital Education Dig-Ed, Solunska glava 3, Skopje, North Macedonia;*
[2] *Faculty of Computer Science and Engineering, Ss Cyril and Methodius University in Skopje, Rugjer Boshkovikj 16, Skopje, North Macedonia*

[3] *School of Computing and Information Technology, University of Wollongong, Northfields Ave, Wollongong, NSW 2522, Australia*

maja@dig-ed.org; ana.madevska.bogdanova@finki.ukim.mk; elenavg@uow.edu.au; vladimir.trajkovik@finki.ukim.mk;

Abstract:
This paper investigates the effectiveness of game-based learning and gamification in teaching coding to primary school students. The paper also examines how teachers' implementation of coding subjects in low grades in primary schools can be supported. The study is motivated by the fact that most primary teachers lack computing backgrounds and may struggle to integrate coding effectively. To address this, a pilot study was conducted in six primary schools in North Macedonia to evaluate a proposed approach for teaching coding using game-based learning and gamification. In addition, the study investigates whether Bloom's taxonomy is suitable for evaluating the learning outcomes of a computational thinking course and which activities are most appropriate for teachers without formal coding training. The findings suggest that game-based learning and gamification can enhance critical thinking and problem-solving skills and help achieve educational goals for primary school students, regardless of teachers' coding skills.

Keywords:
coding, computational thinking, game-based learning, Bloom's taxonomy, teacher digital competences

1. Introduction

Computational thinking and understanding how computers execute programs are necessary 21st-century skills that can drastically influence the future labour market. Moreover, the reliance on computer code is not limited to the technology sector anymore, with an increasing number of businesses relying on it. Thus, understanding basic coding principles is becoming a necessary skill.

Educators face the challenge of defining the right learning objectives and strategies for implementing the subjects that develop students' critical thinking and problem-solving skills [1]. Since most students use the internet and mobile technologies daily, using familiar technology to learn coding seems natural. However, a logical framework for teaching computational thinking in primary education needs to be based on carefully selected tools and practices in teaching, learning, and instruction [2]. Furthermore, the subjects should be presented to address students' possible change of focus from general to specific knowledge [3].

In most countries, primary teachers who need to integrate basic coding skills in their classes have no computing background [4]. Additionally, there is little evidence of the problems teachers face and how they can integrate coding effectively [5]. Based on observations, it is crucial to support teachers in implementing the coding subject and provide different approaches for learning how to code. Furthermore, it is essential since educators generally lack training in different digital tools and regularly face the need to improve their digital literacy.

The main goal of this paper is to investigate which activities are most appropriate to be implemented in coding subjects for teachers without formal coding training. Furthermore, the paper investigates whether the educational goals' complexity concerning coding can be treated similarly to other subjects.

The test framework used in this paper builds on the approach published in [6] and includes cooperative and competitive cycles spanning several school years. Its initial cycle introduces game-based tools that lead the students to develop their computational thinking. Then, the students are acquainted with more challenging problems that require algorithm-based problem-solving skills using block-based coding. In the final phase, the students have already obtained the prerequisite knowledge and can start to learn and use different programming languages connecting coding with real-life situations.

The teacher's (educator) role in these processes varies from the facilitator in the initial phase to the mentor in the final phase of the methodology. However, the approach corresponds with teacher qualifications since the initial phase covers most (if not all) the students at a younger age, while the final phase is for some students in higher classes. Therefore, the teachers acting as facilitators do not need specific programming skills, while the teachers acting as mentors need suitable qualifications for teaching coding. That is the case with higher classes coding teachers.

Student's achievement of the educational goals while implementing the proposed methodological framework and defining the development of coding and computational thinking skills determines the framework's suitability in coding subjects. According to Bloom's taxonomy, teachers define these cognitive and educational goals, carefully adjusting learning outcomes to different complexity levels.

Bloom's taxonomy is a widely used framework for assessing students' depth of knowledge. It establishes a hierarchy of six levels (remembering, understanding, applying, analyzing, evaluating, and creating) for increasing the degree of cognitive expertise of the student. Every level assumes the student's capability for the lower levels. Research in implementing Bloom's taxonomy in teaching programming fundamentals reveals different interpretations and revisions of Bloom's taxonomy [7]. In this paper, we aim to discuss the relevancy of Bloom's taxonomy for defining educational goals concerning the development of computational thinking and coding skills.

2. Methodology

Implementing innovative practices in elementary education should prioritize student-centred approaches facilitated by Information and Communication Technology (ICT) with a coaching role for teachers. The curriculum for ICT subjects and topics integrated into various subjects should adhere to principles outlined in the European Union Digital Competence Framework [8]. The suggested framework aims to provide fundamental skills and knowledge to a large student population while enabling higher programming skills for advanced students. The spiral curriculum enables revisiting specific topics in higher classes to enhance previous knowledge, increasing students' complexity levels and knowledge throughout their schooling. This progression applies to coding, with students starting with simple programs and advancing to more complex coding in subsequent years. This approach allows for minimum coding competencies in large classes while providing specialized classes with more advanced coding competencies. The proposed methodology fosters teamwork and individual learning by using various tools and educational paradigms.

Following the constructivism theory [9], a methodological approach is suggested for teaching coding by incorporating students' experiences and reflections [10]. Game-based learning [11] and project-based learning [12], and gamification elements [13] are used to enhance students' computational thinking skills by encouraging them to explore, practice, and collaborate. The aim is to increase motivation, persistence, and engagement in learning while enabling collaboration among students with varying knowledge backgrounds. The flipped classroom approach [14] allows students to utilize additional knowledge to achieve learning outcomes related to coding, with new knowledge built upon reflecting and applying previously acquired knowledge. By using collaboration, game-based, and project-based learning, student-centred learning is established, and the teacher's primary role is to facilitate this collaboration and encourage students' reflection on their learning. After introducing coding concepts through a collaborative approach, students master the learning topics individually.

The methodology for developing students' coding skills consists of three mandatory cycles and one optional cycle, with students repeating practice on coding principles using different tools in each cycle. This graduate approach enables students to progress according to their interests. Many games, programming platforms, and tools can be utilized in this approach, with "Scottie Go!" [15], "Scratch" [16], "Micro:Bit" [17], and "Python" [18] being the specific tools selected in this instance (Figure 1). However, the framework focuses on skills and learning approaches rather than tools.

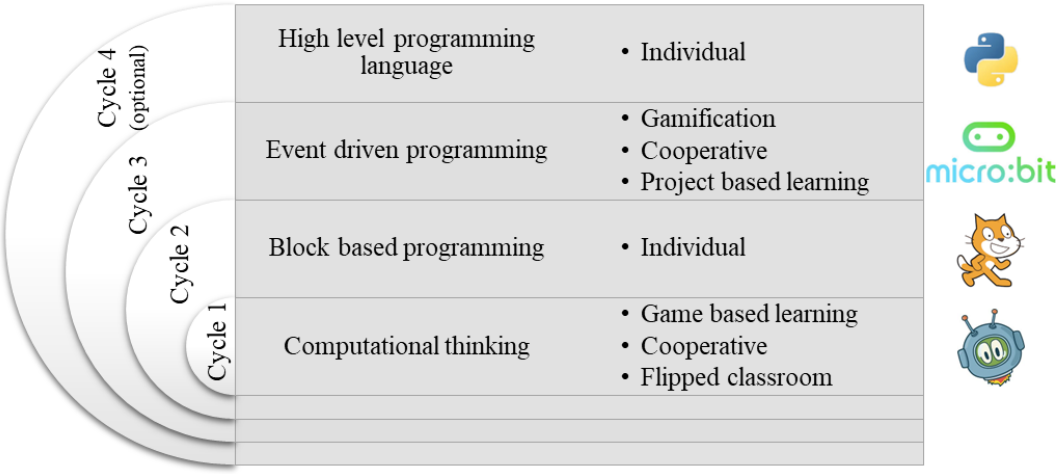


Figure 1. A methodological framework for developing students' coding skills

The proposed approach aims to teach coding using games and gamification techniques to engage students in learning [19, 20]. Through this process, students develop problem-solving skills, computational thinking, knowledge of algorithms and programming concepts, and the ability to create programs in various programming environments. However, the level of acquired coding knowledge and skills according to the educational goal complexity is the topic of further discussion. Furthermore, obstacles such as teachers' lack of digital competencies [21] can prevent the successful integration of the proposed methodology. To overcome these challenges, the approach is designed to progress students from understanding basic programs to upgrading their skills by creating solutions and then mastering their knowledge by demonstrating the practical use of their programming skills in new situations. However, the primary goal of the approach is to teach children how to think differently and experiment creatively to solve problems, which builds their confidence in a fun and exciting way.

Therefore, the research questions raised by this approach are whether the Bloom taxonomy is relevant for describing educational goals for coding skills and whether this approach can minimize the level of required coding skills among teachers.

2.1. Participants

A mixed-method pilot study was conducted in six primary schools in North Macedonia, using interviews with the teachers and a survey for the students. The study evaluated a proposed approach for teaching coding using game-based learning and gamification. The approach was implemented as part of the compulsory coding subject for students from the third grade (8 years old).

The study used "Scottie Go!" and "Micro:bit" as tools for learning to code. 121 students participated in the survey about "Scottie Go!" with 52% male and 48% female students, 76% from urban and 24% from rural areas. The "Micro:bit" survey was shared among 78 students, with 50.6% male and 49.4% female, as well as 66.2% from urban and 33.8% from rural areas. The survey was conducted in four well-equipped schools in urban areas and two not-well-equipped schools in rural areas.

Six teachers teaching coding classes in the surveyed schools were interviewed. The evaluation of the methodological framework aimed to investigate the relevance of Bloom's taxonomy for describing educational goals for coding skills and whether this approach minimizes the level of needed coding skills among the teacher population.

2.2. Procedure and Instruments

The study used two tools: the “Scottie Go!” board game and the BBC “Micro:bit” platform, along with the popular coding environments “Scratch” and “Python”. The focus is put on “Scottie Go!” and “Micro:bit” as tools that promote collaborative learning environments through game-based learning and gamification.

The “Scottie Go!” game combines physical cardboard tiles with a mobile app that sets coding tasks and scans proposed code solutions. Students work in groups to create coding instructions and solve tasks, using their previous experience playing mobile and board games to engage with the learning process. The teacher acts as a facilitator, helping students share solutions and develop reflection skills without requiring additional teacher training in coding.

The “Micro:bit” tool is presented to the students as a tool needed to develop an interdisciplinary project. Students work in teams, with a small competition for the best project. In this case, the teacher also acts as a facilitator, although some understanding of “Scratch” is required since “Micro:bit” can be programmed using blocks, similar to “Scratch”. The study found no need for additional teacher training in coding, especially because “Scratch” was used in a previous methodological framework cycle.

The teaching materials used in the study cover the rules (syntax), environment, and problems that need to be solved. Students develop computational thinking skills during the play, demonstrating decomposition of the problem, logical thinking, abstraction, finding patterns, creating algorithms, and evaluating the created program. The students' learning tasks are typical problem-solving assignments, such as how to most efficiently get from one point to another using a predefined set of rules and avoiding obstacles.

The study [22] aimed to evaluate the effectiveness of this methodology in achieving educational goals of varying complexity. The evaluation was conducted through a short test consisting of 10 multiple-choice questions concerning the linear structure in programming, which aimed to measure the level of retention of students' learning with a playful approach. According to Bloom's taxonomy, the questions in the test refer to different levels of achieving learning outcomes: remembering facts, comprehension, applying the knowledge in new situations and higher-thinking skills (analysis, evaluation and creating).

The interview with the teachers that implemented the proposed methodology in their schools was conducted to investigate the acceptability of the methodological approach and the level of teachers' competencies necessary for its implementation. Additionally, the teachers' opinions regarding the benefits of this approach for developing students' computational and coding skills were collected.

3. Results and discussion

The study results concerning the students' achieved learning outcomes are presented in Figure 2, showing the percentage of students' correct answers on questions with different complexity levels. The results revealed that students using "Scottie Go!" exhibited more significant achievement in more complex educational goals than "Micro:bit". Specifically, game-based learning and flipped classrooms (“Scottie Go!”) led to more significant achievement of more complex educational goals. These results are expectable because “Micro:bit” is proposed for use in the third cycle, which specifically emphasizes the comprehension of programming tools and requires prior coding knowledge and skills. In addition, all students achieved educational goals that required only the reproduction of factual information, and the percentage of correct answers decreased with the questions' complexity increment.

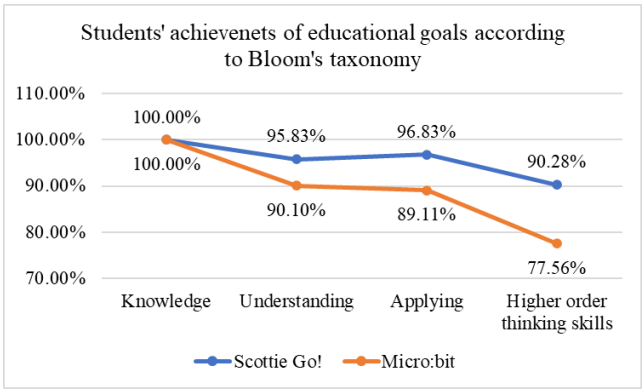


Figure 2. Students' achievement of educational goals with different complexity levels

The results indicate that students had slightly better outcomes in applying their knowledge in new scenarios than in comprehending coding concepts when learning with "Scottie Go!". This finding raises an important question regarding the suitability of Bloom's taxonomy, commonly used to evaluate students' cognitive knowledge in assessing the learning outcomes of a computational thinking course. Notably, experts in the field are divided on interpreting Bloom's taxonomy in computational thinking tasks [23]. For instance, the taxonomy considers creating as a more challenging task than understanding. However, it is debatable whether creating a basic project, such as moving a sprite from one point to another, is more cognitively demanding than fully comprehending the concept of concurrency [24]. One possible explanation for our results could be attributed to the complexity of the assigned tasks concerning applying gathered knowledge in new situations. Research indicates that in disciplines such as computer science, the new category "higher application" should be added at the top of Bloom's taxonomy as the highest level of gathered knowledge, where evaluation and creation are used during the process of applying knowledge in new situations (e.g. complex programs) [6]. However, this is inconclusive for coding using "Micro:bit", although students' answers do not differ much in these two categories (understanding and applying). This can be a result of the tasks given to students using "Micro:bit" at applying level, which require elements of creation and evaluation

The interviews were conducted with teachers to investigate their attitudes towards using the tools and teaching methods in the classroom.

The findings from the interviews with the teachers indicate that the educational tools under research were perceived to be user-friendly, and students did not encounter any difficulties using them in the classroom. All teachers who participated in the study strongly agreed that integrating the tools into their teaching practices positively impacted student motivation and engagement with learning. Furthermore, the teachers reported that the tools fostered a collaborative and stimulating learning environment and that working in teams and learning from mistakes helped students to persist in achieving their learning objectives. The teachers appreciated that using a flipped-classroom approach and gamification eliminated the need for additional training in coding. They also highlighted that their role as facilitators allowed them to emphasize problem-solving and critical thinking skills, in addition to developing coding skills, by focusing on learning outcomes. One of the respondents also reports that it is a useful example of using technology the students are more familiar with.

Regarding the use of technology as a pedagogical tool, the technology was used to assess students' learning of applying theoretical concepts in practice and to detect where they have challenges, establishing guidelines for improvement. In cases where the students have difficulties using the tool, the role of the teacher is to combine the students with difficulties with students who work on similar projects but have no difficulties implementing them. This approach is an excellent way to promote student collaboration. However, from the teachers' perspective, the students are more motivated and inspired, and they focus more on learning outcomes by establishing needed functionalities with different tools. In parallel, students master the technological tools as well. The competitive cycle enhances the engagement and students' eagerness to do better.

Based on the interview results, the teachers believe that "Scottie Go!" and "Micro:bit" can effectively enhance critical thinking and problem-solving skills and achieve educational goals. Interestingly, some teachers who lacked prior coding skills expressed hesitancy in implementing these tools in the

classroom. However, after seeing the students take the lead and collaborate to achieve the learning outcomes, they were satisfied with the results. The teachers could serve as facilitators, and their lack of coding skills did not hinder the attainment of learning objectives. Consequently, it can be inferred that the proposed framework reduces the coding skills required among teachers.

4. Conclusion

This paper explores the effectiveness of game-based learning and gamification in teaching coding skills to primary school students. The study investigated whether Bloom's taxonomy can be applied to describe educational goals for coding skills and to determine the most appropriate activities for teachers without formal coding training. The results indicated that students using more straightforward tools achieved significant educational goals, especially applying their knowledge to new scenarios. The findings suggest that Bloom's taxonomy may not be suitable for evaluating the learning outcomes of a computational thinking course.

Additionally, the study shows that teachers who lack coding skills could effectively facilitate the implementation of these tools in the classroom, and their lack of coding skills did not hinder the attainment of learning objectives. This outcome reduces the level of needed coding skills among the teachers' population. Overall, this study provides empirical evidence that game-based learning and gamification can enhance critical thinking and problem-solving skills and help achieve educational goals for primary school students, regardless of teachers' coding skills. This study contributes to understanding the effectiveness of game-based learning and gamification in teaching coding skills, which has practical implications for educators, policymakers, and researchers.

References:

[1] D. W. Johnson, R. T. Johnson, Cooperative, competitive, and individualistic learning, Journal of Research & Development in Education 12(1) (1978) 3–15.

[2] S. L. Mason, P. J. Rich, Preparing elementary school teachers to teach computing, coding, and computational thinking, Contemporary Issues in Technology and Teacher Education 19(4) (2019) 790-824.

[3] E. V. Soboleva, T. N. Suvorova, A. V. Grinshkun, M. I. Bocharov, Applying gamification in learning the basics of algorithmization and programming to improve the quality of students' educational results, European Journal of Contemporary Education 10(4) (2021) 987-1002. doi:10.13187/ejced.2021.4.987.

[4] P. J. Rich, B. Jones, O. Belikov, E. Yoshikawa, M. Perkins, Computing and engineering in elementary school: The effect of year-long training on elementary teacher self-efficacy and beliefs about teaching computing and engineering, International Journal of Computer Science Education in Schools 1(1) (2017) 1-20. doi:10.21585/ijcses.v1i1.6.

[5] European Schoolnet, Computing our future - Computer programming and coding Priorities, school curricula and initiatives across Europe, 2015. URL: http://fcl.eun.org/documents/10180/14689/Computing+our+future_final.pdf/746e36b1-e1a6-4bf1-8105-ea27c0d2bbe0.

[6] E. Vlahu-Gjorgievska, M. Videnovik, V. Trajkovik, Computational thinking and coding subject in primary schools: Methodological approach based on alternative cooperative and individual learning cycles, 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), Wollongong, NSW, Australia, IEEE, 2018, pp. 77-83. doi:10.1109/TALE.2018.8615334.

[7] S. R. Sobral, Bloom's Taxonomy to improve teaching-learning in introduction to programming, International Journal of Information and Education Technology 11(3) (2021) 148-153. doi:10.18178/ijiet.2021.11.3.1504.

[8] European Commission – EU Science Hub, The Digital Competence Framework 2.0, 2017. URL: https://ec.europa.eu/jrc/en/digcomp/digital-competence-framework.

[9] M. Muhajirah, Basic of learning theory: (behaviourism, cognitivism, constructivism, and humanism), International Journal of Asian Education 1 (2020) 37-42. doi:10.46966/ijae.v1i1.23.

[10] Y. Tran, Computational thinking equity in elementary classrooms: What third-grade students know and can do, Journal of Educational Computing Research 57(1) (2019) 3-31. doi:10.1177/0735633117743918.
[11] V. Trajkovik, T. Malinovski, T. Vasileva-Stojanovska, M. Vasileva, Traditional games in elementary school: Relationships of student's personality traits, motivation and experience with learning outcomes, PloS one 13(8) (2018) e0202172. doi:10.1371/journal.pone.0202172.
[12] S. Safaruddin, N. Ibrahim, J. Juhaeni, H. Harmilawati, L. Qadrianti, The effect of project-based learning assisted by electronic media on learning motivation and science process skills, Journal of Innovation in Educational and Cultural Research 1.1 (2020): 22-29. doi:10.46843/jiecr.v1i1.5.
[13] M. Beatriz, J. Frez, J. Cruz-Lemus, M. Genero, An empirical investigation on the benefits of gamification in programming courses, ACM Transactions on Computing Education (TOCE) 19(1) (2018) 1-22. doi:10.1145/3231709.
[14] J. L. Bishop, M. A. Verleger, The flipped classroom: A survey of the research, in: Proceedings of the ASEE national conference, 30(9), 2013, pp.1-18. Doi:10.18260/1-2—22585.
[15] M. Videnovik, V. Trajkovik, Using Scottie Go! as Game-Based Learning Tool for Computational Thinking Course, in: ICERI2018 Proceedings, IATED, 2018, pp. 1337-1344. doi:10.21125/iceri.2018.1304.
[16] J. Fagerlund, P. Häkkinen, M. Vesisenaho, J. Viiri, Computational thinking in programming with Scratch in primary schools: A systematic review, Computer Applications in Engineering Education 29(1) (2021) 12-28. doi:10.1002/cae.22255
[17] M. Videnovik, E. Zdravevski, P. Lameski, V. Trajkovik, The BBC micro: bit in the international classroom: learning experiences and first impressions, in: 17th International Conference on Information Technology Based Higher Education and Training (ITHET), IEEE, 2018, pp. 1-5. doi:10.1109/ITHET.2018.8424786.
[18] N. Tabet, H. Gedawy, H. Alshikhabobakr, S. Razak, From Alice to Python. Introducing text-based programming in middle schools, in: Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education, 2016, pp. 124-129. doi:10.1145/2899415.2899462.
[19] M. Videnovik, V. Trajkovik, L. V. Kiønig, T. Vold, Increasing quality of learning experience using augmented reality educational games, Multimedia tools and applications 79(33-34) (2020) 23861-23885. doi:10.2196/32095
[20] T. Vasileva-Stojanovska, M. Vasileva, T. Malinovski, V. Trajkovik, The educational prospects of traditional games as learning activities of modern students, in: Proceedings of the European Conference on Games-Based Learning 2, 2014, pp. 746-759.
[21] E. J. Instefjord, E. Munthe, Educating digitally competent teachers: A study of integration of professional digital competence in teacher education, Teaching and teacher education 67 (2017) 37-45. doi:10.1016/j.tate.2017.05.016.
[22] M. Videnovik, E. Vlahu-Gjorgievska, V. Trajkovik, To code or not to code: introducing coding in primary schools, Computer Applications in Engineering Education 29(5) (2021) 1132-1145. doi:10.1002/cae.22369
[23] C. G. Johnson, U. Fuller, Is Bloom's taxonomy appropriate for computer science?, in: Proceedings of the 6th Baltic Sea conference on Computing education research: Koli Calling ACM, New York, USA, 2006, pp. 120-123.
[24] O. Meerbaum-Salant, M. Armoni, M. Ben-Ari, Learning computer science concepts from scratch, Computer Science Education 23(3) (2013) 239-264. doi:10.1145/1839594.1839607

Development of Human Resources in the Digital Age

Svetlana Stojkov[1], Mila Kavalić[1], Edit Terek[1]

[1] *University of Novi Sad, Technical faculty "Mihajlo Pupin", Zrenjanin, Serbia*

svetlana.stojkov@tfzr.rs; mila.kavalic@tfzr.rs; edit.terek@tfzr.rs

Abstract:
This paper provides an overview of existing literature on the role of human resource management in the digital age. The aim of this work is to indicate what human resources face in the digital age and what competencies are needed to master new technologies. The importance of professional training was emphasized, in order to achieve a better performance of the company. The application of new technologies in organizations enables a faster flow of information, but also leads to a change in management methods. It also affects the motivation and development of employees' creativity. In order for the organization to operate efficiently and productively, it is necessary to create a strategy plan for implementing digital transformation, with the use of effective digital tools for implementing business processes.

Keywords:
HR, HR management, digitalization, digital transformation, competencies

1. Introduction

Conceptual components such as "transformation" and even more "digitalization of human resource management (HR)" imply essential changes in human resource management, hinting at the importance of these concepts. This statement is confirmed by research such as [1, 2, 3, 4]. Human resources face numerous challenges when it comes to the application of new technologies. In order to facilitate and adapt to rapid technological changes and developments, organizations are expected to develop procedures and establish practices for continuous review of employee competencies. They are also expected to introduce new forms of work organization. Using effective tools enables employees and human resources management to carry out business processes more efficiently, create additional value for business and provide organizations with a competitive advantage.

Research questions:

1. Do organizations see digital transformation as an opportunity or an obstacle?
2. Do organizations invest in the education and professional development of their employees?
3. Does the use of digital tools have a positive effect on optimization and achievement of goals?
4. Do employed managers have adequate skills to master digital competencies?

2. Methodological research framework

In the following subsections, the authors present the methodological research framework.

2.1. Problem and subject of research

The changes, brought by the new digital age, are increasingly intense and require adaptation of all sectors. Classical managerial skills, as the most important economic resource, are no longer sufficient to achieve the necessary levels of digital competence. To create the workplace of the future and effectively facilitate digital work, organizations provide new and innovative information technologies. Mastering digital competences is especially important for the development of skills and additional training of population groups that need this kind of education. The subject of the work is competences and the development of human resources in the digital age. Organizations are forced to change the

traditional way of doing business to a digital one. This leads to a gap in the human resource management sector. Education and adequate skills of the employees become imperative for the survival of the organization in the modern era.

2.2. Research objective and research questions

The aim of this paper is to indicate what human resources face in the digital age, what competencies are needed to master new technologies, how education can help in the further development of human resources. It was emphasized that the success and professional development of an employee depends on the knowledge he possesses. HR managers and project managers need to influence and change the mindset of their employees to prepare them for the future.

3. Theoretical framework

In the following subsections, the authors present the theoretical framework of the research.

3.1. Competences and development of digital business transformation

From a business perspective, a digital business transformation strategy aims to transform products, markets, business processes and organizational aspects using digital technologies. Wade et al. (2015) define digital transformation as: "organizational changes through the use of digital technology and business models to improve business performance". Similarly, Ismail et al. (2017) define digital transformation as the process of incorporating multiple new digital transformations, with the intention of achieving superior performance and sustainable competitive advantage. Analogously, researchers like Li et al (2018); Burchardt et al. (2019) ; Cichosz et al. (2020) define the concept of digital transformation as a "fusion of advanced technologies" that integrate physical and digital systems. Semečenko and Vasilić (2020), define digital transformation as a complex, dynamic, continuous and necessary process of transformation of all organizational aspects. With a strategically designed integral application of modern digital business transformations, it can result in the creation of a new business model. This puts the customer at the center of all activities and decisions the organization makes. The ultimate goal is to create conditions for improving innovation, a better market position, and thus improving overall business results. Digital transformation also implies new business concepts and rules both internally and externally [5, 6, 7]. Roughly speaking, such concepts denote the increasing use of technology and corresponding fundamental changes in numerous domains of business and society. Therefore, the process of organizational changes must be continuous because without it business in today's world becomes impossible [8, 9, 10]. Since the so-called Covid-19 crisis of 2020, companies have been forced to introduce at least temporary solutions in the way organizations work in all sectors and regions. Today, we cannot imagine business as it was before the pandemic. In the conditions of the fourth industrial revolution (I4.0), the company is rapidly developing and marketing products with a developed working structure. In order to achieve these goals, companies are forced to make radical changes in the way they do business, to restructure business strategies and business models [11, 12, 13, 14, 15]. Such an approach to business enables companies to integrate digital products into their business processes and products. Research such as: Rodriguez-Molano et al. (2018); Savić and others. (2019); Zehir (2020) and Kahrović and Kahrović (2021), show us that a large number of organizations are looking for a way to successfully use digital transformation as an opportunity. This is why they invest in education, train their employees, work on the creation and development of innovative products. The rapid development and adoption of technologies such as Cloud Computing, Edge Computing, mobile applications, social networks, big data analytics and the Internet of Things (IoT) are helping organizations leverage the features of such transformative technologies. The development of technologies helps organizations to improve business agility, accelerate innovation and transform business processes and introduce new business models.

3.2. Digital economy - Challenge or obstacle

The digital economy, an economy that functions primarily through digital technology, has enabled organizations to develop their own capabilities, infrastructure and competitiveness. In this way, it influenced the creation of the concept of digital business transformation. The digital economy can be an obstacle or a challenge for organizations, as confirmed by research such as [16, 17, 18]. Even Don Tapscott in the first published book on the digital economy in the world "The Digital Economy: Promise and Peril in the Age of Networked Intelligence" emphasizes the importance of the digital economy, which in his opinion is an economy based on intellectual property and knowledge workers. Business transformation requires the integration of business transformation in all areas. Digital transformation provides superior employee analysis and reporting features. It also enables performance measurement and forecasting of employee activities. This significantly improves interaction and collaboration with HR professionals and organization management [19; 6]. Employees can use various digital tools for work and reporting, which have a positive effect on the optimization of work processes. These tools enable faster and more efficient achievement of the organization's goals, and managers are given the freedom to search for and hire people with adequate qualifications and skills [20, 21, 22]. Stable organizations that manage to maintain their business have understood the difficulties and the need to have different digital options to reach the customer. Consequently, organizations provide outstanding and innovative products ahead of the competition. Companies whose operations are based on the principles of the digital economy are extremely dependent on the stability of their employees. That is why the success of the company largely depends on the structure of human resources. Loyalty to the company, motivation and willingness to cooperate and accept innovations are important segments that lead to ideas and solutions through teamwork [23, 24].

3.3. Digital HR management - professional training

The digital age has changed the business framework for organizations where the imperative of a competent project manager is education. Some of the previous studies such as [25, 26] indicate that the integration of digital technology affects the simplified and improved work experience of human resources employees. The role of the human resource management function in this context refers to the transformation of jobs and organizational structures. It also affects the development of procedures and standards that support employees in changing working conditions and changing customer demands [27, 8, 28]. Continuous education is a necessary condition for the development of human resources, and essentially it implies work in several different areas in project management. Research conducted by Savić et al. (2019) show that a large number of organizations are trying to successfully use digital transformation as an opportunity. For this reason, they invest in the education and training of their managers. They also work to create and develop their innovative products and services. Adopting new technologies enables the HR function to develop planning strategies by working to identify and audit critical roles and establish contingency plans. Knowledge should be constantly updated and increased in order to ensure the necessary level of personality development. A key element to the success of this process is effective communication to ensure employee readiness. However, there is a problem in the labor market even though there is a need to hire new experts. Consequently, with the application of digital technology, many jobs are performed by machines and robots instead of people. That is the reason why a large number of people remain unemployed. These people need to develop some other, additional competencies and skills in order to continue to be competitive in the labor market. That is why the top management must know how to put together a good team, whose knowledge, expertise, motivation and communication will come together and create a new dimension of the company's success. [29, 30, 31]. It is necessary for employees to participate in various forms of professional development of employees and this should be an integral part of maintaining the level of knowledge and skills. This will additionally affect the improvement of human resources and the necessary monitoring of the development of science and technology. Accordingly, it will affect the improvement of productivity and the overall quality of business. It will also be easier to follow the goals and development of the organization, the development of the market and the changes that occur in it.

3.4. Project organization - mastering digital skills

The new age project approach has a flexible and up-to-date attitude towards problems. However, only a competent person who possesses the knowledge and skills of the modern era can provide answers to current changes and emerging problems. As technology evolves, new jobs require more skilled workers whose job it is to ask the right questions and identify critical problems. Access to more data increases the need for human intelligence to understand how to use it. When a project is implemented in a functional organization, the project organization requires the project manager to design and structure the activities. This includes planning, implementation and control of project implementation and forming a team that will work on the implementation of the existing project. However, organizations often do not have enough human resources. That is why employees perform existing tasks in the organization and activities on projects in which they participate. To be able to drive future organizational performance, HR managers and professionals need to improve skills and competencies and acquire new ones. However, only a small part of HR managers use domestic companies to use these technologies [32]. Also, in order for managers and employees to be ready for action, they need to direct knowledge and development towards training and mastering digital skills. It is not only necessary to hire people who are trained to work in the digital age, but also to develop the necessary skills in existing employees. The authors of Elia et al. (2020) believe that it is necessary to adopt collective intelligence in order for the digital system to develop. In this regard, four dimensions were created [33]:

- digital participants (DC),
- digital activities (what),
- digital motivations (why) i
- digital organization (how)

Strategically significant technologies in the digital age, which can significantly affect the corporate market in the coming years, are social networks, mobile devices, cloud computing and large database analytics. The role of digital technologies in the implementation of increasingly diverse functional requirements of business systems and their clients is particularly significant, such as:

- electronic advertising and promotion,
- electronic contacts between business partners,
- online sales, delivery and payment activities,
- electronic ordering of products,
- single system business process support such as inventory query capabilities and
- updating delivery status and orders in commercial transactions, etc.

Modern society, due to rapid and sudden changes, has to adopt, process and distribute a huge amount of information. This leads all economic subjects to constant learning in order to be ready to quickly discard what they have learned and, if necessary, adopt new information ready for processing. Technologies such as artificial intelligence (AI), virtual reality (VR), augmented reality (AR), and blockchain supplemented with data science can be used to improve all HR functions [34, 35, 31]. Digital platforms such as cloud computing, Internet of Things, big data analytics increase the capacity of employees. As a form of digital business models, they represent a key element in building successful digital ecosystems [36]. Authors like Hanandeh et al. (2022) indicate that it is necessary and urgent for organizations to start the process of digital transformation and begin to fully rely on artificial intelligence tools and operations. This includes all processes of business transactions, documents and orders. Effective management of big data provides companies with numerous advantages. Some of them are more competitive advantages and new knowledge, supporting and developing the decision-making process and providing users with more services and new products [37].

4. Final considerations

In the following subsections, the authors present the final considerations of the research.

4.1. Scientific and social justification of research

The research provides a framework that facilitates the connection between digital transformation, the use of digital tools and the role of HR in that process. This involves the cooperation of organizational members, skills, learning and IT support.

4.2. Applicability of research results

The paper clarifies the role of human resources and the effect between company performance. In this process, the use of digital tools, cooperation of organizational members, skills, learning and IT support are particularly important. Modern society, due to rapid and abrupt changes, has to adopt, process and distribute a huge amount of information. It is necessary for employees to participate in various forms of professional training, and this should be an integral part of maintaining the level of knowledge and skills. This leads all economic subjects to constant learning in order to be ready to quickly discard what they have learned and, if necessary, adopt new information ready to be processed. This will further influence the improvement of human resources and the improvement of productivity and overall quality of business. The paper also emphasizes the importance of the essential elements of transformation related to operational processes, business models and digital platforms. The prominent elements of business transformation relate to competition, innovation and values. Digital transformation provides superior functions of analysis and reporting of employees, and continuous education is a necessary condition for human resource development. In order for companies and HR sector management to accept and effectively use digital tools in business, they must become aware of some basic concepts. Knowledge, expertise, motivation and communication create a new dimension of the company's success. In order for managers and employees to be ready for action, they need to direct knowledge and development towards training and mastering digital skills. The adoption of new technologies enables the HR function to develop planning strategies by working to identify and review critical roles and establish contingency plans. Strategically designed application of digital tools creates the conditions for improving innovation, better market position, and therefore improving overall business results. The existing literature in the field of digital transformation of human resources and the role of employees in the digital age is not sufficiently designed, which is the main contribution of this paper.

4.3. Limitations and further research directions

Given that the area is complex, extensive and that the modern digital age is changing (in the context of digital technologies and business process management), it is necessary to investigate this area in more detail. Through further research at the level of large and small companies, it is possible to monitor changes in the management skills of employees in organizations and thus to conceive new business management strategies in the digital age.

5. Conclusions

Research by the mentioned authors shows that a large number of organizations are looking for a way to successfully use digital transformation as an opportunity. They also invest in the education and training of their employees and work on the creation and development of their innovative products. Such an approach to business gives the possibility of integrating digital products into your business processes and products. The digital economy can be an obstacle and a challenge for organizations. These statements provide an answer to research question number 1. Furthermore, it is necessary to adequately manage human resources in changing working conditions, invest in education, train employees, work on the creation and development of innovative products. The rapid development and adoption of technologies helps organizations take advantage of the characteristics of such transformative technologies, which answers the second research question. Stable organizations that manage to sustain their business have understood the difficulties and the need for different digital

capabilities to reach customers and provide distinguished and innovative products ahead of the competition. Regarding the third research question, the use of digital tools in this context has a positive effect on the optimization of work processes. In particular, it enables faster and more efficient achievement of organizational goals. The existence of various digital tools for work and reporting gives managers the freedom to search and recruit employees with adequate qualifications and skills. It is necessary for employees to participate in various forms of professional development of employees and this should be an integral part of maintaining the level of knowledge and skills. Regarding skills and the fourth research question, we can conclude that it is necessary to adopt collective intelligence in order for the digital system to develop. In order for companies to be ready for effective management, they need to direct knowledge and development in the direction of training and mastering the digital skills of their managers and all employees. It is not only necessary to hire people who are trained to work in the digital age, but also to develop the necessary skills in existing employees. Research so far shows us that predictors of human capital development and improvement methods, which every organization should use in order to survive and achieve performance, can only be ensured by continuous work.

References:

[1] J. Meijerink, A. E. Keegan, J. H. Marler, M. Boons, Digitization and the transformation of human resource management. The International Journal of Human Resource Management (2018) 6-12. doi:http://dx.doi.org/10.1080/09585192.2018.1503845

[2] I. Pantelidis, Digital Human Resource Management. Human Resource Management in the Hospitality Industry (2019) 3-6. doi:10.4324/9780429441400-21

[3] S. Strohmeier, Digital human resource management: A conceptual clarification. German Journal of Human Resource Management (2019) 345–365. doi: https://doi.org/10.1177/2397002220921131

[4] P. Vardarlier, Digital Transformation of Human Resource Management: Digital Applications and Strategic Tools in HRM. Digital Business Strategies in Blockchain Ecosystems (2020) 239-264. doi:10.1007/978-3-030-29739-8_11

[5] Rodríguez-Molano, Contreras-Bravo, Rivas-Trujillo, Industry Knowledge Management Model 4.0, in:Proceedings of the International Conference on Information Technology & Systems (ICITS 2018), Springer, 2018.

[6] Demartini, Pinna, Tonelli, Terzi, Sansone, Testa, Food industry digitalization: from challenges and trends to opportunities and solutions. IFACPapersOnLine (2018) 1371-1378. doi: https://doi.org/10.1016/j.ifacol.2018.08.337

[7] Rachinger, Rauter, Müller, Vorraber, Schirgi, Digitalization and its influence on business model innovation. Journal of Manufacturing Technology Management (2019) 114-120. doi: https://doi.org/10.1108/JMTM-01-2018-0020

[8] V. Milovanović, IMPACT OF TOTAL QUALITY MANAGEMENT ON COMPANY BUSINESS PERFORMANCE. UNIVERSITY OF KRAGUJEVAC, FACULTY OF ECONOMICS (2018) 3-5. URL:https://nardus.mpn.gov.rs/bitstream/handle/123456789/10517/Disertacija.pdf

[9] R. Z. Szabo, H. I. Vuksanovic, R. Hanák, Industry 4.0 Implementation in B2B Companies: Cross-Country Empirical Evidence on Digital Transformation in the CEE Region. Sustainability (2020) 3-5. doi: https://doi.org/10.3390/su12229538

[10] A. Dias, A. Carvalho, P. Sampaio, Quality 4.0: literature review analysis, definition and impacts of digital transformation processes on quality. International Journal of Quality & Reliability Management (2022) 1312-1335. doi: https://doi.org/10.1108/IJQRM-07-2021-0247

[11] F. Ponsignon, S. Kleinhans, G. Bressolles, The contribution of quality management to an organization's digital transformation: a qualitative study. Total Quality Management and Business Excellence (2019) 15-34. doi: 10.1080/14783363.2019.1665770

[12] N. Radziwill, S. in Woodhouse, Connected, Intelligent, Automated - the Definitive Guide to Digital Transformation and Quality 4.0, Quality Press, Milwaukee, Wisconsin, 2020, 19-20

[13] H. Rowlands i S. Milligan, Quality-driven industry 4.0. Sustainability and Innovation in the Fourth Industrial Revolution (2021) 3-30. doi: 10.1142/9789811230356_0001

[14] D. Bonnet, G. Westerman, How to Embrace Digital Transformation — First Steps. MIT SLOAN MANAGEMENT REVIEW (2021) 1-7. URL: https://blog.sodipress.com/wp-content/uploads/2021/03/how-to-embrace-digital-transformation.pdf

[15] J. Reis i N. Melão, Digital transformation: A meta-review and guidelines for future research. Heliyon, (2023) 4-10. doi: 10.1016/j.heliyon.2023.e12834

[16] T. Hess, A. Benlian, C. Matt, F. Wiesboeck, Options for formulating a digital transformation strategy. MIS Quarterly Executive (2016) 123-139. doi: 10.4324/9780429286797-7

[17] Lanzolla, Lorenz, Miron-Spektor, Schilling, Solinas, Tucci, Digital transformation: What is new if anything. Academy of Management Discoveries (2018) 378-387. doi: 10.5465/amd.2020.0144

[18] A. G. Frank, L. S. Dalenogare, N. F. Ayala, Industry 4.0 technologies: Implementation patterns in manufacturing companies. International Journal of Production Economics (2019) 15-26. doi: https://doi.org/10.1016/j.ijpe.2019.01.004

[19] J. H. Marler, E. Parry, Human resource management, strategic involvement and eHRM technology. The International Journal of Human Resource Management (2016) 2233-2253. doi: 10.1080/09585192.2015.1091980

[20] J. Schwanholz, T. Graham, Digital Transformation: New Opportunities and Challenges for Democracy?, Springer, Cham, 2018, 1-7, doi: https://doi.org/10.1007/978-3-319-61708-4_1

[21] L. Cortellazzo, E. Bruni, R. Zampieri, The role of leadership in a digitalized world, A review. Frontiers in psychology (2019) 1-21. doi: https://doi.org/10.3389/fpsyg.2019.01938

[22] J. P. Sankar, G. Yogan, Human Resource Digital Transformation of IT Sector in India. Webology (2021) 219-232. doi: 10.14704/WEB/V18I1/WEB18085

[23] G. Vial, Understanding digital transformation: a review and a research agenda. Journal of Strategic Information System (2019) 118-144. doi: 10.1016/j.jsis.2019.01.003

[24] G. Tortorella, A. M. Vergara, J. A. Garza-Reyes, R. Sawhney, Organizational learning paths based upon industry 4.0 adoption: an empirical study with Brazilian manufacturers. International Journal of Production Economics (2020) 284-294. doi: https://doi.org/10.1016/j.ijpe.2019.06.023

[25] J. Goldstein, How digital Technology Demand is Transforming HR. Workforce Solutions Review (2015) 28-29. URL: https://www.shrm.org/resourcesandtools/hr-topics/technology/pages/how-digital-technology-demand-is-transforming-hr.aspx

[26] Y. Yusoff, M. Nejati, D. M. Kee, A. Amran, Linking green human resource management practices to environmental performance in hotel industry. Global Business Review (2018) 10-14. doi: https://doi.org/10.1177/0972150918779

[27] S. Stepanov, D. Paspalj, M. Butulija, MOTIVACIJA ZAPOSLENIH KAO KLJUČNI POKAZATELJ ZADOVOLJSTVA POSLOM. EKONOMIJA TEORIJA I PRAKSA (2016) 57-61. URL: http://scindeks.ceon.rs/article.aspx?artid=2217-54581602053S

[28] Branca, Fornai, Colla, Murri, Streppa, Schröder, The challenge of digitalization in the steel sector. Metals (2020) 288-311. doi: https://doi.org/10.3390/met10020288

[29] V. Buha, R. Lečić, M. Bjegović, M. Gajić, Ekonomska analiza potreba za usavršavanjem ljudskih resursa. Tehnika (2015) 337–345. URL: https://scindeks.ceon.rs/article.aspx?artid=0040-21761502337B

[30] V. Gerasimenko, O. Molchanova, Global pressure or national identity: strategic challenges for Russian business schools. Serbian Journal of Management (2017) 133-141. doi: https://doi.org/10.5937/sjm12-12048

[31] D. Kurek, Use of Modern IT Solutions in the HRM Activities: Process Automation and Digital Employer Branding. European Research Studies (2021) 152-170. doi: 10.35808/ersj/2035

[32] J. Odegov, E. Malakhova, A. Garnov, Development of HR processes and the use of digital tools in Russian companies. in:International Conference on Communicative Strategies of Information Society (CSIS 2018), Holandia, 2019, doi: 10.2991/csis-18.2019.19

[33] G. Elia, A. Margherita, G. Passiante, Digital entrepreneurship ecosystem: How digital technologies and collective intelligence are reshaping the entrepreneurial process. Technological Forecasting and Social Change (2020) 119-791. doi: https://doi.org/10.1016/j.techfore.2019.119791

[34] C. Zehir, T. Karaboğa, D. Başar, The Transformation of Human Resource Management and Its Impact on Overall Business Performance: Big Data Analytics and AI Technologies in Strategic HRM. in:Contributions to Management Science, Springer, Cham, 2020, 25-30. doi: 10.1007/978-3-030-29739-8_12

[35] M. Götz, B. Jankowska, Adoption of Industry 4.0 Technologies and Company Competitiveness: Case Studies from a Post-Transition Economy. Foresight and STI Governance (2020) 61-78. doi: 10.17323/2500-2597.2020.4.61.78

[36] O. Valdez-De-Leon, How to Develop a Digital Ecosystem – a Practical Framework. Technology Innovation Management Review (2019) 43-54. doi: 10.22215/timreview/1260

[37] A. Hanandeh, M. Hajij, The impact of artificial intelligence, big data analytics and business intelligence on transforming capability and digital transformation in Jordanian telecommunication firms. International Journal of Data and Network Science (2022) 727–732. doi: 10.5267/j.ijdns.2022.3.009

An Integrated System for Efficient Student Attendance Management

Piroška Stanić Molcer[1], Robert Pinter[1], Sanja Maravić Čisar[1] and Zlatko Čović[1]

[1] *Subotica Tech-College of Applied Sciences, Subotica, Serbia*

piroska@vts.su.ac.rs; pinter.robert@ vts.su.ac.rs; sanjam@ vts.su.ac.rs; chole@ vts.su.ac.rs

Abstract:
The usual way to record students' attendance by using Excel or application sheets is a slow process. The attendance data is usually used primarily as a bonus-based motivation tool. In this paper, the authors present the possibility of performing counting by those who enter the classroom by real-time video and by using a trained neural network. The study enhances an attendance recording system employing Arduino, and Raspberry Pi, aiming to validate attendance data through an alternative method using a camera-based system. The paper provides experimental results and operational insights, demonstrating potential for accurate attendance tracking.

Keywords:
recording attendance, neural network, top of head detection, Raspberry Pi, Arduino

1. Introduction

Tracking attendance is a crucial activity seen in many areas of life: employee attendance has to be monitored at the workplace, presence at classes is vital on all levels of educational institutions, from elementary school to university. The teacher needs to be able to trace the classroom presence as well as the activity of any given student, for any given class. If a particular class is attended by a large number of students, the teacher would generally resort to the age-old method of roll call and marking attendance on paper. However, this is both tedious and time-consuming, thus a novel method of tracking attendance has to be introduced.

While academic freedom at college and university institutions ensures the students' rights to miss a certain number of classes during the active semester, attendance at class still remains one of the crucial factors for students' learning achievement.

Available literature in the field of the relationship between student attendance and exam achievement indicates that there is some empirical evidence pointing towards the fact that attendance likely is a determining factor in academic performance and progression [1, 2, 3, 4].

Using real-time video for counting people is not a rare task. In study [5] the authors presented a cloud-based people counting system employing a Raspberry Pi embedded system integrated with OpenCV and Python. The methodology encompassed three crucial phases: the development of a people counting algorithm, integration into Raspberry Pi, and integration of data on ThingSpeak. The system utilized a Pi camera for real-time video capture and monitoring, while the Raspberry Pi served as the microcontroller for video processing and subsequent people counting. The resultant data was seamlessly transmitted to ThingSpeak, an IoT platform, for visualization. The authors performed experiments to evaluate the system's performance under various real-world scenarios, including live video footage and saved video, subsequently visualizing the data on the ThingSpeak platform.

In [6] the authors proposed a real-time human detection and flow estimation method utilizing depth images captured by a top-view TOF camera. The algorithm comprised distinct stages, including head detection based on local pooling and searching, classification refinement based on human morphological features, and tracking assignment filter based on dynamic multi-dimensional features. To validate their approach, the authors established a depth image dataset with over 10,000 entries and departure events, complete with detailed human location annotations. The proposed algorithm exhibited high-accuracy human detection and people counting, achieving an accuracy of 97.73%. Furthermore, it

demonstrated robust performance in complex scenarios, including fast walking, occlusion, and crowded scenes.

The study conducted by the authors in [7] emphasized automatic human detection within office environments, highlighting the technology's critical applications in security, healthcare, and service industries. They addressed challenges arising from changing geometric shapes due to body posture and partial loss of depth information caused by occlusions and the absorption of infrared light. The authors conducted extensive experiments using actual office data spanning 100 hours (equivalent to 10 workdays) to validate the effectiveness of their proposed algorithm. The results indicated a high accuracy of 97.7% in human detection. Two innovative ideas were proposed: a feature set describing the upper-back shape of humans, incorporating roundness and size of a height-continuous region, and an adaptive feature adjustment algorithm utilizing implicitly included information in the missing region.

The approach introduced by the authors in [8] leverages top-view camera data and oriented trajectories for effective audience measurement. It is applicable in various environments, offering robust, cost-effective, and adaptable solutions. The study reviews state-of-the-art audience measurement techniques and defines key concepts. The method involves head detection, orientation estimation, and trajectory computation. The experiments demonstrate its effectiveness in relative attention distribution.

In [9] the authors introduce a cost-effective people counting system for retail analytics using edge AI. It leverages real-time processing at the edge, reducing reliance on cloud services. The system accurately counts people and calculates conversion rates, providing valuable insights for retail optimization. It offers a promising solution for efficient retail analytics.

The authors of [10] propose a low-cost bidirectional people counter device using infrared sensors and a microcontroller. It accurately counts people in indoor spaces. The device's implementation is straightforward and cost-effective. Experimental results demonstrate high accuracy and real-time performance, making it suitable for applications in access control, public transportation planning, and security systems.

This paper is organized as follows. In Section 2 we introduce attendance recording system to enhance attendance recording through real-time top-view camera footage and deep learning-based head detection and tracking. Section 3 discusses the training process of the neural network and evaluates its performance metrics. Finally, Section 4 draws conclusions about the effectiveness and potential applications of our system for accurate attendance recording in educational and professional settings.

2. Research motivation

In this work, we present the enhancement of a custom-developed attendance recording system. The existing system is described in [11]. Using a mobile phone and an Arduino microcomputer system, the existing system can identify individuals present in the room through a Bluetooth connection and communication. The system enhancement corresponds to the verification of the recorded data (number of the students). Its aim is to count the attendees using an alternative method. In this experiment the authors have installed a camera above the entrance door, and then a Raspberry Pi system processes the streamed video in real-time and determines how many people have entered the room. In theory, the attendance values obtained by the two methods should be identical. The model of the system is shown in Figure 1.

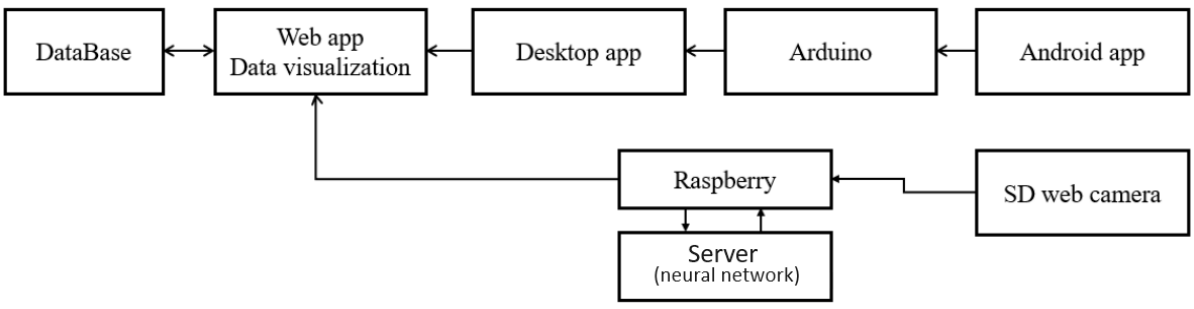


Figure 1: The schema of the attendance recording system

The placement of the camera and the system is shown in Figure 2. In the next section we will focus on the image processing and shape recognition system running on the Raspberry Pi.

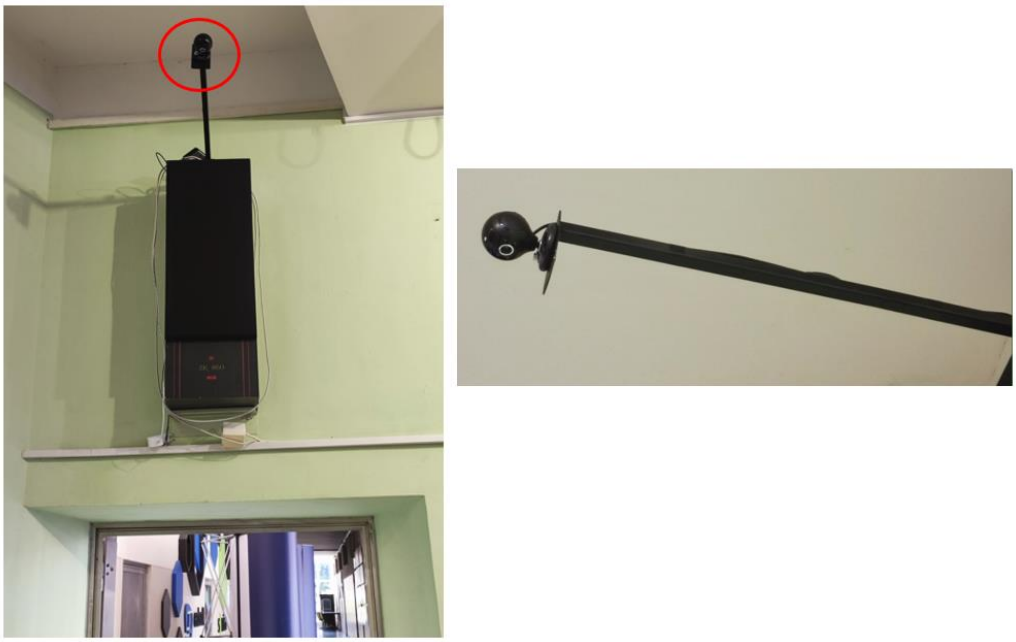


Figure 2: The mounted system above the entrance and the arm which holds the camera system

3. Presenting the experiment

We collected the necessary data for training the system in June 2023. Using the camera attached to the Raspberry Pi 4 system, we recorded footage of over 100 students individually entering the room. Recordings were also made when they exited the room. To determine the applicability of the method, it is important to note that the lighting conditions and the students' attire will be different in the fall or winter. Therefore, after the initial data processing and evaluation, we will conduct data collection specifically focused on these condition changes. Next paragraph introduces the decision-making part running on the server.

The decision-making part of the people counting software consists of two parts: the head detector and the head tracker. The detector is a neural network trained by deep learning. Images in the data set intended for teaching must be manually annotated, that is, mark where a head is visible in the image. At this step, questions arise as to whether the semi-visible heads should be annotated or not. This choice later plays a role in determining the quality of the detector. It is possible that a partially visible head is not annotated, but the detector recognizes the head. In this case the confusion matrix will report that there is no head in the image, but the net saw a head, i.e., a plus False Positive case will be added. However, the truth is that the net was not mistaken, and the head was indeed there. Such a case can be seen in the photo collections Figure 3 and Figure 4. Figure 3 shows the manually annotated heads, and Figure 4 shows the head locations detected by the neural network. Although with little confidence, the model also found heads that were not annotated by humans, and there really were heads there.

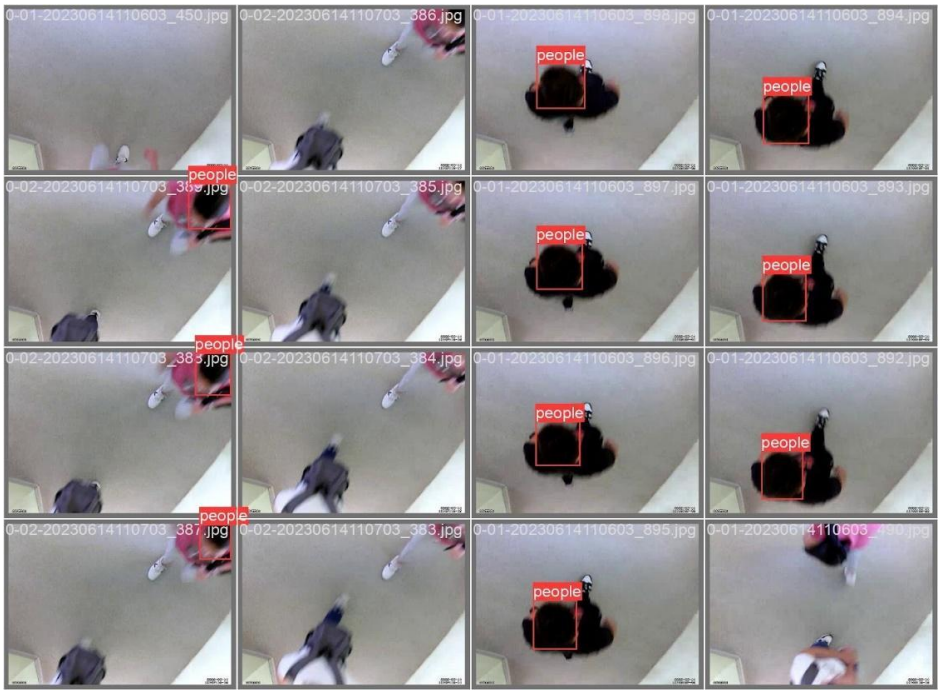


Figure 3: Test set annotations

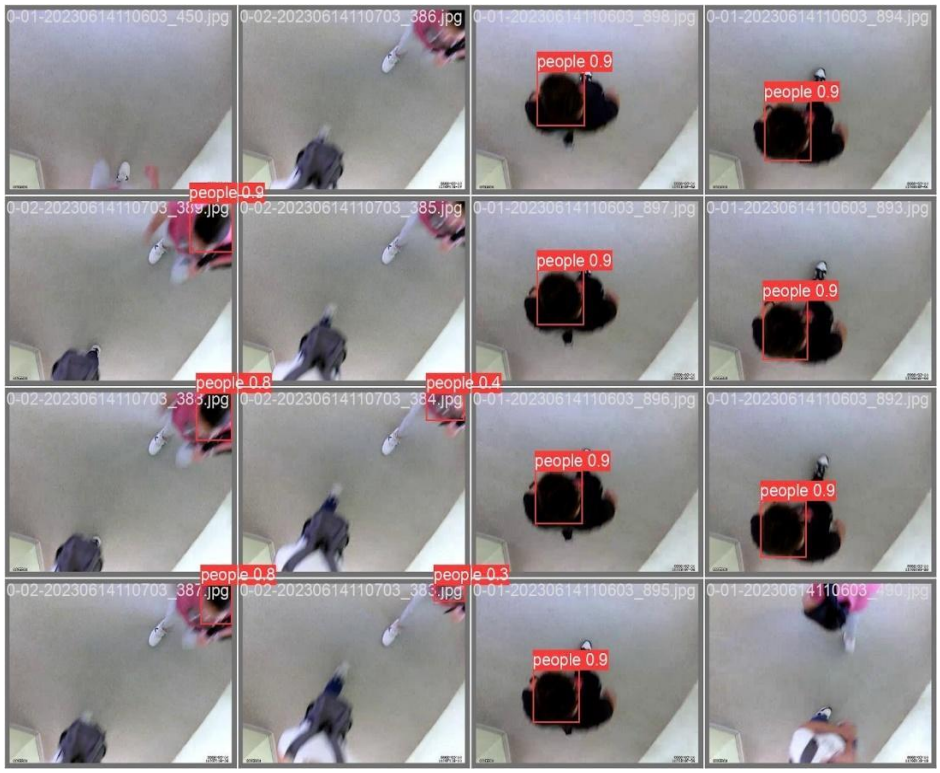


Figure 4: Test set - boxes given by the model with the adherent confidences

The output of the detector is the input of the head tracker. The head tracking part is not based on artificial intelligence, it just uses its results, so its output depends indirectly on the output of the neural network. This relationship is not simple, because if the model is wrong and does not detect a head where there is a head, it does matter what location the head was. Since the goal is to determine how many people entered and exited the room, a selector in the form of a dividing line between outside and inside is needed. If the net is wrong near the dividing line, the counter will also make a mistake, but if the missed head is far enough from the dividing line and is detected again in one of the following frames, the counter will notice that it has crossed the dividing line.

The classification loss shows the dissimilarity between the predicted class probabilities and the true class labels for each bounding box. Distribution focal loss ensures that the model prioritizes learning rare classes helping to improve overall model performance on imbalanced datasets with varying class distributions. The bounding box loss shows the difference of the predicted bounding box coordinates and the original box coming from annotation. During training, the goal is to change the model parameters to minimize the average of these losses.

For the head detection a deep neural network was used, namely yolov8 which is a pre-trained model. The further training with the collected and preprocessed dataset took 20 epochs. After 20 epochs the loss curves of the validation data set no longer improved. The decreasing losses during training for the train and the validation data set can be seen in Figure 5.

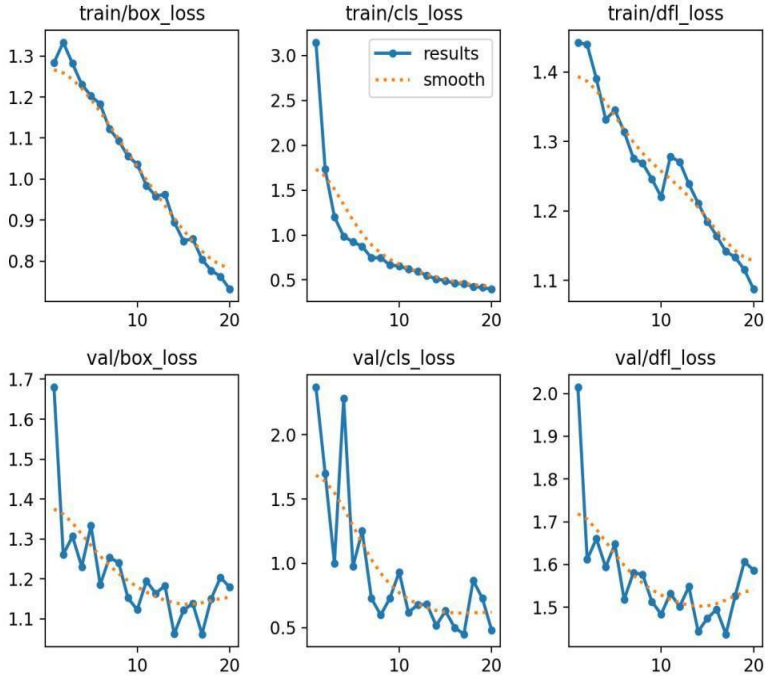


Figure 5: Losses as a functions of teaching epochs

The performance metrics of the model depend on the confidence level. Figure 6 shows how the recall value changes with confidence, so we can choose the appropriate threshold value for the application.

The recall value depends on the number of correctly predicted heads (True Positives) and the number of unrecognized heads (False Negatives):

Recall = True Positives / (True Positives + False Negatives)

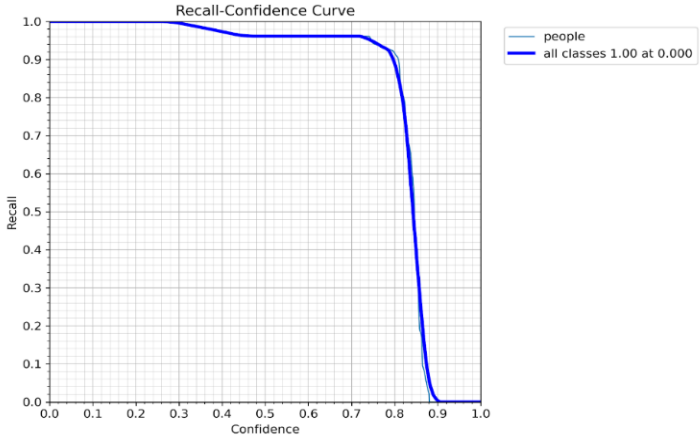


Figure 6: The recall performance value of the head detector in the function of confidence

Figure 7 shows how the value of precision changes as a function of confidence. Precision is a performance metric that, in addition to well-predicted heads (True Positives), also takes into account cases in which the model predicted a head where it was not marked (False Positives):

Recall = True Positives / (True Positives + False Positives)

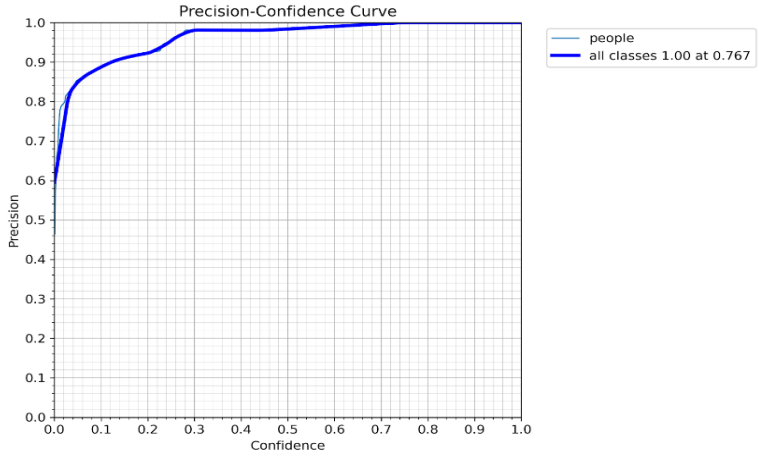


Figure 7: The precision performance value of the head detector in the function of confidence

The F1 score takes into account both performance metrics:

F1 = 2*Precision * Recall / (Precision + Recall)

Figure 8 shows the value of the F1 score as a function of confidence.

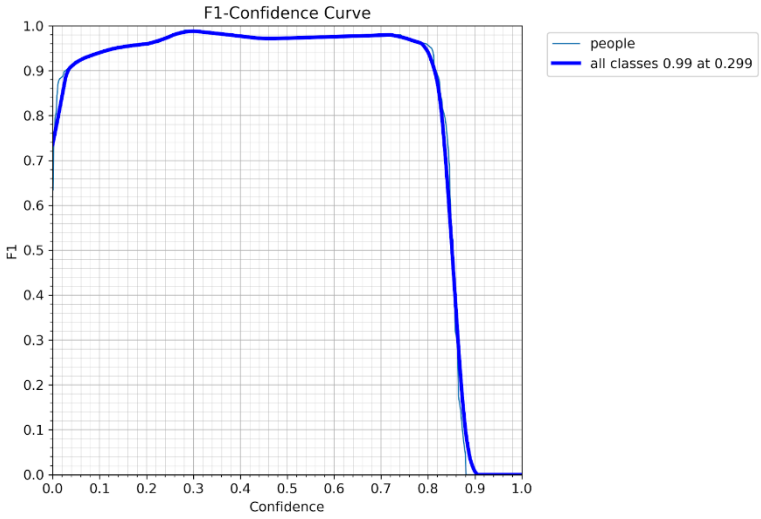


Figure 8: The F1 score of the head detector in the function of confidence

In addition to the heads found, the value of recall also depends on how many times the model failed to detect heads. The value of precision, on the other hand, depends on the number of times the model found a head where there was none. It should be noted here that the model also found the partially visible heads that were not annotated, and this should not actually be considered an error, but it still has an impact on the precision score. Nevertheless, it is most appropriate to work with the best confidence value shown according to the F1 score, since F1 is the one that takes both types of errors into account. Based on the F1 curve, the best confidence is 0.29, at which the F1 value is maximal, in our case 0.99.

4. Conclusions

The work presented here is part of the goal of creating an easy-to-use and affordable tool that would help track student attendance at lectures. The hardware part of the device is a Raspberry Pi and a camera that films the top of the heads. The software, the head detector, was created using a deep learning based model. Since the performance of the people counter largely depends on the performance of the head detector, in the work we illustrated the operational results of the head detector. Recall, precision and F1 score were measured at different output confidence levels.

Acknowledgment:
This work was created within the framework of the project under the title "Development of a real-time presence recording system using camera and machine learning", and number 142-451-2371/2023-02/2. The project is financed by The Provincial Secretariat for Higher Education and Scientific Research of Vojvodina.

References:

[1] G. C. Durden and L. V. Ellis: The Effects of Attendance on Student Learning in Principles of Economics, The American Economic Review, Vol. 85, No. 2, Papers and Proceedings of the Hundredth and Seventh Annual Meeting of the American Economic Association Washington, DC, January 6-8, 1995 (May, 1995), pp. 343-346
[2] D. O. Allen, D. J. Webber: Attendance and Exam Performance at University: a Case Study, Research in Post-Compulsory Education, Published by Taylor & Francis, 2010, Online ISSN: 1747-5112, Print ISSN: 1359-6748
[3] A. S. N. Kim, S. Shakory, A. Azad, C. Popovic, L. Park: Understanding the impact of attendance and participation on academic achievement. Scholarship of Teaching and Learning in Psychology, 2019, Published by American Psychological Association, Online ISSN: 2332-211X, Print ISSN: 2332-2101
[4] Why Is School Attendance Important? The Effects of Chronic Absenteeism, January 14, 2021, https://soeonline.american.edu/blog/importance-of-school-attendance/, visited in September 2023
[5] A. K. Mahamad, S. Saon, H. Hashim, M. A. Ahmadon, S. Yamaguchi, Cloud-based people counter, Bulletin of Electrical Engineering and Informatics, Vol. 9, No. 1, February 2020, pp. 284-291, SSN: 2302-9285, DOI: 10.11591/eei.v9i1.1849
[6] W. Wang, P. Liu, R. Ying, J. Wang, J. Qian, J. Jia, J. Gao: A High-Computational Efficiency Human Detection and Flow Estimation Method Based on TOF Measurements. Sensors (Basel). 2019 Feb 11;19(3):729. doi: 10.3390/s19030729. PMID: 30754685; PMCID: PMC6387275.
[7] Y. Fujimoto, K. Fujita: Depth-Based Human Detection Considering Postural Diversity and Depth Missing in Office Environment, ISSN: 2169-3536, DOI: 10.1109/ACCESS.2019.2892197
[8] R. Nakatani, D. Kouno, K. Shimada and T. Endo: A Person Identification Method Using a Top-View Head Image from an Overhead Camera." J. Adv. Comput. Intell. Intell. Informatics 16 (2012): 696-703.
[9] K.R. Kanjula, V. V. Reddy, K.P. Jnanesh, J. S. Abraham and K Tanuja: People counting system for retail analytics using edge AI. ArXiv abs/2205.13020 (2022): pp. 1-5.
[10] R.L. dos Santos, H.C. de Oliveira, M.C. de Almeida, et al. A Low-Cost Bidirectional People Counter Device for Assisting Social Distancing Monitoring for COVID-19. J Control Autom Electr Syst 33, 1148–1160 (2022). https://doi.org/10.1007/s40313-022-00916-z
[11] R. Pinter, S. Maravić Čisar, Z. Balogh, H. Manojlović: Enhancing higher education student class attendance through Gamification, Acta Polytechnica Hungarica, Journal of Applied Sciences, 2020, Vol. 17, No. 2, pp. 13-33, ISSN 1785-8860

EasyLoanDecision: A Expert System for Consumer loan

Natasha Blazheska-Tabakovska[1], Lijeta Hodja[1], Igor Nedelkovski[1], Mimoza Bogdanoska Jovanovska[1] , Marina Blazekovic-Toshevski[1]

[1] *University "St.Kliment Ohridski" – Bitola,7000, Partizanska bb, R. of North Macedonia*

natasa.tabakovska@uklo.edu.mk; lijeta_123@hotmail.com; igor.nedelkovski@uklo.edu.mk, mimoza.jovanovska@uklo.edu.mk, marina.blazekovic@uklo.edu.mk

Abstract:
The banks, due to the large supply and demand for loans, have a great need for a more efficient and effective way of working, and at the same time to offer high quality to their customers. Banks in R. Macedonia uses information technologies in their operations, but there is room and opportunities for their application on a larger scale if mistrust is reduced and the awareness of employees and authorities is increased. Expert systems are an appropriate way to support lending decisions, as they use expert knowledge in decision-making. Due to their interface, these systems are easy to use and do not require much credit expertise from their users. In this paper we proposed an expert system, for supporting the decision-making process for the application for consumer credit.

Keywords:
expert system, knowledge-based system, banking sector, consumer loan

1. Introduction

Banks are the most important participants in the financial system and the largest financial institutions, measured by their participation in the total assets of the financial sector. They are the main pillar of the financial sector in the country, contribute to the distribution of liquidity in the economic sector, and indirectly stimulate the economy of a country through lending. A bank's main activity is the collection of deposits and the granting of loans to make a profit, as well as the execution of payment transactions. Lending loan to anyone could be a risky activity for the banks. Therefore**,** there should be an appropriate method through which the risk of lending loan to any individual can be evaluated. In credit assessment of customers number of factor are considered like sex, age, education, marital status, career, no return check, deposit account, deposit period (month), average salary [1].
Also, the application of IT improves processes in many segments of operations. Expert knowledge is a combination of a theoretical understanding of the problem and a collection of heuristic problem-solving rules that experience has shown to be effective in the domain [2]. So, the expert systems dramatically expand computer efficiency, utility, and range in financial sector. Processes such as credit decision-making, but also other activities such as various transactions, expert system facilitate and contribute to the performance of those tasks with greater accuracy, speed, and without bias. The primary goal of the expert systems is to make expertise available to decision makers and technicians.
The paper proposes an expert system to support decision-making for consumer credit applications. The system will assist the loan officer in risk assessment and completion of the loan package and application of the proposed expert systems can improve efficiency in credit decision-making. The paper is organized as follows. The second section depicts used research methodology, upgraded knowledge of important topics, and gives a short overview of realized research. The following section describes the proposed expert system EasyLoanDecision, for supporting the decision-making process for approval/rejection the application for consumer credit. Finally, the last section provides concluding remarks.

2. Background

2.1. Function of banks in the financial sector

Banks have a central role in the economic system, because they perform certain activities, which are of vital importance for the functioning of the economy as: financial intermediaries to customers and partners; bearers in payment transactions in the payment sector; have a special role in monetary policy; and offer some additional or special services to economic agents [3]. According to the Macedonian Law on Banks "bank is a legal entity with permission from the governor of the National Bank of Macedonia established in accordance with the provisions of this law, whose main activity is collecting deposits and other returnable sources of funds from the public and granting loans in its own name and for own account" [4]. However, the legal definition does not cover the non-banking activities that banks perform in large numbers. According to the functional approach, banks are institutions that offer clients various banking services such as: financial intermediation between clients and partners; they are carriers in the payment turnover in the payment sector; have a special role in monetary policy; and offer some additional or special services to economic agents [3]. The functions performed by a bank can be divided into three categories: active, passive, and neutral banking operations. Our focus is on active banking operation, the operations in which the bank is a creditor of its client, for example issuing a loan. The word loan ("credere"), which means trust, refers to legal certainty. It is a private legal act by which an economic entity, for a certain period, transfers certain real economic goods or money into the ownership of another economic entity, with the obligation to return them with a certain compensation (interest) [5].

A natural or legal entity can apply as a borrower. Loans to the population represent financial assets that banks, and other financial institutions lend to individuals and households. However, according to the purpose, there are several types of loans, which are intended for financing current consumption or for the purchase of durable products, such as: Consumer loans (dedicated consumer loans or non-purpose consumer loans, revolving loans, car loans, housing loans, etc.

The procedure for concluding a loan agreement takes place in several stages: it starts with submitting a request from the client to the bank; the client submits the necessary/requested documentation together with the request; verification of submitted documents, which include: personal identification document, employment certificate and salary certificate, certified by the employer, latest invoices from overhead expenses confirming that they are regularly paid; based on the attached documents, the bank analyses financial aspects, evaluates the creditworthiness of the client and collateral analysis (possibility of administrative ban on salary or guarantors, or secured deposit or pledge of securities or real estate mortgage); if the assessment is positive, the next phase is reached - concluding a loan agreement; monitors the transfer of the credit and enabling the borrower to use it; the last stage is the repayment of the loan by the borrower [6].

2.2. Expert system in banking sector

Expert systems are knowledge-based systems that attempt to act as experts in a problem domain. They are intelligent computer programs that mimic the way experts solve problems, make decisions, or perform a task that requires domain expertise, using both facts and heuristics [7]. The expert system is based on the paradigm Expert + Knowledge = Advice. From the user's point of view, the expert system acts as an intelligent consultant in a specific area.

Goodwin and Wright point out that two types of expert systems can be distinguished: The first are basically academic research projects where difficult or potentially intractable problems are solved so that new ways of representation or extraction must be developed of knowledge; The second set of systems are those built by consultants using commercially developed expert system shells. They are easy to program in the same way that word processing or spreadsheet programs provide easy-to-use tools [8]. Based on the problem they address and the area in which they are applied Several generic categories of expert systems can be defied: Classification Systems, Diagnostic Systems, Monitoring Systems, Design Process Control Systems, etc. [9], [10].

The banking industry has a wide variety of business lines and needs urgent management of its expertise. For banking industry more appropriate is the second set of system, according to Goodwin and Wright. They need to make expertise cheap and accessible and ensure uniformity in decision-making across the organization. The literature suggests that expert systems have a major role and application in the banking sector in the US, Europe, and Japan. The list of applications for expert systems is extensive: Credit analysis for commercial and financial loans; Risk analysis and loan clearance, including loan monitoring; Client Profile and Investment Advisory; Portfolio management; Risk analysis for securities offerings; Analysis of securities and evaluation of investment exposure; Estate planning and tax consulting; Financial planning assistance, including resource allocation; Support for foreign exchange trading and analysis of foreign exchange exposure, etc. [8].

The credit risk estimation problem is a very challenging and important financial analysis problem, and research shows that expert systems perform very well for this complex and unstructured problem compared to more traditional statistical approaches. Research on the credit risk of banks in different countries, regardless of whether they are developing or developed, shows that in many countries exposed to financial crises, financial defaults on overdue loans have a great impact [2].

Expert systems solve the problem of credit risk assessment and credit package completion by applying expert knowledge and used as a decision support system, making expertise available to decision makers. The expert knowledge embedded in expert systems is a combination of a theoretical understanding of the problem and a set of heuristic rules for solving problems that experience has shown to be effective in the domain. Expert systems covering the field of lending have the ability to recognize a typical task in a particular problem they are solving. They also possess some qualities, which are due to heuristic knowledge. Based on this knowledge, they can recognize the fastest way to reach a solution, as well as the right approach in solving problems, even if the data is incomplete [9].

For the development of the Expert System for evaluation and support of credit decisions, two activities are crucial: acquiring knowledge for building the knowledge base in the development environment; and integrating the knowledge base with the user interface in the consultation environment. In the development environment, the knowledge engineer first analyzes all the available documents related to the subject area which constitutes indirect knowledge acquisition, and then through knowledge acquisition sessions, he extracts knowledge from experts by directly asking them how they do their work in the domain. The extracted knowledge is checked and updated by experts in the domain and then that knowledge is incorporated into the system [7]. One of the most challenging activities in building expert systems is the representation of knowledge in the system [13].

C Language Integrated Production System (CLIPS) is a rules-based programming language and is a low-cost option for developing and delivering expert systems across a wide range of hardware platforms as well as other programs where heuristics are easier to implement and maintain than algorithmic solutions [14]. However, our pick for expert system's development tool is Exsys Corvid because it is designed to enable experts to develop powerful, interactive advisory applications quickly and easily through a powerful development environment that can be quickly learned and implemented. This tool allows converting complex decision-making processes into an interactive form that can be easily embedded in a website [15]. The steps in building the system are very similar to an expert's explanation of how a decision is made. A rule-based system can be simply created by using a set of assertions and a set of rules that specify how to act on the assertion set. Rules are expressed as a set of If-Then statements (called If-Then rules) [16]. By applying some of (or all) seven different types of variables, of which the most used are: Static List, Numeric and Confidence; and logical blocks that allow to define, organize, and structure information for decision-making in logically connected blocks, is very easily developed the expert system. The inference engine of the expert system, imitates the human brain and intuitively combines heuristics, combining individual rules to solve larger problems, makes a combination of solutions, analyses all the answers received, as well as all the data from each point of view to get best solution.

3. Research Methodology

The first step, before starting the development of an expert system, was to collect and analyses data about the situation in the banking sector in the Republic of Macedonia and the need for such a system.

For this purpose, 3 questionnaires were used as research instruments: Questionnaire number 1 and Questionnaire number 2 consisting of 10 questions for employs in the banking and financial sector. The third survey questionnaire covered 150 respondents who are direct or indirect applicants for loans on the market, persons who have already used loans, use, or intend to use loans from banks and other financial institutions in the Republic of Macedonia.

The research was conducted over a period of 5 months and cover 6 different cities in the southwestern region of the Republic, with effort not to select banks or savings banks based on a certain criteria. In the research process, 6 banks and 6 financial companies that function and work on the financial market in the Republic of Macedonia were included. The sample is representative and geographically distributed across several cities in the country, so the conclusions can be generalized.

The data, collected by the research instrument, after their systematization and grouping, were statistically processed.

After the analysis of the collected data and the obtained presentation of the situation in the banking sector in the Republic of Macedonia in the part of the application of IT during the processing of credit requests, we moved to acquiring knowledge for building a knowledge base for credit decision. Then we followed the selection of a tool for development of an expert system and finally we moved on to develop an expert system for consumer loan.

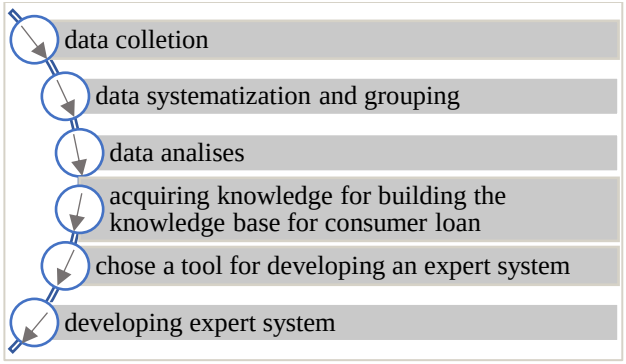


Figure 1: Step of expert system development for consumer loan

4. Expert system for consumer credit: EasyLoanDecision

The conducted analysis of the collected data gave an overview of the conditions in the banking sector in the Republic of Macedonia in the area of IT application, credit processing time, the need for staff to work on credit requests, the application of expert systems in the process of processing credit requests, etc.

In general, the largest percentage of the respondents agree that the established written credit policy allows them to be unified in lending and that in this way the bank will ensure that the regulatory standards are met. Respondents (over 90%) fully agree that banks should aim to achieve profit with the lowest possible risk through the marketing of products and services.

According to a large part of the respondents, consumer loans are offered mostly by banks in the Republic of Macedonia. IT is used in the process of making decisions on credit approval in the banking sector. However, a large part of the respondents stated that they are not sufficiently familiar with the functionality and functioning of expert systems in the lending process.

Regarding the duration of the process from application to receiving an answer for the requested loan, they stated that it lasts from 10-30 days. A decision according to the respondents, is mainly made by the official with the help of computer tools. Half of the respondents answered that the training for one person takes between 9-18 months, which is quite a long period considering that 90% of the respondents

answered that it is difficult to find a person with adequate knowledge. Almost all respondents stated that in the long term, it is necessary to use expert systems when making a decision to approve or reject a loan request.

Citizens and legal entities in the Republic of Macedonia are interested in loans, and the majority of respondents answered that they are currently consumers of credit, but as the biggest problem they point to the high interest rates (more than 60% of the respondents) and the lending conditions that offered by the banks, especially the long procedure for the loan and the duration of the loan approval.

To facilitate and speed up the lending procedure, this paper proposes an expert system for supporting the decision-making process for approval or rejection of the application for consumer credit.

In addition, based on the detailed analysis of the procedure for obtaining a consumer loan, the various tools for developing an expert system were considered. An Exsys Corvid shell was chosen.

The result of realized research and upgraded knowledge is Expert system EasyLoanDecision. The decision tree for Expert system EasyLoanDecision is given on Figure 2.

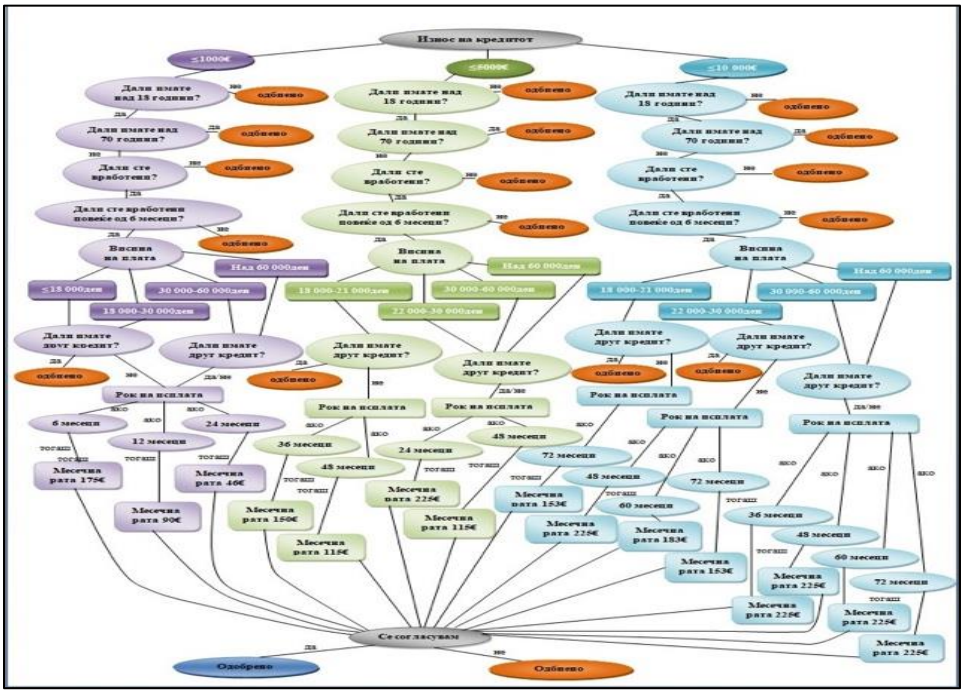


Figure 2. Decision tree for a decision making for a consumer loan.

The decision for a positive/negative response to the loan application depends on several parameters such as: the age of the applicant, the work status (employed/unemployed), the duration of the work status as an employee, the amount of personal income, other debts (loans) as well as the number of loan repayment installments (Figure 2).

Two type of variables was used in ExsysCorvid for EasyLoanDecision: Static list, Numeric value, and Confidence value. Total number of used variables is 17. The decision-making logic is described and built using nodes that represent the statements in the If..Then rule (Figure 3).

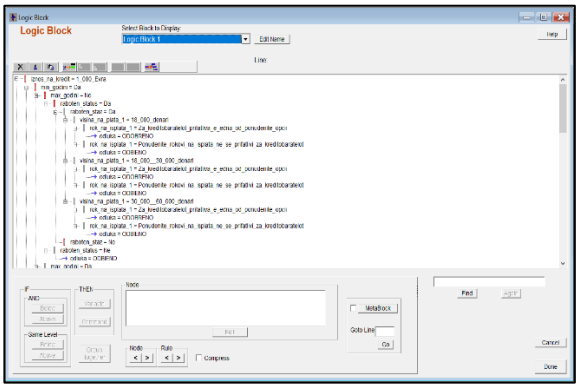


Figure 3. The logic block of EasyLoanDecision.

The command block controls the system. Unlike, the logic block in a system tells it HOW to do something and the way the activities should be carried out, the command block tells it WHAT to do and WHEN the given activities are to be carried out. The EasyCreditDecision based on the built-in logic (If..Then rules) and the answers given by the credit analysts to the questions asked by the expert system for the specific credit request, are helping the decision making of the request (Figure 4 and Figure 5).

Figure 4. Communication between ES and credit analysts

Figure 5. Expert system recommendation

5. Conclusions

The conducted research shows that the citizens and the legal entities in the Republic of Macedonia are interested in loans, and many respondents answered that they are currently consumers of credit. Mostly, the loan's users are not satisfied of the lending conditions that offered by the banks and the big procedure for the loan, as well the duration of the loan approval.

The professional staff of the bank should make a good analysis of each borrower and evaluate whether the potential borrower is a good payer, i.e., that there is no risk (the risk is relatively small) that the borrower will not pay the principal and interest on the approved loan on time. Banks use IT in the decision-making process during the credit approval process, but do not use expert systems. There are not enough professionals on the market who are suitable for the position of considering requests and making credit decisions, and the training for one person to be qualified to perform that task is on average between 6 -10 months.

To facilitate and speed up the lending procedure, this paper proposes an expert system EasyLoanDecision for supporting the decision-making process for approval/rejection of the application for consumer credit. This expert system embeds the procedure for concluding a loan agreement in accordance with the established written credit policy. The application of an expert system in the process of processing the credit request will contribute to increase the effectiveness and efficiency of the banks in this segment of their operations.

Acknowledgment:

The authors want to acknowledge the expert who works as credit officer at one Macedonian bank for his support and explanation of the procedure for concluding a loan agreement.

References:

[1] A. Kaur, V. Madaan, P. Agrawal, R. Kaur and S. K. Singh, "Fuzzy Rule based Expert System for Evaluating Defaulter Risk in Banking," *Indian Journal of Science and Technology, ,* vol. Vol 9, no. 28, pp. 2-6, 2016.

[2] M. Mahmoud, N. Algad and A. Ali, "Expert System for Banking Credit Decision," in *International Conference on Computer Science and Information Technology*, 2008.

[3] G. Petrevski, Bank Management, Skopje: Faculty of Economics, 2008.

[4] "Law on Banks," *Official Gazette of the Republic of Macedonia,* no. 67/2007.

[5] V. Madzova, Credit Policy, Shtip: Gotse Delchev University, 2011, p. 12.

[6] C. I. Group, "Consulting Inter Group," Консалтинг Интер Груп, 2011. [Online]. Available: http://cig.com.mk/kreditirane-i-vidovi-kredit/. [Accessed 06 2023].

[7] N. Blazheska-Tabakovska, Business Information Systems, Bitola: Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" , 2022, pp.151-153.

[8] P. Goodwin and G. Wright, Decision Analysis for Management Judgment (4th ed.), Wiley, 2009.

[9] I. Nedelkovski, Engineering expert systems, Bitola: Technical Faculty, 2009.

[10] N. Singh and A. Agarwal, "A Comparative Study on Artificial Intelligence and Expert Systems," *The International journal of analytical and experimental modal analysis,* vol. 12, no. 7, p. 282–288, 2020.

[11] D. N. Chorafas and H. Steinmann, Expert Systems in Banking: A Guide for Senior Managers, Macmillan Academic and Professional LTD , 1991.

[12] J. Giarratano and G. Riley, Expert Systems: Principles and Programming, Thomson Course Technology, 2005.

[13] F. Puppe, Systematic Introduction to Expert Systems: Knowledge Representations and Problem-Solving., Berlin: Springer-Verlag., 2012.

[14] H. K. Jabbar and R. Z. Khan, "Tools of development of expert systems: A comparative study," in *3rd International Conference on Computing for Sustainable Global Development (INDIACom), (pp. 3947–3952.) IEEE*, New Delhi, India, 2016.

[15] E. Corvid, "Exsys® Corvid Knowledge Automation Expert System Software Developer's Guide," 2010.

[16] C. Grosan and A. Abraham, "Rule-Based Expert Systems. I," in *ntelligent Systems. Intelligent Systems Reference Library*, Berlin, Springer, Berlin, Heidelberg, 2011, p. 149–185.

Intellectual Capital and its Importance for an Entrepreneurial IT Company in the Period to Come

Bozidar Milenkovski[1], Sasho Nikolovski[2], Nikola Rendevski[3]

[1] *Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" –Bitola, Macedonia*

[2] *Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" –Bitola, Macedonia*

[3] *Faculty of Information and Communication Technologies, University "St. Kliment Ohridski" –Bitola, Macedonia*

bozidar.milenkovski@uklo.edu.mk, sasnik@gmail.com, nikola.rendevski@uklo.edu.mk

Abstract

The main forces today that initiate changes are globalization, the consequences of the health crisis caused by the Covid-19 pandemic, energy crisis, economic crisis, new technologies, increased competition, changes in consumer demands, etc. All these factors force entrepreneurial IT companies to adapt better, to respond faster to business challenges and proactively to shape the activities in which they operate.

Additionally, in the last few decades, there are large differences between the market and accounting values of the companies in general. Many authors believe that this difference is due to the value of the intangible assets of the companies i.e. their intellectual capital. Therefore, it is needed more detailed analysis of the impact of IT company's intellectual capital on its business performance because IT companies that use intellectual capital in their operations and at the same time practice an entrepreneurial way of managing processes (innovation, creativity, taking risks) are in a better position to maintain their current competitiveness as well as to develop a new one.

Hence, this paper aims to analyze the importance of intellectual capital for entrepreneurial IT companies in the period to come especially when the future seems more uncertain and unpredictable than ever before.

Keywords: intellectual capital, entrepreneurship, IT company, knowledge, competitiveness.

Introduction

The change, in general, has been a central theme for centuries. The ancient Greek philosopher Thales of Miletus argued that change is ubiquitous and inevitable and that it is one of the fundamental characteristics of the world. There is no other constant reality, except the reality of change (Latin: "Panta Rei"). Hence, success is achieved by one who changes faster and smarter. Today, knowledge is increasing at a tremendous rate and theoretically, people should perceive and understand things better and faster. But the exact opposite is happening. New knowledge leads to faster economic, social and political changes, and to understand what is happening, people accelerate the accumulation of knowledge, which, in turn, leads to faster and greater changes. This, above all, is mostly due to the digital revolution and the massive application of new information and communication technologies (ICT) in all spheres of social life. Robotics, Artificial Intelligence (AI), Nanotechnology, Biotechnology, Blockchain Technology, Internet of Things, Quantum Computers, 3D printing, 5G, Virtual Reality, Augmented Reality, Big Data etc. are all part of this revolution. In the period to come by 2030, it is expected that more than 500 billion devices (smartphones, drones, autonomous cars, home appliances, sensors, wearables) will be connected on Internet as a result of the concept "Internet of Things (IoT) [1]. As a consequence, people are less and less able to understand the present or to predict the future. For example, in 1023 it was relatively easy to predict how Europe would look like in 1050. Indeed, royal dynasties might change, unknown invaders might invade, and natural disasters might occur, yet it was clear that in 1050 Europe would still be ruled by kings and priests, that society would be agricultural, that most of the inhabitants will be fortified peasants who will suffer greatly from famine, diseases and wars. In contrast, in 2023, it is not known exactly what Europe will look like in 2050. Nobody knows what political systems there will be, how labor markets will be structured, what medicine will be like, or what health, economic, environmental, or energy crises will exist. Therefore, a more detailed analysis of intellectual capital and entrepreneurship is needed as concepts

that would help individuals, companies and societies to deal more easily with the large number of challenges in the coming period.

The relationship between intellectual capital and entrepreneurs

Uber, Airbnb, Oculus VR, WhatsApp, Tesla Motors are some of the startups that have managed to acquire a multi-billion dollar market value in just a few years (although it should be mentioned here that up to 90% of newly established startups fail over the long run) [2]. The speed of wealth acquisition for this new type of company is something without precedent in the business world. These companies rely on fast-growing and functionally sustainable exponential assets such as the large amount of data (Big Data), access to and seeking opinion from the "Community" or fast-growing technologies (Accelerating Technologies). Unlike linear growth, which is the result of adding a constant (arithmetic progression) and which applies to traditional companies, exponential growth is multiplication by a constant (geometric progression) and mostly applies to these exponential companies that operate in the IT sector. For example, imagine that you have to walk along a road with steps of one meter in length. As soon as you take 6 steps, you have progressed six meters (1, 2, 3, 4, 5, 6). After 24 steps you are already 30 meters from where you started. It's easy to predict where an extra 30 steps will take you. This is linear growth. However, let's imagine that, after each step is taken, we can double the length of our step. In this way, after taking 6 steps, we will actually advance 32 meters (1, 2, 4, 8, 16, 32), which is significantly more than the 6 meters we would have covered with an equal step length. Unbelievable, but after step number 30, doubling our step, it will take us further than a billion meters from where we started (2^{30} =1073741824 meters). This is actually the surprising power of accelerated exponential growth. Today, the future is not unfolding linearly, but exponentially, which makes the process of predicting the future of technological trends and business models significantly difficult. Hence, the need for a more detailed analysis of the intellectual capital of IT companies is imposed. But what is intellectual capital? The intellectual capital of an IT company represents its intangible assets as an important part of its total assets. This capital has a specific power creatively to turn the various types of knowledge, abilities, experiences, skills, technologies, etc. within an IT company into products that have real value. The essence of intellectual capital can be presented by its contribution to value creation and knowledge-based competitive advantage for an IT company. Mathematically, the simplest way to calculate a company's intellectual capital value is:

Company's intellectual capital value = Company's market value − Company's accounting value.

Moreover, the basic elements of intellectual capital are the human capital, the structural capital and the relational capital. Each of these components of intellectual capital contributes significantly to the success of an entrepreneurial IT company.

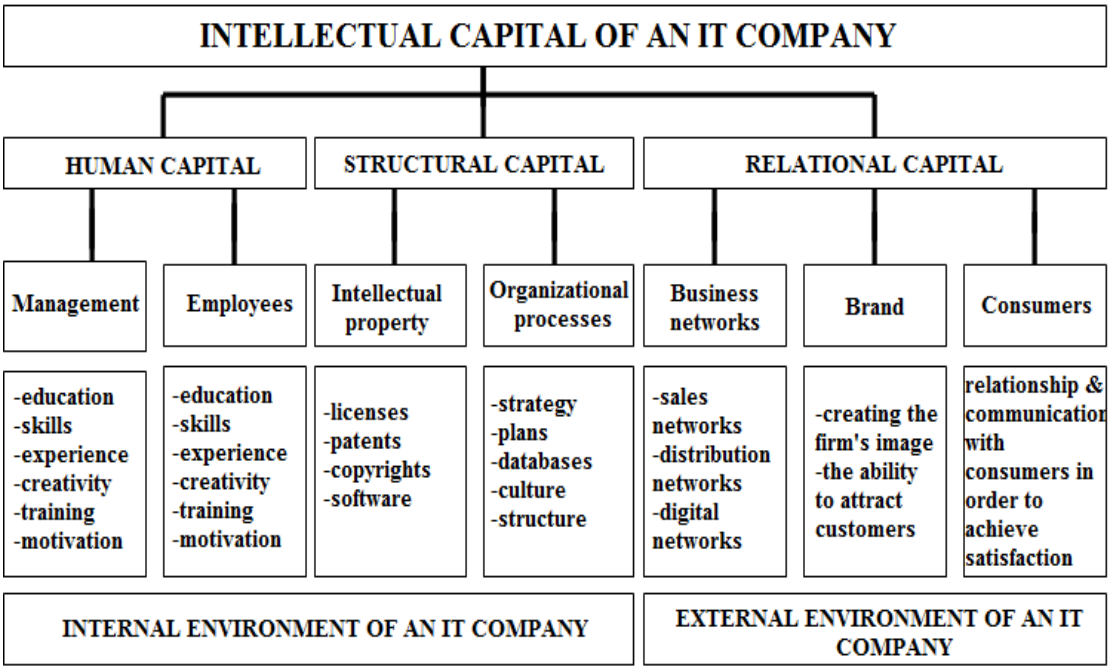


Figure 1. The intellectual capital structure of an IT company [3]

On the other hand, entrepreneurs are individuals who allocate economic resources from a level of lower productivity to a level of higher productivity. So for example, the human capital of the entrepreneur (education, business experience, skills, training, level of motivation), structural capital (the capacity of the company to adapt quickly to changes, organizational structure, organizational culture, the ability to implement successful organizational strategies) and of relational capital (development of productive business networks, quick access to key external stakeholders of the company, etc.) are important intangible resources that are related to the achievement of positive business performances. This shows the interrelation and mutual influence between entrepreneurs and the intellectual capital in their companies as well as their contribution to the success of the entrepreneurial IT company. In this context, a simple metaphor that can be used to explain the connection between entrepreneurs and intellectual capital is the one according to which the entrepreneur can be considered as a sort of "juggler", who juggles different parts of intellectual capital: experience, training, skills, education, networks, organizational structure, organizational culture. Such a metaphor implies interesting aspects. For example, a juggler, just like an entrepreneur, must have the necessary skills and abilities. The more skills he has, the more things he can juggle. Likewise, if an entrepreneur can "juggle" more aspects (parts) of intellectual capital, then he will be more successful than other entrepreneurs who know how to "juggle" fewer parts of intellectual capital [4].

Determinants of a successful entrepreneurial IT company in terms of its intellectual capital

Entrepreneurship is not exclusively bound to small or newly founded companies and the private sector, but it is also found in large companies, as well as in the public and non-governmental sectors. At the same time, it should be emphasized that, regardless of the size or the sector in which it operates, an entrepreneurial company is usually characterized by the ability to innovate, initiate changes and quickly respond to changes in the environment in a flexible manner.

There are three phases that an entrepreneurial company goes through:

(1) the ex-ante period, (2) the gestation period, and (3) the consolidation period. All this also applies to an entrepreneurial IT company [5].

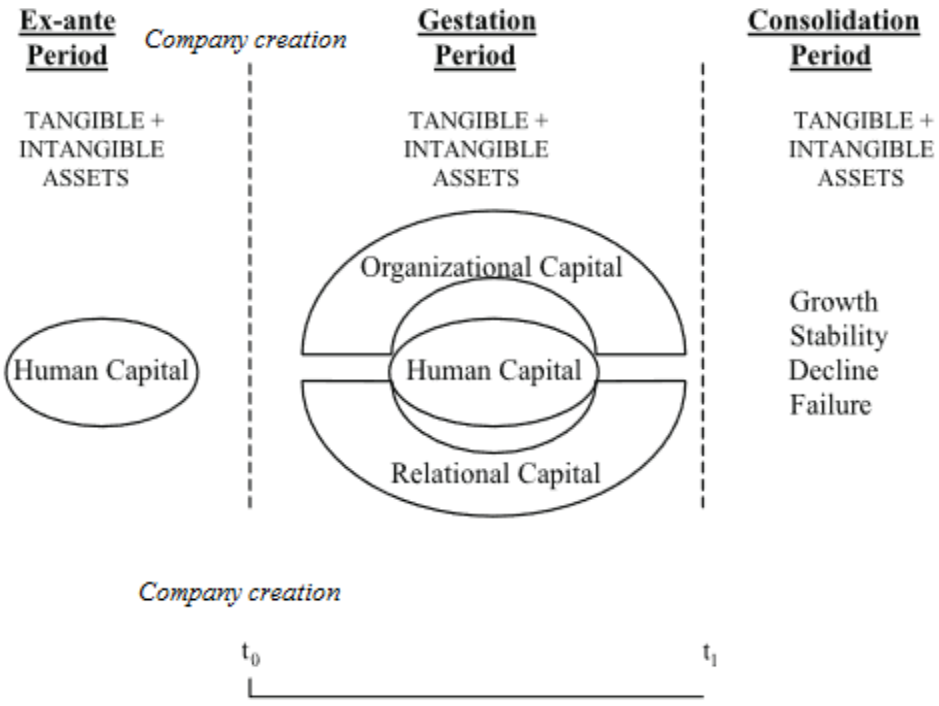


Figure 2. The three stages that an entrepreneurial company goes through [5]

During the first phase, the entrepreneur has a business idea and decides to establish a company. At this stage, only exists the entrepreneur's human capital because the company has not yet been established. The entrepreneur's human capital in this phase may initially have been acquired through formal education, training, previous work experience, different life experiences and so on.

The gestation period begins with the founding of the company. During this phase, which extends from period t_0 to period t_1 (Figure 2), the entrepreneur has to make a great effort to adapt the newly founded company to the market conditions and to turn it into a successful business organization. In doing so, structural and relational capital emerges. Structural capital includes procedures, processes, plans, business strategies, organizational culture, databases, patents, copyrights, software, etc. Relational capital includes all the resources that are related to the relations that the company has with external factors. Human capital, structural capital and relational capital are interrelated and each of these three elements of intellectual capital contributes to the development of the other two elements. Here, the interaction between the three constituent elements of intellectual capital is unique and unlike common material goods where "1+1 equals 2", the nature of intellectual capital is that "1+1 can yield 3, 4 or more". In other words, intellectual capital is characterized by the synergetic effect. In the company's gestation period, the entrepreneur decides how much time and how many resources to invest in each of the components of intellectual capital to develop and strengthen the key competencies of his company.

The company's consolidation phase is a consequence (result) of the company's gestation period. After several years of managing the company and continuous learning, the entrepreneur can assess (predict) the future of his business. The consolidation period can be in the form of: (1) growth, when the company advances and increases its intellectual capital, (2) stability, when the company did not progress as expected of it and the intellectual capital remained at the same level as it was at the beginning, (3) decline, when the entrepreneur's expectations are not met, but the company continues to operate and achieve poor business results and (4) failure, when the entrepreneur decides to close the company.

Which of these four outcomes will occur depends on the utility that the elements of intellectual capital have for the company and the value that is accumulated in them. If the result of the consolidation phase is the growth of the company, the entrepreneur will have to change over time and adapt to new situations. The qualities and skills that he possessed at the beginning change over time. Also, organizational structure and organizational culture, as elements of structural capital, are changing. At the same time, the management should become more formal, but not bureaucratic.

Negative intellectual capital and loss of intellectual capital in an entrepreneurial IT company

The failure rate is significantly high in many research projects and new products. Many of the new drugs are ineffective, many of the new consumer products are of poor quality, and many of the new artistic achievements are not good. Investments in intangible assets have a much higher level of risk and uncertainty than investments in tangible assets. So, for example, if a software project fails, the costs of its development are lost forever, unless the knowledge gained during its development is used for a new entrepreneurial attempt. On the other hand, if an entrepreneur invests in real estate and that property loses value, the investor will be able to recover at least part of his investment. Hence, it should be kept in mind that there is negative human capital, negative structural capital and negative relational capital. In that context, it should be pointed out that not everything people know is useful. Such is the case with the intellectual human capital possessed by a criminal or the entrepreneur who wishes to undertake a criminal enterprise. It is knowledge just as flying an airplane or programming is knowledge. An example of negative structural intellectual capital at the company level is a situation where there is strict adherence to old organizational methods in which organizational hierarchy is rigidly respected and which prevents employees from achieving their work potential. An example of negative relational intellectual capital at the company level might be the situation in a company's marketing, where there is a product focus rather than a consumer focus.

A loss of intellectual capital can occur as a result of wars, environmental disasters, health crises, or economic crises when there is a loss of human lives or displacement of people, destruction of infrastructure and production facilities, the collapse of businesses, etc. (for example, the irreversible loss of intellectual capital in Ukraine today). Also, loss of intellectual capital can occur when an employee who has specific, relevant knowledge for the company, leaves the company for various reasons such as retirement, dismissal, transfer to another company or death. Moreover, this situation will not be recorded in accounting because when an employee with expertise leaves the company, no financial-accounting report will register it. Simply, the modern accounting system measures and provides information only about the visible (material) assets (there are no accounting positions that indicate the invisible assets (intellectual capital)). In this context should be mentioned the forgetting curve of the German psychologist Hermann Ebbinghaus as well. This curve can partially explain the loss of intellectual capital due to forgetting and resistance to acquiring new knowledge, which, ultimately leads to knowledge obsolescence. This is especially important for IT companies that, due to the constant development of ICT, are very susceptible to the risk of obsolescence of the knowledge

they have. Namely, this author discovered as early as 1885 that people forget about 75% of everything they learn in a very short time: in the first 24 hours more than half, and in less than a week they forget about 75% of all information that they received[1].

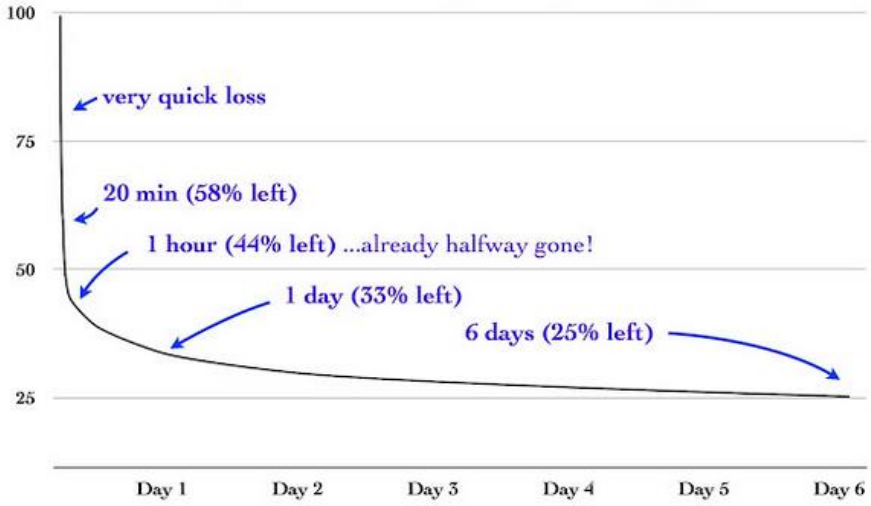


Figure 3. Ebbinghaus' Forgetting Curve [6]

With the penetration of digital transformation into the daily operations of IT companies, in addition to the mentioned scenarios that endanger intellectual capital, business processes are constantly exposed to the danger of its loss. The possibility of cyber attacks that could lead to outages with catastrophic consequences precisely in the area of intellectual capital (encryption and deletion of data, damage to data structures and so on), are one of the main reasons why modern companies in their business operations pay special attention to the protection of intellectual capital through protective plans and procedures for maintaining business continuity in the firm (Business Continuity and Disaster Recovery - BCDR). These plans and procedures are aimed at maintaining business continuity through the protection, repair, and recover of the infrastructures on which the information systems are placed, as well as the data and information as carriers of the intellectual capital. Considering that modern information systems are data-centric in their operation and functioning, the protection of intellectual capital is key to minimize losses from system outages within firms (Cost of Downtime-CoD). According to the Ponemon Institute's newest report, for the year 2022 average total cost of a data breach is 4.35 million dollars, a value that is constantly increasing from year to year [7].

IT entrepreneurship in the period to come

Economic development revolves around the human mind and its ability to transform material resources and intangible ideas into valuable goods and services. The animal world also reorganizes material resources, often with incredible precision (for example, birds make nests and bees make hives). But unlike the animal world, humans create wealth and prosperity through their ability to dream, fantasize, invent, experiment, think, and take risks. A market economy creates wealth and prosperity, not only because of private property, the free market, and the profit motive, but also because of people's willingness to invent and develop new ideas. Human flexibility and ability to distort facts in reality through dreaming, fantasizing, inventing, experimenting, thinking, etc., make homo sapiens the dominant species on Earth. Humans are the only creatures that can experiment inside their minds, which is their specific evolutionary advantage. In that context, it should be emphasized that the social environment such as the family, the educational system, cultural norms and customs, the value system, religious communities, social networks, etc., have a great influence on the formation of entrepreneurs. Biological predispositions are not a decisive factor in whether someone will be a successful entrepreneur or not [8]. Also, entrepreneurship is an economic-social and psychological phenomenon that does not choose space, time, religion, gender, age, level of education, nationality, income, or

[1] The opposite of the forgetting curve can be considered the experience curve, which describes the empirically verified phenomenon that by repeating things, they are performed faster and more efficiently as a result of accumulated experience. This is an extension of the concept of the "learning curve" which describes the reduction of the time required to produce certain products based on increased sophistication and experience.

sexual orientation and which aims to satisfy human desires, needs and demands. Entrepreneurship always and everywhere seeks the latest combination of opportunities and risk. It is a creative response, an innovative solution to new challenges, and a new initiative to create new markets, products and processes. The less favorable, riskier and more unpredictable the circumstances, the more challenging it is for true entrepreneurs. In the context of IT entrepreneurship, it should be noted that some of the most successful IT companies today emerged in times of crisis. Google, eBay, and LinkedIn emerged after the dot.com bubble in 2000. Uber, Airbnb, Instagram, and WhatsApp appeared after the economic-financial crisis in 2007. Zoom Video Communications, the company that provided the most used video conferencing application during the Covid-19 pandemic, in September 2021 reached a market value of $82.19 billion and 4,422 employees. This company had started to operate only ten years earlier, in 2011, with one founder and 40 engineers [9]. This only confirms the fact that there are great business opportunities for IT entrepreneurs even in conditions of great crises when many new human desires, needs and demands are expected to appear. The global pandemic with Covid-19 or the war in Ukraine have posed challenges to all humanity in the search for quick, global and unique solutions because for the first time in human history, a health, military, economic and energy crisis affects everyone equally: rich and poor, developed and undeveloped, successful and unsuccessful. In this new business and cultural environment, in which the only certain thing is that nothing is certain, it is necessary to redesign the phenomenon of IT entrepreneurship and the conditions and ways of its further development. Such redesigning should go in the direction of creating initial strategies, goals, missions and visions in which IT entrepreneurs will incorporate values that should ensure the sustainability of their IT businesses in the long run. Also, their basic motives such as self-interest and profit, should be complemented by environmental ethics, concern for employees and concern for the wider social community.

Conclusion

Today, the future seems more uncertain and unpredictable than ever before. The fact is that the modern business environment is turbulent, unpredictable and risky, which certainly makes it difficult for IT companies to operate. On the other hand, acting in conditions of increased risk and uncertainty is a main characteristic of enterprising people with divergent (creative) thinking.

Additionally, the development of new technologies has contributed to modern economies being largely based on intangible resources. In that context, it should be mentioned that intellectual capital has a strong influence on achieving business success in an entrepreneurial IT company because it contributes to the creation of value and competitive advantage based on knowledge.

Finally, it should be emphasized that all this comes into force even more in the context of health crises, energy crises and economic crises when discontinuity occurs in the business environment and when the need for knowledge exchange has been increased. Hence, future entrepreneurs should take advantage of the experiences of such crises and develop IT businesses whose main goal will not be only to make a profit but also to achieve broader social goals.

References

[1]. htpps://www.cisco.com/c/en/us/solutions/internet-of-things/overview.html.
[2]. https://explodingtopics.com/blog/startup-failure-stats#top-startup-failure-stats.
[3]. Sundać Dragomir, Švast Nataša, Intelektualni kapital-temeljni ćimbenik konkurentnosti poduzeca, Ministarstvo Gospodarstva, Rada i Poduzetništva, Zagreb, 2009.
[4]. Michalski Markus, Vazques Javier, The Importance of Intellectual Capital for the Entrepreneurial Firm, Mälardalen University, Västerås, Sweden, 2008.
[5]. Peña Iñaki, Intellectual capital and business start-up success, *Journal of Intellectual Capital*, 3(2), 2002.
[6]. http://www.senseandsensation.com/2013/03/retrieval-getting-and-forgetting-part-8.html.
[7]. https://www.ibm.com/downloads/cas/3R8N1DZJ (last accessed 23.06.2023).
[8]. Petković Saša, Preduzetništvo i inovacije u digitalnoj eri, Univerzitet u Banjoj Luci, Ekonomski fakultet, 2021.
[9]. https://en.wikipedia.org/wiki/Zoom_Video_Communications.

Software Testing Strategies, Approaches, Methods and Techniques - Overview

Zoltan Kazi[1], Maria Kazi[2]

[1] *University of Novi Sad, Technical faculty "Mihajlo Pupin" Zrenjanin, Republic of Serbia*
[2] *University of Novi Sad, Faculty of Sciences, Novi Sad, Republic of Serbia*

zoltan.kazi@tfzr.rs; mariakazi023@gmail.com

Abstract:
Software products should pass the process of evaluating the quality during the development project in order to ensure as few defects as possible and guarantee correct functioning. This paper presents testing strategies and approaches, lists testing methods and techniques that are classified in two basic groups: static and dynamic techniques. Also, software testing can be classified in two ways: according to the testing approach, and according to the level of testing. Software testing is divided into two large groups of techniques according to the method of execution: manual and automatic. The most important software testing methods and techniques are presented and explained. An overview of comparative analyzes of software testing techniques is also presented with conclusion that present the current state in this important area of software development.

Keywords:
Software, testing, techniques, methods, automation

1. Introduction

Software testing is the process of evaluating software that is created during a software development project in order to determine whether the software conforms to the expected requirements and needs of users, whether it is fit for purpose, and whether there are any defects in the software. Malfunctioning software with defects causes user dissatisfaction and financial, material losses, and in extreme cases it can even cause human casualties.

The main goal of software testing is to deliver to the market, customer, user or client a software product that meets their needs or requirements within a time frame that ensures satisfactory quality. In a narrower sense, testing is checking whether certain software is fully implemented according to the original user requirements. In a broader sense, testing is a quality control system that not only checks software, but also all its accompanying components and features.

The largest number of software errors, according to [1], occurs due to problems in the specification of user requirements (56%). In many cases, the specification is not even written, or it is not precise enough, or it is constantly changing, or there is not good enough communication with the project team. The next biggest source of errors and defects in software is due to design during the software planning (27%). Errors in programming and coding process account for only 7% of error sources while all other reasons account for 10% of errors in software.

Most of the defects in software are design errors, not manufacturing. Good testing provides measures for the following relevant factors: correctness, efficiency, flexibility, reliability, testability, reusability, usability, documentation, maintainability, and integrity structure. [2]

The rest of the paper is organized as follows: section 2 presents classification of software testing techniques; section 3 closer describes static testing techniques, as well as section 4 for dynamic techniques; section 5 lists automated testing techniques and most popular software tools for tests

automation; section 6 is a review of a testing methods comparative analyzes; and final section 7 brings the conclusion.

2. Classification of software testing techniques

All software tests could be divided into the following groups of testing approaches that are occasionally inseparable [3]:

- Static Testing Techniques
- Dynamic Testing Techniques (Figure 1)

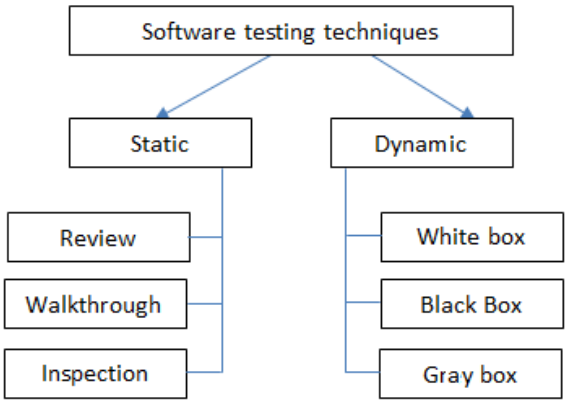


Figure 1: Software testing techniques classification

On Figure 1 we can see the division of static testing methods into review, walkthrough and inspection methods while dynamic testing includes three large groups of techniques: White box, Black box and Gray box techniques. These methods are explained in section 4 of this paper.

Software testing can be classified in two ways:

- According to the testing approach
- According to the level of testing

According to the approach, testing is divided into [4]:

- Functional – testing based on the specification
- Structural, i.e. non-functional testing based on the source code of the software

According to the level, testing is divided into [4]:

- Unit testing
- Integration testing
- System testing (Figure 2)

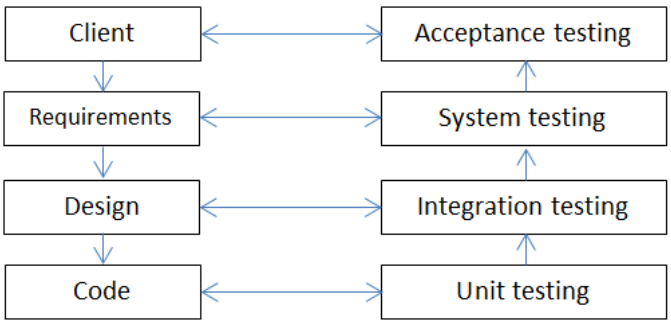


Figure 2: Software testing levels [3]

On Figure 2 we can see the classification of testing into levels. Based on the client's needs, the specification of the user's requirements is formed. Besides that the planning and design of the software, its elements and architecture are carried out. Based on the design, programs are written and the first tested parts are code with unit tests. After that, the parts of the program must be connected into larger units, modules and integration testing is done based on the previous design, models and different specifications. This is followed by system testing of the entire software to see if it complies with the specification, and at the end, acceptance tests are performed to determine whether the software meets its purpose, the basic task that the user needs.

A software testing strategy integrates various test case design methods into a series of steps that result in successful testing of software. Software testing strategies are developed by testing specialist, project managers and software engineer. There are four software testing strategies: Client, Requirement, Design and Unit testing. [5]

Software testing is divided into two large groups of techniques according to the method of execution:

- Manual - performed by manual execution of test cases according to the defined scenario and test plan
- Automatic – it is done with the help of a code, i.e. a script that is executed in a specialized software tool or development environment [4]

3. Static testing techniques

Static testing is a way of testing software in which the program is tested without executing the code, i.e. without starting it, manually or through specialized software tools. This type of testing can be performed in the early stages of software development, before running dynamic tests. They can be carried out on the requirements specification, software design, models, architecture, functional requirements, but also on the source code, and during the test it is possible to detect very early various errors and problems such as: errors in the system architecture, design errors, deviations from standards, unmaintainable and complicated code, deviations from specifications and requirements. The main goal of testing is to find defects at the earliest possible stage of software development.

Static testing approach involves source code investigation that deals with program and symbolic analysis, error handling, model verification, and code inspection to ensure functional requirements, design, and coding standards are observed and estimate software quality without any reference to actual executions. [3]

Static software testing techniques [4] are:

- Review
- Walkthrough
- Inspection

These three groups of techniques are focused on evaluating the material results of work on software development (documentation, models, and specifications) and are carried out in meetings, where one of the participants must have the role of a leading person, and one has the role of recorder.

4. Dynamic testing techniques

Dynamic testing is carried out by testing software that is being developed and can already be executed, that is, it tests the behavior of the software with various variables in the set of input data and program settings. Tests seek to find software weaknesses and defects in a real environment, where after entering input data and selecting input settings, the results are compared to expected results. Dynamic testing methods are:

- Functional testing using Black box methods
- Non-functional, i.e. structural testing

Functional testing views the software as a closed, Black box and the implementation of the program in this case is unknown. This is why the name Black box testing comes from. Software is viewed as a function that maps an input set of values to an output set of values. Tests are determined and formed based on the software specification, are independent of the implementation, and are usable even if the implementation changes. Test development can take place in parallel with software development. The disadvantage of functional testing is that part of the implemented functionalities will not be covered by tests if they are not specified in the specification.

Non-functional, i.e. structural testing, often called White box testing is highly effective in detecting and resolving problems, because faults that are a manifestation of an error in a software, can often be found before they cause trouble. [6]

Third group of techniques is called Gray box testing. It is a combination of Black and White box testing. With the Black box method, the tester does not know the internal structure of the software he is testing, while with the White box method, the structure of the software is known. In Gray box testing, the structure of the software is partially known. This way enables testing to be carried out for web-based applications, for regression testing, for functional tests. The main goal of this method is to find errors that occur due to improper code structure or improper use of the application.

Gray box testing is considered that software tester already has some knowledge of its underlying code or logic. It is based on the internal data structures and algorithms for designing the test cases more than Black box testing but less than White box testing. This method is important in cases when integration testing between two modules of code written by two different developers, where only interfaces are exposed for test. This method can include reverse engineering to determine boundary values. [6]

Advantages of Gray box testing [7]:

- It provides benefit of Black and White box testing techniques
- Tester can design excellent test scenarios
- Unbiased testing
- Intelligent test authoring

Disadvantages of Gray box testing [7]:

- Test coverage that is limited as the access to source code is not available
- Many program paths remain untested
- The test cases can be redundant

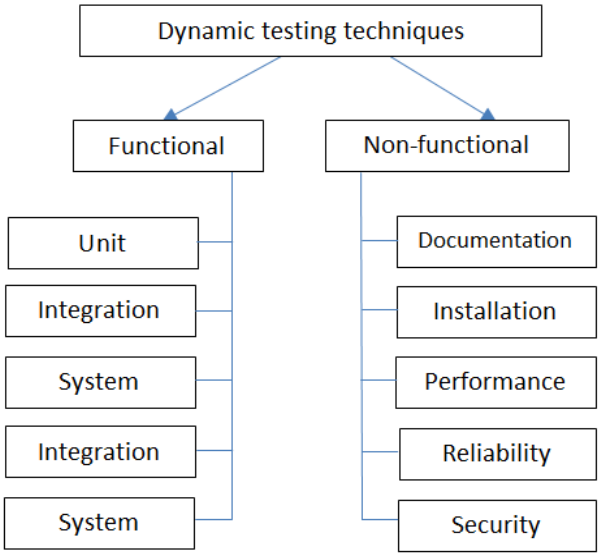


Figure 3: Dynamic software testing techniques

Advantages of Black box testing [8]:

- The number of test cases is reduced to achieve reasonable testing
- The test cases can show presence or absence of classes of errors
- Black box tester has no “bond” with the code
- Programmer and tester both are independent of each other
- More effective on larger units of code than clear box testing

Disadvantages of Black box testing [8]:

- Test cases are hard to design without clear specifications
- Only small numbers of possible input can actually be tested
- Some parts of the back end are not tested at all
- Chances of having unidentified paths during this testing
- Chances of having repetition of tests that are already done by programmer

Types of functional tests [4] (Figure 3):

- Unit testing - unit testing
- Integration testing - integration testing
- System testing - system testing
- Interface testing - connectivity testing (checking how connected software parts communicate with each other - via services, APIs, servers, drivers, etc.)
- Regression testing - regression testing (checking whether the software works as expected after changes in the code, insertion of plugins ("Updates"), improvements in functionality, new requirements, etc.)
- User acceptance testing - acceptance testing for the user (performed by the end user, the client in order to verify whether the requirements have been met, whether everything has been specified and implemented, it is done after unit, integration and system tests)

Structural testing focuses on the implementation of the program and the available, accessible, source code. That's why it got another name - white box testing. The focus of this testing is on the execution of all program structures and data structures in the software under test and the tests are determined accordingly. In this type of testing, software specification, functions and user requirements are not checked. Therefore, it is not possible to detect whether the specified functionalities are actually implemented in the program. Types of non-functional tests (Figure 3):

- Documentation testing - documentation testing (checking whether the documentation (technical and user) used matches what the software, i.e. the system does, whether subsequent changes are documented)
- Installation testing - installation testing (checking whether the software is installed and uninstalled correctly by the user, whether the procedures within the installation process take place as intended)
- Performance testing – testing of characteristics and performance, sub-techniques:
    - Load testing - load testing (checking how the system functions with a large number of virtual users executing a large number of transactions in a short period of time)
    - Stress testing - load testing (robustness testing, when the software is pushed beyond the limits of normal functioning with large amounts of data and a large number of open connections, e.g. servers, OS, network software)
    - Endurance testing - endurance testing (checking how the system functions during a long period of use under load, working with memory, a larger number of users in the software)
    - Spike testing - testing the so-called "spikes" (checking how the system works when there are extreme variations in network "traffic", sudden drops or increases in the number of users, how long is the "recovery")
- Reliability testing - reliability testing (checking that there will be no software crashes and failures or the probability that this will not happen, the way the user uses the software)

- Security testing - security tests (is there a possibility of loss of data and information, unauthorized access to software, possible "attacks" and "intrusions" on the system, software in the network environment, especially with web applications, etc.)

Unit Testing ("Unit Testing") - Refers to the testing of individual units of the source code, such as classes or parts of classes, although this depends on the technique and way of creating the program. The smallest functional unit of source code is usually a single method within a class. Unit tests can be used by developers, not just software testers, to test their own written code. For this testing, a particular environment for writing unit tests is most often used (e.g. Junit within the Java programming language). A unit test is a piece of code that tests some other piece of program code. Execution of tests is most often automated and can be executed and repeated several times, i.e. as many times as necessary and necessary.

Module testing or unit testing is a process of testing the individual sub-programs, subroutines, classes, or procedures in a program. More specifically, rather than initially testing the program as a whole, testing is first focused on the smaller building blocks of the program. [9]

Integration Testing (Engl. "Integration Testing") - After unit testing is completed; the units are integrated into the whole. The main focus of Integration Testing is on the verification of functionality and interfaces between connected and integrated modules.

System Testing (Engl. "System Testing") - When all parts of the software are finally integrated into a complete system, system testing checks the behavior of that system as a whole in relation to the system specification. When the majority of functional requirements have been verified at lower levels of testing (integration tests), then the emphasis is placed on non-functional requirements, such as the speed of the software, its security, reliability, i.e. robustness, etc.

5. Automated testing methods and software

In recent years, most of the software testing is done with the software tools for tests automation which lessens the number of people working on software and finds the errors that can be escaped through the eyes of the tester. Automation testing contains test cases which makes the work easy to capture different scenarios. It plays a vital role in the software testing success. [10]

Automatic software testing is done with the help of code, i.e. a script that is executed in a specialized software tool or development environment with the goal of automating the steps in executing a specific test case. Files containing automated logic are called test scripts and can be written in all modern programming languages, regardless of programming paradigm. Automated testing is often used so that test cases within a particular test suite that are slow to execute manually can be quickly executed automatically. This is important for regression testing (improvements, additions, new functions, additions, "updates", etc.), when the testing contains all the test cases that were previously performed on the software, but there are newer tests that test additional functionality. Such tests can include thousands of test cases when the need for automatic testing is very clear, with the motive to save time and therefore money. Automated testing software and tools should have the ability to enter input data into the system, compare expected and obtained test results, and ultimately generate a detailed test execution report. [4]

Which test cases should be automated?

- Those of high risk - test cases that are critical for business
- Repeatable test cases
- Those cases that are too tedious to perform manually or are difficult to perform manually
- Test cases that are long and time consuming

Which test cases should not be automated?

- Test cases that are newly designed and have not been manually executed at least once
- Test cases for which requirements change frequently
- Test cases run on an ad-hoc basis

The following steps are taken in the test automation process:

- Selection of testing tools (it depends on the technology under which the software was developed, different tools support different testing techniques)
- Defining the scope of automation (critical functions, data-intensive scenarios, common functions, complex test cases, same tests for different functions, technical feasibility)
- Test planning, design and development (creating test execution schedules, determining when tests will be executed, preparing the execution space)
- Test execution (loading input data, executing automated test script, generating test execution results)
- Test maintenance (checking whether new functionalities have been added to the software and therefore (or to) the test(s) or not, checking changes in previously performed tests and their re-execution, revision of performed tests in order to increase their effectiveness)

Tools for automatic software testing: testRigor, Ranorex Studio, Kobiton, LambdaTest, Avo Assure, Subject7, Selenium.

6. Review of testing methods comparative analyzes

Nowadays, it is very important that the software industry develops quality application programs. This is why the testing phase plays a vital role in improving user satisfaction with software applications. Various studies show that almost 30% of the entire software development effort is used to perform testing activities. Every software company or application developer follows a certain set of testing strategies and uses some software testing tool to ensure quality. A testing strategy must be decided between manual and automated testing. Paper [11] reviews and analyzes the performance metrics of software testing tools and testing strategies used to improve the quality of an application being developed. The result of the review indicates that it can guide the project manager to make trade-off decisions to select the appropriate test tools and test strategies applicable to their project domain. In this analyzes it was concluded that automated testing is faster than manual, more cost effective, easily reusable, high infrastructure is needed, more programming, to high need for tools usage, with very high quality of testing and low human resources. Manual testing do not need special training, needless infrastructure and usage of tools, without programming but is too flexible, need more human resources and the quality of testing can be lower than with automated testing.

The purpose of software testing is to check whether the software meets the requirements, needs and expectations of the users. Software testing should provide an independent view of the software to enable implementation risk to be understood and observed. Therefore, research [12] describes and compares three main software testing techniques: Black, White and Gray boxes. Black box testing is least exhaustive and time consuming, while White box is most exhaustive and time consuming, while Gray box is somewhere in the middle. White box is better for testing data domains and internal boundaries. Black box testing can be tested only by trial and error method. Gray box also can be used for testing data domains and internal boundaries.

According to [13] Black box testing advantages are:

- Code knowledge is not required; tester's perception is very simple.
- User's and developer's view are separate.
- Access to code is unrequired, quicker test case development.
- Efficient and suitable for large parts of code.

Black box testing disadvantages are:

- Limited coverage, few test scenarios are designed and performed.
- Some parts of the backend are not tested at all.
- Inefficient testing due to the limited knowledge of code possesses by a tester.
- Test cases are difficult to design without clear specification.

White box testing advantages are:

- Code optimization can be performed.
- Easy to identify data and cover more test cases due to tester's knowledge of the code.
- Errors in hidden codes are revealed.

White box testing disadvantages are:

- Specialized tools are required such as debugging tools and code analyzers.
- It's often expensive and difficult to maintain.
- Impossible to find and test all the hidden error and deal with them without going out of time.

Paper [14] provides comparative analysis between White box, Black box, Gray box, and Regression techniques for software testing. Conclusions are following:

- The main features of White box testing are full coverage of the code.
- Black box testing is efficient when it is used on larger systems.
- While Black box testing can still help one fulfils a software requirement of a user. Developers have more time to fix bugs in Gray box testing and it is much more effective in integration testing.
- Regression testing is most important when new changes are incorporated with the existing system and mostly when the changes affect different old modules which are not being modified.

Summarizing the analysis of the current development of the testing industry, the authors in [15] noted the fact that there is still no software that will enable testing and recording its results in full. For developers, there are many different testing environments on different platforms with diverse interfaces and programming languages, but for software testers today there is no complete software, which poses a problem for the implementation of the most important phase of the design and implementation of a software product. If it is necessary to quickly prepare new software versions without quality degradation, testing as part of development must not lag behind. This requires a transition from slow, labor-intensive testing methods to faster and fully automated testing. It is concluded that the development of information technology and software contributes to the widespread use of software testing and will only expand this area, which requires an increasing number of experts.

7. Conclusion

Every software product should be released or delivered to the users after it has passed the proper review of the development, testing, documentation and defect correction process. Testing observes and analyzes software performance, stability and error handling by setting up test scenarios under controlled conditions with evaluation of the results. Software must be tested to see if it meets the needs and requirements of customers, as well as if it conforms to certain standards. Because of that aim of this paper is to bring an overview of the most common testing approaches, methods and techniques.

The contribution of this paper is the systematization of knowledge and information in the field of software testing, since it has been noticed in the available literature that classifications and divisions of other authors are not comprehensive, but are partial and incomplete.

From the review of testing methods comparative analyzes it can be concluded that on the market are many different testing environments on different platforms with diverse interfaces and programming languages, that there is still no software that will enable testing and recording its results in full but, for software testers today there is no complete software tool for their job.

References:

[1] R. Patton: Software Testing, Second Edition, SAMS, 2006.

[2] P. Jiantao: Software Testing, Carnegie Mellon University, 18-849b Dependable Embedded Systems, Spring 1999.

[3] U. Mubarak Albarka: Comprehensive study of software testing: Categories, levels, techniques, and types, TechRxiv, 2020.

[4] M. Živković: Software Testing, Singidunum University, Belgrade, 2018, (on Serbian).

[5] A.A. Sawant, P.H. Bari, P.M. Chawan: Software Testing Techniques and Strategies, International Journal of Engineering Research and Applications (IJERA), Vol. 2, Issue 3, May-Jun 2012, pp. 980-986.

[6] R.K. Chauhan and I. Singh: Latest Research and Development on Software Testing Techniques and Tools, International Journal of Current Engineering and Technology, Vol.4, No.4 (Aug 2014), pg. 2368-2372.

[7] A. Bansal: A Comparative Study of Software Testing Techniques, International Journal of Computer Science and Mobile Computing IJCSMC, Vol. 3, Issue. 6, June 2014, pg.579 – 584.

[8] I. Jovanović: Software Testing Methods and Techniques, The IPSI BgD Transactions on Internet Research, Volume 5 Number 1, 2008, pg. 30-42.

[9] G.J. Myers, T. Badgett, C. Sandler: The Art of Software Testing, Third Edition, John Wiley & Sons, 2012.

[10] S. Karuturi, M.M. Gowda: Research on software testing techniques and software automation testing tools, International Conference on Energy, Communication, Data Analytics and Soft Computing, 2017, pg. 1-4.

[11] M.J. Pramod, S. Priyadarsini, R.V. Renju, S. Sumisha, M. Prasanna: A Comparative Analysis On Software Testing Tools And Strategies, International Journal of Scientific & Technology Research, Volume 9, Issue 4, April 2020, pg. 3510-3515.

[12] S. Rani, D. Gupta: A Comparative Study of Different Software Testing Techniques: A Review, Journal of Advances in Shell Programming, Volume, Issue 1, 2018, pg. 1-8.

[13] A Comparative Study of Dynamic Software Testing Techniques, International Journal Advanced Networking and Applications, Volume 12, Issue 3, 2020, pg. 4575-4584.

[14] S.W.G. AbuSalim, R. Ibrahim, J.A. Wahab: Comparative Analysis of Software Testing Techniques for Mobile Applications, Journal of Physics: Conference Series 1793, 2021, pg. 1-11.

[15] A. Fedosov, D. Eliseeva, N. Khodakova, O. Mnatsakanyan, N. Kulikova: Comparative Analysis of Products for Testing Software, Book chapter, Springer Nature Switzerland, AISC 1294, December 2020, pp. 242–252.

Edge Computing System to Form a Data Center on Air Pollution in the Traffic Environment of Smart Cities

Gordana Jotanovic[1], Aleksandar Damjanovic[2], Goran Jausevac[1], Zeljko Stojanov[3], Vladimir Brtka[3], Dragan Perakovic[4], and Miroslav Kostadinovic[1]

[1] *University of East Sarajevo, Faculty of Transport and Traffic Engineering, Vojvode Misica 52, 74000 Doboj, Bosnia and Herzegovina*
[2] *Cape Ann Enterprises, Stara Trznica 14, 75000 Tuzla, Bosnia and Herzegovina*
[3] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Djure Djakovica bb, 23000 Zrenjanin, Republic of Serbia*
[4] *University of Zagreb, Faculty of Transport and Traffic Sciences, Borongajska cesta 83a, 10000 Zagreb, Croatia*

gordana.jotanovic1@sf.ues.rs.ba; aleksandar.damjanovic4@hotmail.com; goran.jausevac1@sf.ues.rs.ba; zeljko.stojanov2@uns.ac.rs; vladimir.brtka2@uns.ac.rs; dperakovic3@fpz.unizg.hr; miroslav.kostadinovic1@sf.ues.rs.ba

Abstract:
The paper proposes the architecture of the Edge Computing System, the purpose of which is to form a Data Center of air pollution in the traffic environment of smart cities. The data collected in the Data Center over a long period of time enables the monitoring of air quality trends. Which is important for assessing the effectiveness of pollution reduction policies and monitoring progress. However, open data about air pollution in the Data Center can put into question the security of the data in the Edge Computing System. Therefore, a Data security model was designed to protect data from unauthorized user access. The data security model for protecting data from unauthorized user access is based on the authorization and authentication of user access to data.

Keywords:
Edge Computing System, Data Center, smart cities, Edge Device, Edge Unit, and Entity Relationship Attribute (ERA).

1. Introduction

In urban areas such as smart cities, a major problem is the exposure of the population to air pollution. Air pollution has a significant impact on the health of the population and negative consequences for the entire ecosystem. Therefore, continuous monitoring of air pollution with the aim of detecting critical areas in urban environments is essential for the quality of life. For this purpose, according to the recommendations of the European Environmental Agency, it is necessary to monitor the concentration of certain particles, such as PM10, PM2.5, O3, NO2, BaP and SO2, and based on that, determine trends and recommendations for improving air quality. Air pollution due to the high concentration of particles in the last few years exceeds the limit values set by the directives of the European Union on ambient air quality, as well as the recommendations of the World Health Organization on air quality [1]. Emissions of exhaust gases from vehicles, as well as from the transport sector, are not a negligible factor of pollution at the global level [2]. Environments such as smart cities often contain roads near residential areas. In order to improve the quality of life in such residential areas, it is necessary to monitor air pollution along roads. Edge Computing systems in smart cities can be used in healthcare, logistics, traffic, transport, preservation of the human environment, for the purpose of creating value and sustainable business [3]. Therefore, we based our work on the formation of an Air Pollution Data Center based on Edge Computing architecture, in order to solve the problem of population exposure to excessive air pollution in urban areas. Based on the data collected in the Data Center, it is possible to determine trends and recommendations for improving air quality. The established Data Center has a low level of data protection due to open data access by various users (citizens, authorities, scientists, management bodies and environmental protection agencies). Therefore, it was necessary to design a model to protect data from unauthorized user access.

The paper is structured as follows: the first part contains an introduction to the issue. The second part deals with the thematic considerations of systems that enable the monitoring of air pollution in the traffic environment of smart cities. The third part describes the architecture of the Edge computer system for the formation of the Air Pollution Data Center. The fourth part refers to the Data security model of the Edge Computing System. Concluding remarks and future research developments are presented in the fifth part.

2. Previous work

Today, a significant number of systems are in use that enable monitoring of air pollution in the traffic environment of smart cities. Existing solutions are mostly based on mobile technologies, the application of the Internet of Things (IoT), modern distributed software architectures, and complex analyzes of large amounts of data. Based on a literature review of tools and techniques for assessing exposure to air pollution and road traffic noise, it was found that 84% of studies use average datasets on a daily or annual scale, while only 12% of studies use time-scaled traffic data [4].

In [5] the fifth generation (5G) of mobile systems is presented. The characteristics of the 5G network are explained, such as: bandwidth, latency and the connection of a large number of devices using the Internet of Things (IoT). One of the main challenges of IoT applications is solutions to reduce latency. The study proposes an edge computing management mechanism for IoT applications in smart cities. That is. proposes a database (called an information map) that allows computing servers at the edge of smart cities to store information about edge services.

Existing solutions for designing air quality monitoring systems according to [6] can be identified in three basic categories from the aspect of system architecture: (1) systems with centralized architecture - smart devices are data sources for a central computer node that processes the collected information and provides it to users, (2) systems with decentralized architecture - processing is divided into several networked nodes that accept data, combine them, and provide information to users, and (3) systems with cloud-based architecture - processing is performed by computers in the cloud that share resources (data storage, tools for visualization, software services).

The Web of Things (WoT) has the task of creating a decentralized Internet of Things [7]. Edge computing improves IoT computing performance by reducing escalation in resource congestion situations. The goal of the research is to reduce the transmission delay between edge servers and end users which improves the response time for real-time WoT applications. Also. the paper describes how edge computing improves WoT performance and concentrates on aspects of transmission, storage and computation.

Overview of the application of wireless sensor networks for measuring air pollution with special reference to applied research methodologies, used microcontrollers and communication devices, sensors for the detection of air pollutants, test locations and system performance. The authors indicate that the most practical and economically profitable way of monitoring air pollution is based on microcontroller-based IoT systems [8].

When analyzing air pollution, in addition to measured pollutant values, location quality parameters (proximity to an industrial zone or landfill, proximity to shopping centers), the presence of green areas and detailed traffic data should also be taken into account [9].

An air pollution monitoring system was designed at the Indian Institute of Engineering Science and Technology, Howrah, India [10]. The system uses an Arduino UNO as a controller unit, and LoRa technology is used for long-range communication. The system uses gas sensors MQ 135 for the detection of ammonia, sulphide and carbon dioxide and MQ 9 for the detection of carbon monoxide. The ESP8266 Wi-Fi microchip is used to connect to the Internet, while the process and application layers of the system are located in the cloud servers.

The paper [11] is based on the development of the concept of smart cities, especially in applications related to safety (such as emergency fires, patient health monitoring or real-time production). The proposed concept is based on fog computing, i.e. cloud computing, with the idea of locating closer to the end devices.

A decision support system that monitors the state of air quality, predicts air pollution and suggests strategies to optimize traffic in order to reduce pollution, e.g. banning traffic on the road section,

optimizing signaling, banning movement of heavy vehicles, or regulation of parking [12]. The system is based on real-time data collected by various IoT devices, which is then combined with a vehicle emissions model, a traffic model and a weather model.

The paper [13] proposed CONTESS, which is a framework for context-aware edge computing with selective sensing that uses information about the context of the sensing environment to improve its applicability to smart IoT systems. Also, the paper demonstrates the capabilities of CONTESS in the scenario of a parking management system for a smart city environment. Implemented CONTESS using Linked Data and Semantic Web technologies. An ontology based on OWL was designed and then by simulating the proposed scenario using OMNET++ network simulator along with Veins framework and SUMO traffic simulator.

The theme of the paper [14] is to move AI workloads for video analytics to the edge of the network from the cloud, providing improved latency and bandwidth savings. The paper provides an overview of current technologies used in video analytics with the help of Edge AI in smart cities. It also examines various artificial intelligence models and privacy-preserving techniques used in edge video analytics. Including security and surveillance, transportation and traffic management, health, education, sports and entertainment, etc.

Road traffic is not the "number one" polluter in Bosnia and Herzegovina, but its influence cannot be ignored, which indicates the need to systematically approach the problem of measuring air pollution as a consequence of road traffic [15].

3. Edge Computing System architecture for the formation of a data center on air pollution

The Edge Computing System enables the formation of the Air Pollution Data Center. The system provides relevant information that is used to make decisions in order to reduce pollution. The specificity of the system is that there are multiple levels of data access in the Data Center. The Data security model was designed for this purpose. The architecture of the system is distributed and enables the collection, local processing of data, with the possibility of offloading the centralized server. This approach enables a high level of reliability, which is especially important in the context of quick and accurate collection of air quality information for making important decisions in real time. The architecture of the Edge computer system for the formation of the Air Pollution Data Center is shown in Figure 1.

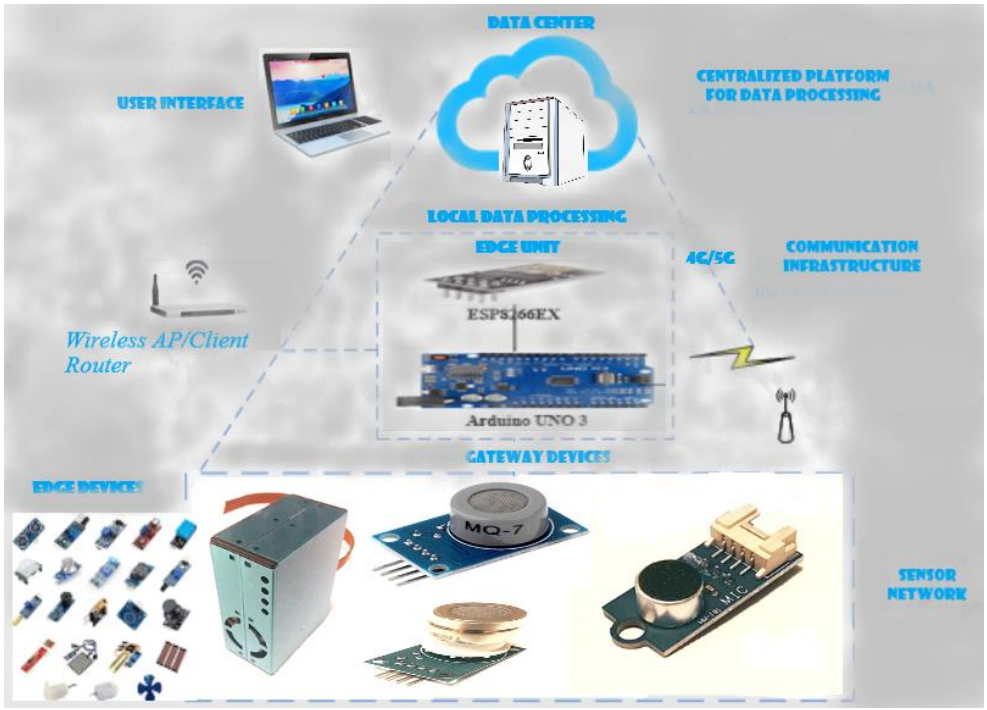


Figure 1: Architecture of the Edge computing system for the formation of the Air Pollution Data Center

The basic components of the Edge Computing System for the formation of a data center on air pollution include:

- Sensors for measuring air pollution, such as PM1, PM2.5 and PM10 particles, air temperature and humidity, carbon monoxide (CO), carbon dioxide (CO2) and noise level measurement. The technical specifications of the sensor include measurement accuracy, measurement range, sampling rate, and the ability to communicate with other devices.
- A network of sensors placed at different locations along roads, intersections, highways, etc. The sensor network enables the monitoring of air quality throughout the area and the identification of local sources of pollution.
- Edge Devices and controllers make up Edge Units. They are located close to the measurement sensors and perform local data processing. Some parts of data analysis and processing are performed on these Edge Units to reduce data traffic to the Data Center.
- Gateway devices are a bridge between local devices and the Data Center. They enable communication between sensors, Edge Units and the Data Center.
- The communication infrastructure that serves to communicate with the central platform for data processing via wireless connections includes Wi-Fi, 4G/5G networks. The accompanying communication protocols comprise the connection between the sensor and the Data Center. Also, they must be efficient, reliable and secure to ensure accurate and secure data exchange. The infrastructure enables the transfer of data in real time from all locations to the Data Center.
- The Data Center represents a Centralized Platform for Data Processing, storing and displaying data to users. A centralized server accepts data from all Edge Devices. Advanced analytics and machine learning can be applied to the data in the Data Center to generate insights and predictions based on the collected data.
- The User Interface includes citizens, scientists, authorities and environmental protection agencies to access pollution data through the application. The application provides information on the current state of air quality, warnings and recommendations for protection. Authorization and authentication of users who can access data in the Data Center has been performed.

The key components of the Edge Computing System for the formation of a Data Center on air pollution in smart cities are:

- Collected data in the Data Center is analyzed to identify causes of pollution, seasonal trends and long-term patterns. Reports can be generated to inform authorities, the public and experts.
- The system can automatically generate warnings and notifications in case of exceeding permitted levels of pollution. Data from the Data Center enables the necessary protection measures to be taken.
- Smart cities often strive to involve citizens in the processes of monitoring and solving air pollution problems. Citizens can submit information to the Data Center about local sources of pollution and contribute to a better understanding of the situation.
- City managers and environmental protection agencies use the collected data to make policy decisions, plan infrastructure projects and implement measures to reduce pollution.
- Open access to the Air Pollution Data Center allows researchers and developers to use this information to analyze and create new solutions.

The data stored in the Data Center can also contribute to a greater understanding of the connections between air pollution, meteorological conditions and other factors that affect air quality in smart cities. The proposed Edge Computing System consists of a Centralized Platform for Data Processing, a Sensor Network that serves to collect air pollution data, Communication Infrastructure (Wi-Fi, 4G/5G) and Local Data Processing that takes place in microprocessor systems.

The Edge Computing System works as follows:

- At key places in the traffic environment (near intersections and places where the frequency of traffic is high), sensors are placed to measure the concentration of pollutants.

- Sensors are connected to Gateway Devices that are placed on traffic lights, light poles or other infrastructure elements. The devices allow connecting sensors to the network and enable Local Data Processing.
- Local Data Processing is performed on Gateway Devices. It includes filtering, aggregating data to reduce the amount of data that will be sent further to the Data Center.
- After local processing, relevant and important data are sent to the Data Center for long-term monitoring of air quality.
- Analysis results and pollution levels are displayed graphically in real time on the User Interface so that they can be informed about air quality.

Data security in the Data Center is of key importance, especially when it comes to sensitive air pollution data. Technical specifications should include authentication and authorization mechanisms to ensure data protection from unauthorized access.

4. Data security model of the Edge Computing System

Security and data protection are essential requirements of the presented Edge Computing System. In order to provide data security in the Edge Computing System, it is necessary to provide data protection systems. The Edge Computing System contains a Data Center accessed by various users (citizens, authorities, scientists, management bodies and environmental protection agencies). Also, the Data Center can include data that may be interesting to users that directly cause pollution in smart cities. Therefore, the presented users should not have equal rights to access and use data from the Data Center. Data protection in the Data Center is implemented in the form of authorization and authentication of user access to data. Security against unauthorized user access to data in the Edge Computing System is implemented using a Data security model designed exclusively for this purpose. The model enables authorization and authentication of user access to data. It also has the possibility for a user with authorized access to record and display the data. The model was designed based on the user requirements shown in Table 1.

Table 1:
Registration of user requests

USER REQUEST	DESCRIPTION
Login	Allows the User to log into the system, with adequate authorization and authentication.
Token refresh	Provides automatic token refresh after the old token expires.
Delivery of all categories	Listing of all categories.
Delivery of a particular category by ID	Listing of a certain category by its ID with associated data.
Add new category	Add new category.
Adding data	Adding data to a particular category.
Display data	Display data by data categories.

Based on user requirements, an Entity Relationship Attribute (ERA) model was created for the authorization and authentication of users with the possibility of accessing the Data Center, (Figure 2).

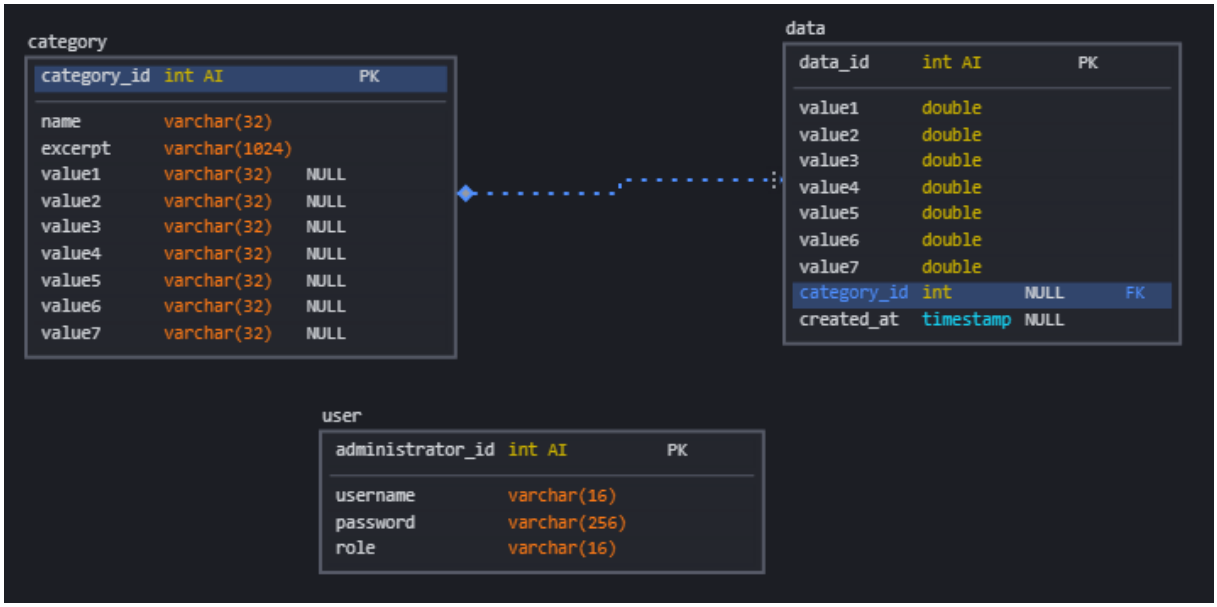


Figure 2: Data security model for authorization and authentication of users

The model is made up of three entities: user, data, and category. Each entity is protected by authorization and authentication. The system user can also have the role of system administrator. The Data Center contains fundamental user data required for authorization and authentication, namely username, password, and roles. Users gain access to the system by registering. Each registered user has the right to create a new category, view the data belonging to the given category at the permitted level of access. The first step in creating a Data Center is to create a user with access rights. System entities are located in the Data Center. The Data Center was created using the phpMyAdmin interface. Access to the Data Center is protected by user authorization and contains a table of users. A system user can have only one role. The user in the Data Center contains fundamental data for authorization and authentication, namely username, password, and roles. Users gain access to the system by registering, and each registered user has the right to create a new category, view data belonging to a given category, and display data.

5. Conclusions

Edge Computing System for forming a Data Center on air pollution enables effective monitoring and management of air quality in the traffic environment of smart cities. Which results in the implementation of targeted measures to reduce pollution and improve the environment. The system combines the advantages of IoT technology for data collection with the capabilities of edge computing for local data processing and data management at the measurement point at the edge of the network.

The Edge Computing System for the formation of a Data Center on air pollution in the traffic environment of smart cities has a number of advantages that contribute to a better understanding, monitoring and management of air quality. The system can generate automatic warnings in case of exceeding permitted pollution levels. It also allows residents to take immediate protective measures to reduce exposure to excessive pollution. The data collected in the Data Center over a long period of time enables the monitoring of long-term trends in air quality. Such an approach is important for assessing the effectiveness of pollution reduction policies and for monitoring progress. Open data on air pollution allows researchers to analyze the data and identify new patterns and connections between pollution and other factors. Also, they contribute to a better understanding of the problem and the development of new solutions. The transparency of data on air pollution in the environment of smart cities encourages the authorities to be more responsible towards the living environment. They encourage the adoption of initiatives to reduce emissions and improve technologies. Collected data can be compared with data from other regions and cities, enabling a better understanding of global pollution patterns and sharing information on successful solutions. However, if necessary, the system provides for limited data access.

Access to the Data Center is limited by user authentication and authorization. Certain users such as: citizens, authorities, scientists, and environmental protection agencies have their own levels of access to data in the Data Center. In this way, data is protected that has wider significance for the preservation of the ecosystem and the protection of people.

There are challenges and shortcomings in the operation of the system. Disadvantages relate to sensor accuracy, managing large amounts of data, system maintenance, and ensuring compatibility of different devices and sensors. Considering the mentioned shortcomings, the implementation of the system requires careful planning, management and coordination of all elements of the system in order to achieve the best results for the improvement of air quality in smart cities.

The future development of the Edge Computing System should go in the direction of connecting the system with the infrastructures of smart cities. Such a move would enable better management of traffic, energy consumption and urban planning to reduce pollution. The development of comprehensive open data ecosystems would enable the dissemination of information and collaboration between different organizations, researchers and citizens to solve pollution problems in smart cities. International cooperation and data sharing on air pollution enables a better understanding of global pollution patterns and the development of more effective international environmental protection policies. The future development directions of the Edge Computing System can contribute to the improvement of air pollution monitoring and enable a better solution to this important problem faced by modern society.

References:

[1] Air quality in Europe - 2019 report. European Environment Agency. Publications Office of the European Union, Luxembourg. 2019.

[2] Anenberg, S., Miller, J., Henze, D. A. V. E. N., & Minjares, R. (2019). A global snapshot of the air pollution-related health impacts of transportation sector emissions in 2010 and 2015. International Council on Clean Transportation: Washington, DC, USA.

[3] Bhandari, S. S., Devullapalli, R., Swapnil, A., Karri, R., & Gopi, C. S. (2023). EDGE COMPUTING.

[4] Khan, J., M. Ketzel, K. Kakosimos, M. Srensen and S. S. Jensen (2018) Road traffic air and noise pollution exposure assessment - A review of tools and techniques. Science of The Total Environment,Volume 634, pp. 661-676. doi: 10.1016/j.scitotenv.2018.03.374.

[5] Khanh, Q. V., Nguyen, V. H., Minh, Q. N., Van, A. D., Le Anh, N., & Chehri, A. (2023). An efficient edge computing management mechanism for sustainable smart cities. Sustainable Computing: Informatics and Systems, 38, 100867.

[6] Mokrani, H., R. Lounas, M. T. Bennai, D. E. Salhi and R. Djerbi (2019) Air Quality Monitoring Using IoT: A Survey. In Proceedings of 2019 IEEE International Conference on Smart Internet of Things (SmartIoT), pp. 127-134. Tianjin, China. doi: 10.1109/SmartIoT.2019.00028.

[7] Anees, T., Habib, Q., Al-Shamayleh, A. S., Khalil, W., Obaidat, M. A., & Akhunzada, A. (2023). The integration of WoT and edge computing: Issues and challenges. Sustainability, 15(7), 5983.

[8] Kingsy Grace R. and S. Manju (2019) A Comprehensive Review of Wireless Sensor Networks Based Air Pollution Monitoring Systems. Wireless Personal Communications,Volume 108, pp. 2499-2515. doi: 10.1007/s11277-019-06535-3.

[9] Jain, V., M. Goel, M. Maity, V. Naik and R. Ramjee (2018) Scalable measurement of air pollution using COTS IoT devices. In Proceedings of the 10th International Conference on Communication Systems & Networks (COMSNETS), pp. 553-556. Bengaluru, Inddia. doi: 10.1109/COMSNETS.2018.8328271.

[10] Walling, S., J. Sengupta and S. Das Bit (2019) A Low-cost Real-time IoT based Air Pollution Monitoring using LoRa. In Proceedings of 2019 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), pp. 1-6. GOA, India. doi: 10.1109/ANTS47819.2019.9117963.

[11] Songhorabadi, M., Rahimi, M., MoghadamFarid, A., & Kashani, M. H. (2023). Fog computing approaches in IoT-enabled smart cities. Journal of Network and Computer Applications, 211, 103557.

[12] Miles, A., A. Zaslavsky and C. Browne (2018) IoT-based decision support system for monitoring and mitigating atmospheric pollution in smart cities. Journal of Decision Systems, Volume 27, supplement 1, pp. 56-67.doi: 10.1080/12460125.2018.1468696.

[13] Ben Sada, A., Naouri, A., Khelloufi, A., Dhelim, S., & Ning, H. (2023). A Context-Aware Edge Computing Framework for Smart Internet of Things. Future Internet, 15(5), 154.

[14] Badidi, E., Moumane, K., & El Ghazi, F. (2023). Opportunities, applications, and challenges of edge-AI enabled video analytics in smart cities: a systematic review. IEEE Access.

[15] Jotanovic, G., V. Brtka, G. Jausevac, Z. Curguz, M. Kostadinovic and E. Brtka (2020) Rerouting Traffic Based on Noise Values and Number of Vehicles in Urban Areas. In Proceedings of the 10th International conference on Applied Internet and Information Technologies (AIIT2020), pp. 147-151, ISBN: 978-86-7672-327-3, Zrenjanin, Serbia.

Communication Performance of The Laboratory System for Measuring Fuel Mass Flow

Zoran Ristikić[1], Svetko Milutinović[2], Milan Eremija[3] and Ibrahim Badnjar[4]

[1] *University of East Sarajevo, Faculty of Transport and Traffic engineering, Vojvode Mišića 52, Doboj, BiH*
[2] *University of East Sarajevo, Faculty of Transport and Traffic engineering, Vojvode Mišića 52, Doboj, BiH*
[3] *University of East Sarajevo, Faculty of Transport and Traffic engineering, Vojvode Mišića 52, Doboj, BiH*
[4] *University of East Sarajevo, Faculty of Transport and Traffic engineering, Vojvode Mišića 52, Doboj, BiH*

zoran.ristikic@sf.ues.rs.ba; s.milutinovic@hotmail.com; eremijamilan5@gmail.com; ibrahimbadnjar97@hotmail.com

Abstract:
For precise laboratory measurements, the gravimetric method (mass measurement) is mainly used. Measurement of fuel consumption and temperature in the internal combustion engine testing laboratory of the Faculty of Transport and Traffic Engineering in Doboj is carried out using the "Fuel balance 733S" system manufactured by AVL GmbH. The paper describes the working principles of modern laboratories for engine testing, with special emphasis on the system for measuring the mass flow of fuel. The method of communication between the measuring system and the computer workstation is described. The measuring system "Fuel balance 733S" works on the communication standard of serial communication RS232. In order to enable the connection of the system to modern computer workstations, a "gateway" is used for data translation to the USB protocol standard. Also, the paper shows the performance of communication in this implementation of connecting laboratory measuring systems.
Keywords:
mass fuel flow measurement, laboratory testing, internal combustion engine, serial communication, USB protocol, communication performance.

1. Introduction

Testing an internal combustion engine is extremely important, not only during its development but also after any engine overhaul and testing for scientific purposes [1]. By means of engine testing, the impact of exploitation on its basic performance indicators can be measured. In order to carry out laboratory tests of internal combustion engines, it is necessary to have measuring equipment for the required performance. The engine testing laboratory of the Faculty of Transport and Traffic Engineering in Doboj is equipped with complete equipment for testing internal combustion engines. The measurement of fuel mass flow is of great importance in engine testing. The paper briefly describes the fuel mass flow measurement system AVL "Fuel balance 733S", which communicates with the host computer via the RS232 communication standard. Since there are several measuring devices in the laboratory that communicate with the host computer via the RS232 or RS485 standard, we resorted to using a multiport gateway to switch from the RS232 or RS485 serial communication standard to the USB serial communication standard. The paper presents the method of connecting measuring devices to the host computer as well as the performance of communication with this method of connection.Laboratory equipment for testing internal combustion engines

The engine testing laboratory of the Faculty of Transport and Traffic Engineering in Doboj is equipped to measure all engine parameters in accordance with the ECE83 and ECE85 standards. Integration of measurement of all parameters of the test engine and management of all actuators was realized through the AVL Puma system.The system consists of [7]:

- PUMA Open software and hardware automation system (host computer),
- Engine controller and dynamometer EMCON SW&HW,
- Drive-by-Wire 400 / E-Gas units for electronic engine throttle control,

- Safety module compliant with safety standards ISO 13849-1 and IEC 62061,
- A table for the operator and a cabinet for placing the system,

The following picture shows the displays of the PUMA shift control system.

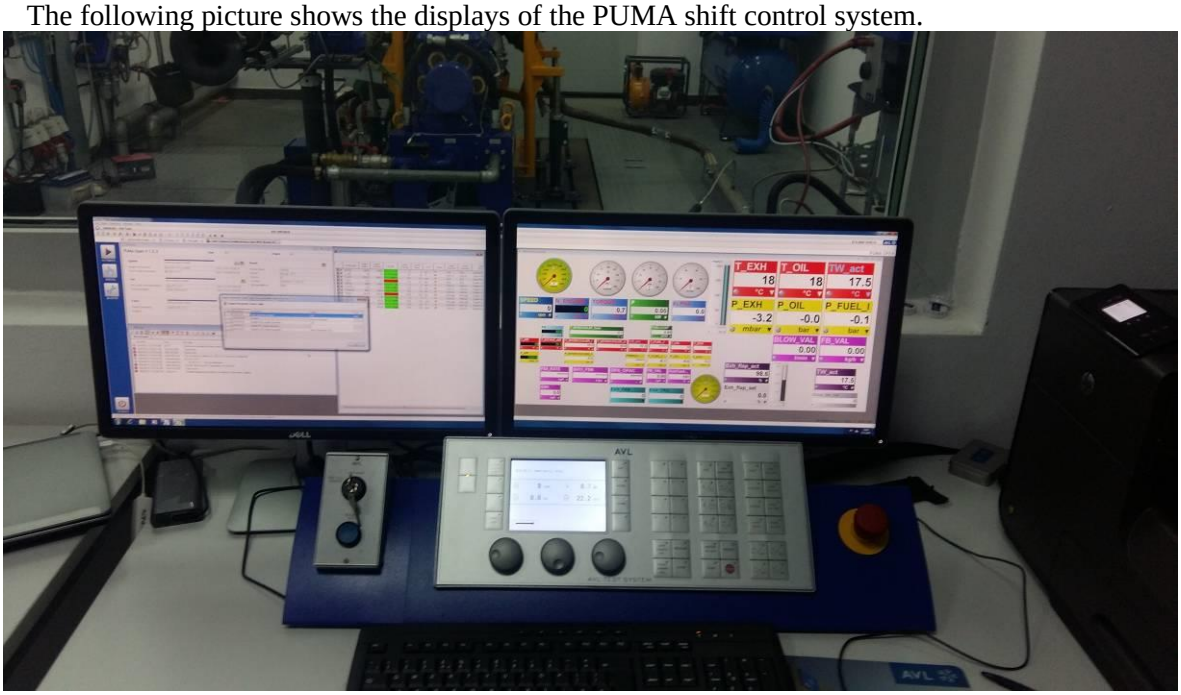


Figure 1. Display of the Puma measuring control system

The test bench work platform, based on the AVL PUMA Open automation system, is used for various purposes, from engine performance testing to vehicle testing. In this case, PUMA Open represents a platform for automating and integrating measuring equipment for engine performance testing and exhaust emission testing [5]. PUMA Open contributes to increasing productivity during testing and enables the user to control the increasingly pronounced complexity of testing tasks and test systems.

2. Implementation of a laboratory measuring system for fuel supply

Fuel consumption can be measured in the laboratory in two ways:

- Gravimetric (mass flow medium)
- Volumetric (by measuring volume flow)

For precise measurement of fuel consumption, the volumetric principle of measurement is used less often because the volume of fuel changes with the change in temperature [5]. For precise laboratory measurements, the gravimetric method (mass measurement) is mainly used. Measurement of fuel consumption and temperature in the engine testing laboratory of the Faculty of Transport and Traffic Engineering in Doboj is carried out using the "Fuel balance 733S" system manufactured by AVL GmbH. The following picture shows the system for measuring fuel consumption "Fuel balance 733 S".

Figure 2. System for measuring fuel consumption "Fuel balance 733S

Gravimetric measurement of fuel consumption is performed using the following formula:

$$M = \frac{1}{3{,}6}\int_{i=n}^{i=1} BH \cdot dt_i = M_1 - M_2 \quad (1)$$

Where is:

BH - Current fuel consumption [kg/h]

M1 - Mass of fuel in the measuring container at the beginning of the measurement [g]

M2 - Mass of fuel in the measuring container at the end of the measurement [g]

The simplified principle of operation of the "Fuel balance 733S" system is shown in the following picture.

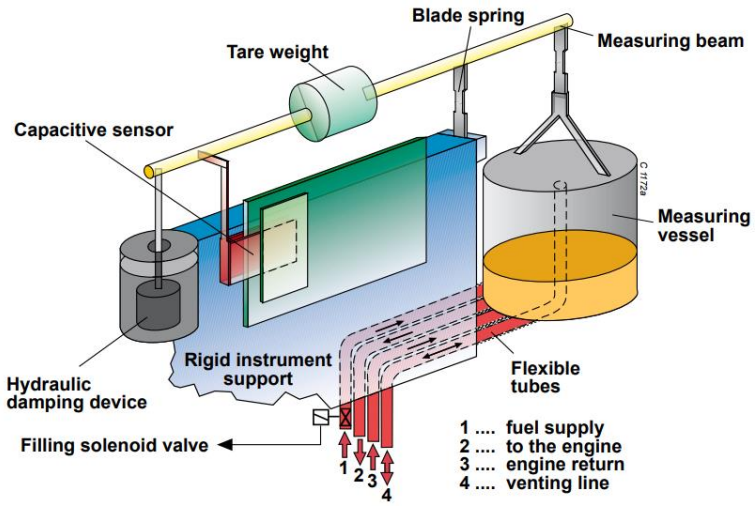


Figure 3. Simplified principle of operation of the "Fuel balance 733S" system [6]

The AVL "Fuel balance 733S" works on the gravimetric measuring principle. Thereby, fuel is supplied to the engine from a measuring vessel, the weight of which is continuously measured. Acquisition of the fuel mass in a measuring vessel that is suspended on a bending. Frictionless and hysteresis-less transmutation of mass into paths [6]. A contactless capacitive displacement pick-up transmutes the path into voltage. A microprocessor evaluates the voltage and carries out fully automatic control of measurement and calibration procedures. Display and operation are accomplished by way of a host computer.

The gravimetric measuring principle allows for a direct measurement of the fuel mass consumed. The acquisition of fuel temperature and density, which, at volumetric measurement procedures, are affected by accuracy-reducing tolerances, is not necessary.

Just as the fuel tank, the measuring vessel is equipped with an engine return line connection and a venting line, which thus provides for continuous air and vapor bubble separation in the measuring circuit (engine feed and return line). An additional separation of bubbles, which would have a negative effect on accuracy and dynamics, is not required. Bubbles are, of course, separated from the fuel supply in the measuring vessel before they get into the measuring circuit.

The capacity of 1800 g is sufficient for a continuous measurement of the consumption of approximately 95 % of all passenger cars during all internationally known car-model test procedures, such as FTP75, ECE, etc.

3. Communication between the measuring device and the host computer

The communication between the host computer and the AVL 733S is accomplished via a serial interface. Data transmission as well as starting and ending functions are exclusively performed at the command of the host computer. The host sends an instruction telegram, and the driver's aid answers with an acknowledgment telegram. The protocol, which is to be kept strictly, has been implemented according to the standards of the „Standardisierung Abgasmesstechnik“ (association of the German automobile industry). In the following, this protocol is briefly called the AK protocol.

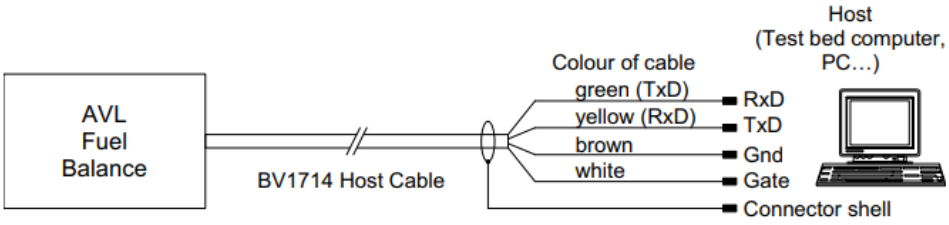


Figure 4. Connecting the measuring device "Fuel balance 733 S" to the host computer

Each telegram of the AK protocol begins with STX (Start of Text) in the first byte. The “don't care” byte can be any ASCII character. Generally, a blank or an underscore (_) is used for readability reasons. The four function bytes represent the AK command [8]. A blank comes next, followed by K and the channel number. The analyzer is a single-channel device, and because of that, the channel number is almost always 0. For delimiting the command parameters from the channel number, another blank follows. This may be followed by command parameters with a variable length. Every telegram ends with the ETX (End of Text) character.

The error status byte in the acknowledgment telegram signals if internal errors in the AVL 733S occurred. It is zero when no error appeared, and it is an unequal zero when one or more errors occurred. Every time a change in the errors happens, the error status byte is incremented by one, no matter if one or several errors disappeared or were added. If it had the value 10, it would be reset to 1. The error status byte does not indicate the real number of errors. If the analyzer does not have errors, the error status byte contains the value 0.

In general, AK commands are subdivided into three classes [8]:

- Control commands (Sxxx)
- Inquiry commands (Axxx)
- Configuration commands (Exxx)

The serial communication speed is selected using jumpers 3 and 4 on the control panel of the AVL 733S device. The following table shows the possible settings for the serial communication speed.

Table 1. Possible settings of the serial communication speed

Baud rate	Jumpers	
	3	4
2400	Off	On
4800	On	Off
9600	Off	Off
19200	On	On

Most modern computers do not have a built-in interface for communication via the RS232 protocol. Several measuring devices are integrated in the engine testing laboratory that communicate with the host computer via the RS 232 or RS 485 serial protocol. In order to enable the connection of all measuring devices that communicate via serial communication, a 16-port gateway was used for exchanging messages from serial RS 232 or RS 485 to the USB interface. The following picture shows the Edgeport/416 DB-9 gateway that translates messages between the mentioned protocols.

Figure 5. Edgeport/416 DB-9 gateway

After the installation of the gateway driver, it is necessary to confirm the configuration of the gateway ports in terms of which COM port each gateway port belongs to. The AVL 733S measuring device is connected to physical port 2 on the gateway, and serial port COM6 is assigned to it in Windows. The following picture shows the setting of communication parameters in Windows.

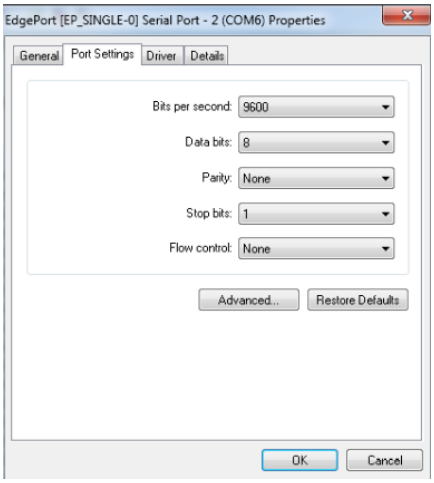


Figure 6. Communication parameters

The following picture shows the communication performance of the AVL 733S measuring device via the AK protocol with the integrated Edgeport/416 DB-9 gateway.

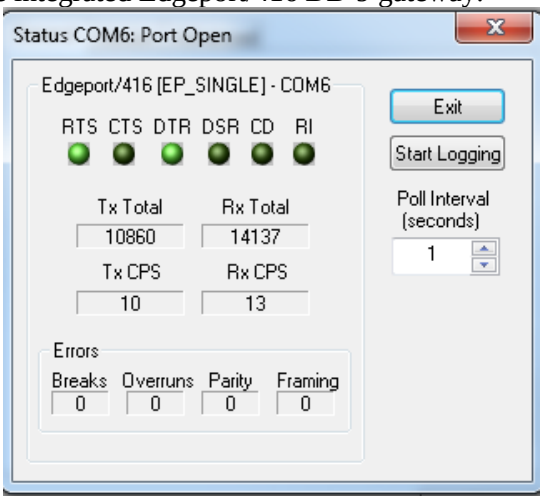


Figure 7. Communication performance

It can be seen from the previous picture that the performance of communication via the gateway measured in CPS (characters per second) units is as follows:

- When receiving messages from the measuring device AVL 733S, the speed is 13 CPS
- When sending messages from the host computer, the speed is 10 CPS.

When measuring communication parameters, there were no registered errors in data transmission.

4. Conclusions

The measurement of the mass flow of fuel is very important in laboratory tests of the engine from the point of view of determining the coefficient of excess air and determining the specific effective consumption of the engine. The gravimetric principle of mass flow measurement is used for precise measurement of fuel consumption. The paper describes the principle of operation of the laboratory device for measuring fuel mass flow, AVL "Fuel balance 733S". This measuring device belongs to the class of the most precise devices for laboratory measurement of mass flow. A partial shortcoming of the device is the connection to the host computer via the RS232 protocol for serial data transmission. The disadvantage is reflected in the fact that most modern computer workstation does not have an RS232 serial port. An even bigger problem is when you need to simultaneously connect several devices that communicate with the host computer via RS232 or RS485. The solution described in the paper enables the connection of up to 16 such devices using a gateway to the USB standard of serial communication. The above solution provides satisfactory performance for connecting this type of

equipment. This procedure can also be used when connecting similar equipment in other areas of laboratory testing.

References:

[1] Đ. Šilić, Ispitivanje motornih vozila, Veleučilište Velika Gorica, Velika Gorica 2010.
[2] M. Lambić, Dizel - motori, Tehnička knjiga, Beograd, 1998.
[3] L. E. Frenzel, Handbook of Serial Communications Interfaces: A Comprehensive Compendium of Serial Digital Input/Output (I/O) Standards, 1st Edition, Newnes, Boston 2015.
[4] B. Sklar, F. Harris, Digital Communications: Fundamentals and Applications (Communications Engineering & Emerging Technology Series from Ted Rappaport), 3rd Edition, Pearson, London 2020.
[5] G. Popović, Tehnika motornih vozila, Hrvatska obrtnička komora, POU Zagreb, Zagreb, 2006.
[6] N. Gunić, Dijagnostika elektronskih sistema motomih vozila, Infoelektronika, Niš, 2016.
[7] AVL PUMA Open 2012 — AVL733s Fuel Balance Integration, AVL List GmbH, Graz, Austria, 2012.
[8] Fuel balance - Operating Manual, AVL List GmbH, Graz, Austria, 2005.
[9] Steady state measurement AVL PUMA Open 2012, AVL List GmbH, Graz, Austria, 2012.
[10] AKprotocol for NDIR (2004) URL: https://www.mikrocontroller.net/attachment/25135/akprotocol_for_ndir_1_1_.700.pdf

Preliminary Research on the Possibilities of PPG (Photoplethysmogram) Signal Analysis of Medical Sensors and Smart Watch Sensors

Ivana Popovic[1], Sonja Djukic Popovic[2], Stefan Popovic[3] and Stevan Ivankovic[4]

[1] *University of Belgrade, Faculty of Physics, Studentski trg 12, Belgrade, Serbia*
[2] *University of Belgrade, Faculty of Mathematics, Studentski trg 12, Belgrade, Serbia*
[3] *Alfa BK University, Palmira Toljatija 3, Belgrade, Serbia*
[4] *Alfa BK University, Palmira Toljatija 3, Belgrade, Serbia*

ivapopovic2305@gmail.com; sonjica27@yahoo.com; stefan.popovic@alfa.edu.rs; stevanivankovic@gmail.com

Abstract:
The development of information and communication technologies has a deep and wide impact on medicine from many different aspects. The technology enables continuous, non-invasive measurement and monitoring of the patient's condition, which contributes to a better overview of the patient's overall condition. Some of the ways in which the advancement of information technology affects medicine are the possibility of better diagnostics and personalized treatment, because progress in data analysis and artificial intelligence enables more accurate diagnostics. Designing PPG signal processing software requires planning and implementation. At the start, we must define what should be achieved by processing the PPG signal (pulse monitoring, oxygen concentration, heart rate variability). We must have PPG data previously recorded or obtained through experiments. In this paper, data obtained with the help of medical sensors and data obtained with the help of a smart watch will be analyzed and compared. Data analysis is of high importance because it allows us to use smart watches to diagnose and monitor certain health conditions in a modern way.

Keywords:
PPG (Photoplethysmogram) signal, medical sensors and smart watch sensors

1. Basic concepts of PPG signals

The theoretical concepts of PPG signals are crucial for the development and application of this technology in medicine and other fields. Understanding how the PPG signal works allows better utilization of its potential in diagnostics and health monitoring. Photoplethysmography is based on the absorption, i.e. the reflection of light that comes out of the tissue during the cardiac cycle, which consists of two phases: the contraction of the heart chambers and the heart atria, which push blood into the blood vessels, and the reverse phase, in which the heart chamber relaxes and the chambers fill from the arteries. [1]

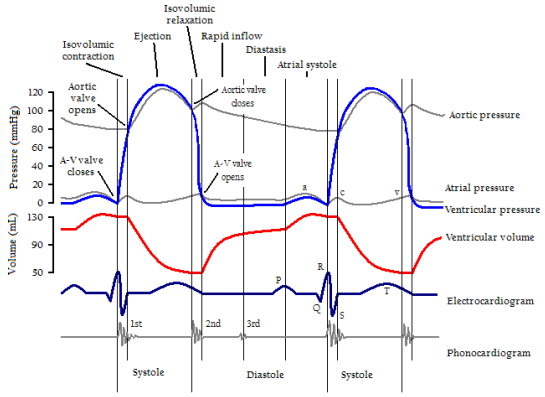


Figure 1: Change in blood volume through the cardiac cycle

Changes in the absorption and reflection of light are synchronized with the cardiac cycle, which is the cause of the characteristic waveform of the PPG signal. Accordingly, it is evident that the PPG waveform is synchronized with the pressure change in the aorta. Light is absorbed, reflected and transmitted as it passes through different layers of tissue and regardless of which way the light wave propagates, a dimmed light intensity is detected at the output using an optical sensor, which is ultimately recorded as a voltage signal [2]. The specificity of the PPG waveform is that it is formed from unabsorbed light detected by an optical sensor such as a photodiode or ambient light sensor. According to the absorption of light in the pulsatile and non-pulsatile parts of the tissue, the PPG intensity can be divided into two components: the non-pulsatile direct current component (caused by light absorption in non-pulsatile tissue such as bloodless tissue) and the pulsatile alternating current component originating from from the absorption of light in the pulsating tissue-artery (synchronous with the heartbeat). The basic idea behind the PPG signal is that light passing through tissue (most often skin) can be absorbed by hemoglobin in red blood cells or reflected by other structures. Changes in the intensity of this light during heartbeats create the PPG signal.

In addition to heart rate, blood pressure, respiration rate, and temperature, pulse oxymeter (PO) is considered to be the “fifth vital sign” of health status. Many vital organs become irreversibly damaged when not supplied with proper amount of oxygen, even for a short period. Among the body organs, the brain is by far the most sensitive to oxygen deficit. [3]

2. Optical characteristics of PPG signals

The optical characteristics of the PPG signal describe how light interacts with tissue and blood during each heartbeat. Light absorption occurs when light passes through tissue and the red blood cells (hemoglobin) in the arteries absorb the light. This absorption increases during systole (when the heart pumps blood into the arteries), which causes a decrease in the intensity of light passing through the tissue. In addition to absorption, some light is reflected from various surfaces within the tissue. Light reflection varies with changes in arterial blood volume during the cardiac cycle.

To fully understand the origin of PPG, a thorough understanding of the interaction of light and tissue is necessary. When a PPG sensor (probe) is used, it is placed next to the light source on the observed part of the body (mainly the skin), and the light that penetrates through a certain volume is modulated by the optical properties of tissue layers and sub-layers. In Figure 2, you can see a diagram of the layers and sub-layers of the skin tissue with illustrations of the depth of penetration of light of different wavelengths. Green and blue light penetrates through the papillary and epidermal layers of the dermal layer of the skin, and light with wavelengths from yellow to red penetrates deeper through the dermal vascular structure. NIR (near-infrared) light penetrates the skin to the sub-dermal layers, and mid-infrared (MIR) light penetrates to shallower tissue depths.

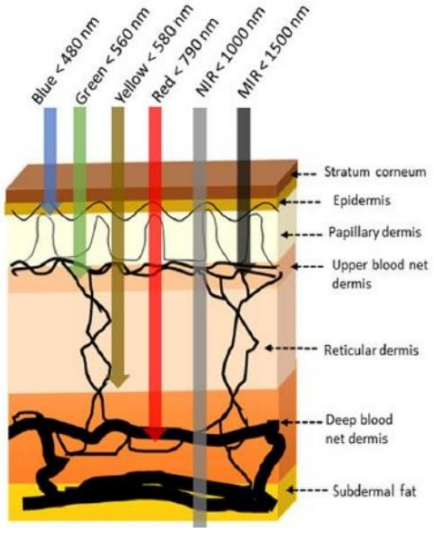


Figure 2: Diagram of skin layers and sub-layers with the depth of penetration of light of different wavelengths

Methods that use multiple simultaneously recorded PPG signals in order to detect markers of vascular pressure, are based on the estimation of the time during which the pulse wave travels between two locations on the vascular tree where the recording is performed. The mentioned method measures the time delay between pulse wave signals at two locations, which is inversely proportional to the pulse wave propagation speed. More advanced PPG model signal than networkless spectrally compressed sensor, can be used to improve spectrum performance estimation and HR performance estimation, leading to reduction of estimation errors. The time delay of the signal, on the other hand, is directly proportional to the distance between the two sensors, so the signals can be recorded at one centrally placed position, and at one peripheral one in order to enhance the time delay. The mentioned method can be used in the diagnosis of peripheral artery disease using comparative wave recording. [1]We can also add that the main advantage of optical sensors for medical applications is their intrinsic safety since there is no electrical contact between the patient and the equipment. An added bonus is that they are less susceptible to electromagnetic interference. [4]

3. Electrical circuits for using PPG signals

The technology and circuitry used for signal acquisition has changed significantly since the discovery of the PPG signal itself in the 1930s. The design of PPG assemblies consists of a light source, a detector, and supporting circuits for electronic signal processing. The original circuit design consisted of a car light bulb and a photoelectric cell used as a light source and photon detector. Today, all components can be incorporated into an integrated circuit and used in portable devices. Major advances in semiconductor and wireless technologies, with high processing power in minimal integrated circuits, have brought PPG to mass use in portable devices. These features allow continuous monitoring of health and activity. High-sensitivity sensors are advancing this technology, and more and more semiconductor manufacturers are developing complete integrated modules for PPG pulse oximetry and health monitoring. Examples of these are integrated pulse oximeters, integrated photoplethysmography and electrocardiogram modules, highly sensitive pulse oximeters and heart rate biosensors for fitness and health monitoring. On the market, manufacturers offer integrated optical chips with two green emitters and a detector (with one or two green emitters and red and infrared emitters integrated or separate wide-spectrum photodiodes, which allows the use of light from multiple wavelengths. Improved performance and processing power for signal processing, in a compact and miniature version, it enables new perspectives and a new start in PPG research. [5]

A classic PPG measurement system that uses light of one wavelength consists of a sensor with optical components (emitter and detector) and signal preprocessing blocks. Signal preprocessing blocks - transimpedance amplifier, bandpass and interference filters, signal amplifier, analog-to-digital converter (ADC), and signal processing electronics such as a microcomputer or computer. The resulting photopleismographic signal consists of a fundamental component (the "DC" component) and a pulsating component (the "AC" component). The main component of the signal is detected due to slow changes in blood volume during the cardiac cycle, caused by processes such as breathing, vasomotor activity of the autonomic nervous system and thermoregulation. The pulsating component of the signal is conditioned by the existence of an arterial pulse wave.

The device used to make the measurements is called a pulse oximeter. In the following lines, we will describe his work. A pulse oxymeter is a medical device that indirectly monitors the oxygen saturation of a patient's blood (as opposed to measuring oxygen saturation directly through a blood sample) and changes in blood volume in the skin, producing a photoplethysmograph [6].

Photoplethysmograph is a non-invasive technique that measures relative blood volume changes in the blood vessels close to the skin [7]. The pulsatile component of the PPG waveform is often called the „AC“ component and usually has its fundamental frequency, typically around 1 Hz, depending on heart rate. This AC component is superimposed onto a large quasi-DC component that relates to the tissues and to the average blood volume. This DC component varies slowly due to respiration, vasomotor activity and vasoconstrictor waves [8] [9]. The time period of each pulse is dictated by the heartbeat and the amplitude by the concentration of various constituent parts of arterial blood and path

length of light travelling through the arteries. After the systole, blood volume increases in the arteries thereby reducing the received light intensity. During diastole, blood volume in the arteries decreases and hence in increasing in light transmission. Thus the PPG signal appears pulsatile in nature at the heart rate [4] [10].

4. Beer-Lambert law

The theoretical explanation of the propagation of light through a medium was first given by Pierre Bouguer in 1729. The loss of light intensity when it spreads through a certain substance is directly proportional to the light intensity and wavelength. The relation of light attenuation is unchanged if the product of concentration and path length is constant was proved years later by August Beer. Today, this law represents a function of absorber concentration and material thickness, with scattering corrections for use in biological media. It represents the basic principle of most bio-optical and chemo-optical processes such as spectrophotometry, colorimetry, infrared spectroscopy, PPG, etc. The reduction in the intensity of the light propagating through the absorber of thickness x is proportional to the light intensity I. The radiation intensity I is the energy that the radiation beam carries per unit time per unit solid angle. [11]

$$\mathrm{dI} = -\varepsilon I c dx \tag{1}$$

If we integrate the left side with the interval of minimum and maximum light intensity, and the right side between the minimum and maximum thickness of the absorber (x=0, x=b), which light passes through, we get:

$$\int_{Io}^{I} \frac{dI}{I} = -\varepsilon c \int_{0}^{b} dx \;\rightarrow \ln\left(\frac{I}{I_o}\right) = -\varepsilon \cdot c \cdot b \tag{2}$$

Light attenuation also depends on the concentration C of the observed substance, so the decrease in light intensity can be expressed as:

$$dI = -\varepsilon I c dx \tag{3}$$

Where ε is the proportionality constant.

According to the Beer-Lambert law, the absorption of light passing through a substance is defined as the product of the molar absorption coefficient a [dm3 mol-1 cm-1], the concentration C of the substance and the optical path length d[cm]:

$$A = -\log T = -\log\frac{I}{I_o} = \log\frac{I_o}{I} = -\frac{\ln\frac{I}{I_o}}{\ln 10} \tag{4}$$

$$A = \frac{\varepsilon}{\ln 10}\, bc \tag{5}$$

$$a = \frac{\varepsilon}{\ln 10} \tag{6}$$

$$A = abc \tag{7}$$

$$T = T_1 \cdot T_2 \tag{8}$$

$$A = -\log T = -\log(T_1 \cdot T_2) = -\log T_1 - \log T_2 = A_1 + A_2 \tag{9}$$

The limitations of this law are real, so due to the high concentration of absorbing species they react with each other. The working principle of PPG is explained at a larger level by a modified Beer-Lambert law where the wavelength of light at each cardiac cycle is given importance, and with each volume change within the observed pulsating portion of the tissue between systole and diastole, the absorption of light varies with time resulting in to the continuous PPG signal. The focus of PPG signal research has shifted more to the area of development of PPG sensors and signal analysis, and over time it has moved away from the research of fundamental characteristics, i.e. PPG signal influence factors. The advancement of technology, electronics and all resources for processing and visualizing signals, lead to a more thorough understanding of information related to the physiological origin and influencing factors. [11]

5. Obtaining representative samples

The data that were compared were collected in the period February-March 2023. The data in question is taken from the KY-018 Fotowiderstand Modul sensor and the data from the Xaomi Watch S1 Active smart watch. Both devices measure in the same way using Schottky photodiodes. Schottky photodiodes are key detection devices due to their easy and inexpensive fabrication process and simple device architecture [12,13,14]. Among these devices, Si-based Schottky photodiodes have drawn particular attention as they present an excellent platform for read-out circuits and large-scale photonic integration [12,14]. Currently, novel semi-metal two-dimensional materials (2DMs) are being used intensively to construct 2DMs/Si van der Waals (vdW) Schottky photodiodes, owing to their unique advantages such as their dangling bonds and high transmittance [12,15,16]. Interestingly, high barrier heights and high responsivity can be achieved from graphene (Gr)/Si and MXene/Si vdW Schottky photodiodes, which can be used to create high-performance and low-cost Si-based broadband photodetectors (PDs) [12,17,18]. As is well known, Si-based photodiodes have been used as key components for devices with applications in near-infrared (NIR) light detection. Nowadays, NIR PDs are attracting much attention in the field of noninvasive medical diagnosis and treatment, including specific molecular detection, imaging, biological detection, and light-based diagnostics and therapeutics [12,19,20,21].

The compared data were analyzed in the python programming language using the Jupyter Notebook working environment. The signal was additionally filtered by Wavelet transforms. The sensor measured the amount of oxygen in the blood and monitored the heart rate of the user in the age of 50. The user used both the sensor and the smart watch at the same time.

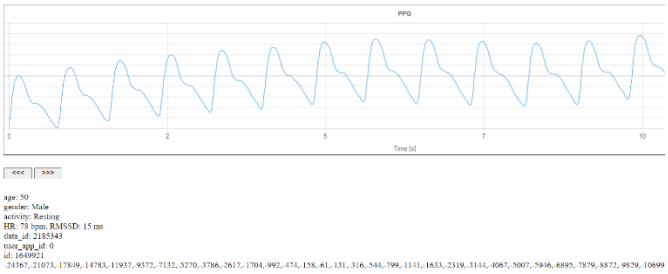


Figure 3: Medical sensor

The data that was analyzed had different values from both the sensor and the smartwatch. The smartwatch showed an average heart rate while wearing the watch in rest periods of 69 beats, while the sensor showed an average of 9 beats 9 beats higher.

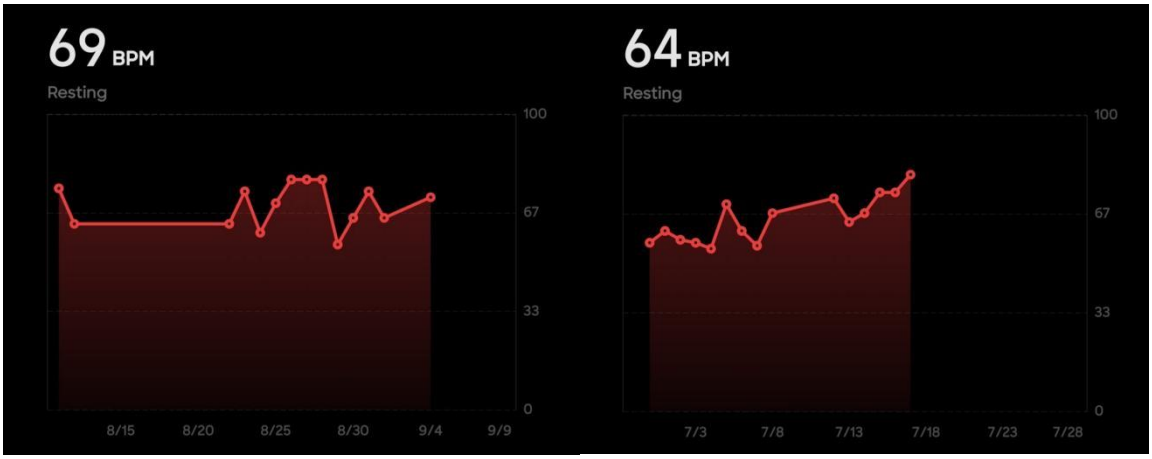


Figure 4: Smart watch

The analysis of the given samples concluded that the accuracy of the sensor for measuring the PPG signal and the smart watch can vary significantly depending on the manufacturer, model and purpose. The medical sensor still gives us more accurate data, which is extremely important for diagnostic purposes, while from the sample itself we see that smart watches have different functions and are not necessarily calibrated for medical precision. Smart watches, while capable of tracking PPG signals, do not have the same level of calibration and updates as medical devices. But for early detection of some irregularities and for regular monitoring of health, regular wearing of a smart watch has great advantages. The PPG data is stored in the cloud in the smart watch application and thus enables the monitoring of heart activity over a long period of time. Changes in the PPG signal can then indicate health problems, such as heart rhythm disorders or hypertension. If an anomaly has already been diagnosed, the smart watch can enable 24-hour monitoring of the patient's condition.

The importance of this testing is to provide us with detailed information about the accuracy of the information provided by smart watches in order to diagnose certain health conditions in patients. In the data analysis, we see that the difference in the data collected is significant in certain situations, such as for example the conditions of PPG signal measurement. For example, it has been observed that hand movements or ambient light can greatly affect the accuracy of the sensors in the smart watch. The reason for this may be in the testing of the device itself, which is present to a lesser extent than with medical sensors for monitoring PPG signals.

6. Conclusion

Medical sensors allow us to measure accuracy and reliability, which is essential, although on the other hand smart watches, although not at the same level of precision as a medical device, can provide a practical solution for monitoring certain health conditions in patients. The combination of sensors and an application that offers insight into the patient's current condition can improve individual monitoring and enable individual research that includes the analysis of certain parameters. When using smart watches for these purposes, the data collection conditions must be taken into account, as some measurement errors may be caused by certain hand movements or the influence of light on the sensor itself. If these interferences were to be removed, the precision of the measurements would largely match the precision of the medical sensors themselves. In addition, we must state that when using medical sensors, we have a certain preparation of the patient, which usually includes calming the patient and placing them on certain parts of the body, which can be very inconvenient and uncomfortable for some patients.

The results derived from the PPG signal motivated the researchers for further work. Here, the need for an additional critical review of the literature was seen, which will help eliminate the conceptual and methodological shortcomings of the research itself. In addition, it is necessary to provide meaningful evidence on the effectiveness of PPG measurement. Finally, there is also the difficulty in finding a few volunteers who would wear a medical sensor and a smart watch and thereby help further research.

References:

[1] Charlton, P. (2019). Modeling arterial pulse waves in healthy aging: a database for in silico evaluation of hemodynamics and pulse wave indexes. American journal of physiology.

[2] K, J. V., Sandhya, T., & Radhakrishnan, P. (2008). Non-invasive Studies on Age Related Parameters Using a Blood. MEASUREMENT SCIENCE REVIEW.

[3] R.W.C.G.R. Wijshoff1, J. Veen, O. Such, R.M. Aarts, J.W.M. Bergmans, "Towards a hemodynamic model to characterize inaccuracies in finger pulse oxymeter". The Netherlands Philips Research, Biomedical Sensor Systems Group, The Netherlands, November 9-10, 2009.

[4] Bagha, S., & Shaw, L. (2011). A real time analysis of PPG signal for measurement of SpO2 and pulse rate. International journal of computer applications, 36(11), 45-50.

[5] Zhilin Zhang, (2015) Photoplethysmography-Based Heart Rate Monitoring in Physical Activities via Joint Sparse Spectrum Reconstruction, PUBLISHED IN IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING

[6] Zhang Jun-feng, Wang Yan-yan,"Design for high data acquisition system based on PCL-816 and LabVIEW", Machinery Design &Manufacture, Vol7, pp.64-66, 2007.

[7] Rajet Krishnan, Balasubramaniam (Bala) Natarajan, Steve Warren, "Two-Stage Approach for Detection and Reduction of Motion Artifacts in Photoplethysmographic Data", IEEE transactions on biomedical engineering, vol. 57, no. 8, August 2010.

[8] M. Nitzan, A. Babchenko, B. Khanokh and D. Landau, "The variability of the Photoplethysmography Signal – A potential method for the evaluation of the autonomic nervous system," Physiol. Meas. Vol.19 pp. 93-102, 1998.

[9] M. Mukunda Rao, V.S. Murthy and S. Ramamoorthy, "Non invasive study of Rhythms and arrhythmia using optical sensors and their biological implications," Int. Conf. On Biomed. Engg. Medical. Phys – World Medical Congress, Chicago, USA. July 2000.

[10] John Allen, "Photoplethysmography and its application in clinical physiological measurement" Physiol. Meas. 28, R1–R39, 2007.

[11] Matija Pauković,(2022) TEHNIKE I OBRADA SIGNALA FOTOPLETIZMOGRAFIJE, Sveučilište u Splitu , Fakultet elektrotehnike, strojarstva i brodogradnje

[12] Xu, M., Cui, Y., Zhang, T., Lu, M., & Yu, Y. (2023). PbS QD-Coated Si Micro-Hole Array/Graphene vdW Schottky Near-Infrared Photodiode for PPG Heart Rate Measurement. Sensors, 23(16), 7214. https://doi.org/10.3390/s23167214

[13] Tao, A.; Yao, T.; Jiang, Y.; Yang, L.; Yan, X.; Ohta, H.; Ikuhara, Y.; Chen, C.; Ye, H.; Ma, X. Single-dislocation Schottky diodes. Nano Lett. 2021, 21, 5586–5592.

[14] Chang, K.E.; Yoo, T.J.; Kim, C.; Kim, Y.J.; Lee, S.K.; Kim, S.Y.; Heo, S.; Kwon, M.G.; Lee, B.H. Gate-controlled graphene–silicon Schottky junction photodetector. Small 2018, 14, 1801182.

[15] Yang, W.; Chen, J.X.; Zhang, Y.; Zhang, Y.J.; He, J.H.; Fang, X.S. Silicon-compatible photodetectors: Trends to monolithically integrate photosensors with chip technology. Adv. Funct. Mater. 2019, 29, 1808182.

[16] Liu, C.; Guo, J.; Yu, L.; Li, J.; Zhang, M.; Li, H.; Shi, Y.; Dai, D. Silicon/2D-material photodetectors: From near-infrared to mid-infrared. Light Sci. Appl. 2021, 10, 123.

[17] Kang, Z.; Ma, Y.; Tan, X.; Zhu, M.; Zheng, Z.; Liu, N.; Li, L.; Zou, Z.; Jiang, X.; Zhai, T.; et al. MXene–silicon van der Waals heterostructures for high-speed self-driven photodetectors. Adv. Electron. Mater. 2017, 3, 1700165.

[18] Wan, X.; Xu, Y.; Guo, H.; Shehzad, K.; Ali, A.; Liu, Y.; Yang, J.; Dai, D.; Lin, C.; Liu, L.; et al. A self-powered high-performance graphene/silicon ultraviolet photodetector with ultra-shallow junction: Breaking the limit of silicon? NPJ 2D Mater. Appl. 2017, 1, 4.

[19] Song, W.; Liu, Q.; Chen, J.; Chen, Z.; He, X.; Zeng, Q.; Li, S.; He, L.; Chen, Z.; Fang, X. Interface engineering Ti3C2 MXene/silicon self-powered photodetectors with high responsivity and detectivity for weak light applications. Small 2021, 17, 2100439.

[20] Song, L.; Xu, E.; Yu, Y.; Jie, J.; Xia, Y.; Chen, S.; Jiang, Y.; Xu, G.; Li, D.; Jie, J. High-barrier-height Ti3C2Tx/Si microstructure Schottky junction-based self-powered photodetectors for photoplethysmographic monitoring. Adv. Mater. Technol. 2022, 7, 2200555.

[21] Ollearo, R.; Ma, X.; Akkerman, H.B.; Fattori, M.; Dyson, M.J.; van Breemen, A.J.J.; Meskers, S.C.J.; Dijkstra, W.; Janssen, R.A.J.; Gelinck, G.H. Vitality surveillance at distance using thin-film tandem-like narrowband near-infrared photodiodes with light-enhanced responsivity. Sci. Adv. 2023, 9, eadf9861.

ChatGPT for EFL Teachers and Students

Lela Ivanovska[1,]

[1]*University St Kliment Ohridski, 1 Maj nn, 7000 Bitola, R. North Macedonia*

lela.andreevska@uklo.edu.mk

Abstract:
The chatbot ChatGPT, which is powered by artificial intelligence, has sparked heated discussions about the potential rise of academic misconduct. ChatGPT "interacts in a conversational way" and can "answer follow-up questions." This implies that students can participate in conversations that prompt meaning-based discussions. Additionally, it can give learners quick feedback, highlighting their mistakes and any gaps in their language proficiency. The comprehensible input hypothesis states that language acquisition occurs when learners are exposed to input that is just one level above their current language proficiency. The comprehensible output hypothesis states that producing language is necessary for success in second language acquisition. The present paper provides a selection of some practical ideas, activities and prompts templates proposed by Nik Peachey in his book "ChatGPT in the Language Classroom" such as: Guest speaker Q&A, Creative writing, Job interviews and applications, and Exploring registers and style. Presented activities will help EFL teachers how to adapt them to use ChatGPT with their own students.

Keywords:
ChatGPT, EFL teachers, student

1. Introduction

The way we learn and communicate is changing at a rapid rate as technology advances. The launch of ChatGPT, a language model developed by OpenAI that has the potential to transform how teachers' approach EFL instruction, is one particularly interesting development in this field [1]. ChatGPT's ("generative, pre-trained transformer")" impact on society may be profound and enormous, similar to the transformation generated with the use of computers. One of the most sophisticated chatbots operated by AI, ChatGPT was created by the Microsoft-backed startup OpenAI and released in November 2022. This technology enhances digital assistants' usage of simple automatic tasks, automated speech recognition, natural language processing, and fundamental AI. Although people have already started to make assumptions about its possible effects, nobody has yet experienced its full potential or strength. This artificial intelligence technology is currently in its early stages and will develop further intelligence. ChatGPT is made to have chats back and forth with users. Despite the fact that chatbots have been employed in educational contexts since the early 1970s [2,3], ChatGPT's capacity to convincingly imitate human interaction creates a brand-new and fascinating field in EFL learning. The paper aims to provide a selection of some practical ideas, activities and prompts templates proposed by Nik Peachey in his book "ChatGPT in the Language Classroom" such as: Guest speaker Q&A, Creative writing, Job interviews and applications, and Exploring registers and style. Presented activities will help EFL teachers how to adapt them to use ChatGPT with their own students.

2. What is ChatGPT?

In its most basic form, ChatGPT is an OpenAI language model that can comprehend, produce, and react to human language. It can be used for many different natural languages processing tasks, including text production, language translation, text summarization, and question answering [4]. It was trained on a sizable dataset of text. The model is able to comprehend the context of a conversation or a text and give a natural and logical response. Additionally, it is capable of producing

text that is similar to the input in terms of both style and content. As a result, ChatGPT can be used for a variety of purposes, including language learning [5].

This is how ChatGPT describes itself: "ChatGPT is a large language model developed by Open AI, based on the GPT - Generative Pre-trained Transformer - archi-tecture. It is designed to generate human-like responses to text-based inputs, by using deep learning techniques, specifically unsupervised machine learning. Its training data includes a vast amount of text from a wide variety of sources, such as books, art-icles, and web pages, which allows it to generate responses on a wide range of topics. It is capable, of understanding natural lan-guage and generating contextually appropriate responses to questions or statements, making it a useful tool for communication and research purposes [6].

In line of the definition, it is clear that ChatGPT is much more than just a tool for study and conversation. It can function as a collaborator when producing factual and artistic work, as well as a translator, teacher, advisor, and even a friend.

3. The ChatGPT interface

The ChatGPT interface appears to be pretty simple at first. You can enter a prompt in a field, and the ChatGPT response will show up underneath the response. ChatGPT analyzes your previous conversations in the conversation trail before responding to you. This connects all of your interactions into a logical conversation. The major components of the ChatGPT interface are numbered in this figure. In this image you can see the major components of the ChatGPT interface numbered [6], Figure 1.

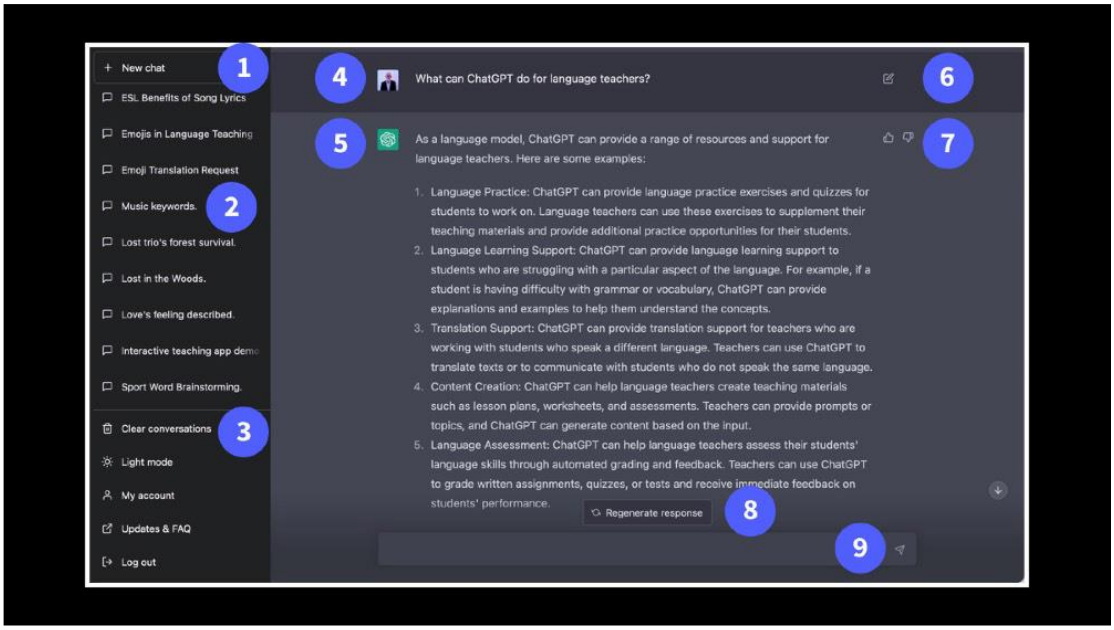


Figure 1: Major components of the ChatGPT interface

1. New Chat -Click on this button when you want to start a new chat about a different topic.
2. Chat History -This is where you find your previous chats with ChatGPT. Just click on these to return to and continue the chat. You can also delete individual chats by clicking on them and then clicking the delete button.
3. Clear Conversations -This button clears all your past conversations in your chat history. Only click this if you are sure you won't want to go back to any of these.
4. Prompt -This is the text or question that you type in ChatGPT to get a response.
5. Response History- This small counter will appear if you regenerate a number of responses to the same prompt. You can use it to scroll back through previous versions of the response.
6. Prompt Edit -This button will appear at the end of your prompt if you place the cursor above it. You can click on it to edit and resubmit your prompt. When you submit the edited prompt, a new

response will be generated. Any responses or prompts that come after it will be deleted when the new response is generated.

7. Feedback -You can click these thumb icons to send either positive or negative feedback. Clicking the thumbs up icon will open a field where you can tell Open AI what you like about the response. Clicking on the thumbs-down icon will enable you to tell Open AI what was wrong with the response. Sending feedback can help Open AI to improve ChatGPT.

8. Regenerate Response - Click on this button to get ChatGPT to regenerate your response. ChatGPT tends to give different responses each time you prompt it, so clicking this button is likely to change the response. You may well get a better response, or one that you prefer, by clicking here. You can get back to previous responses by clicking on the response history.

9. Prompt Field -This is where you type your prompts for ChatGPT. These can be questions, statement, instructions or requests. ChatGPT will respond to these prompts.

4. What is a PROMPT?

A prompt is a piece of text, such as a question, that users type into ChatGPT in order to get a response. These prompts can be very simple, or they can be quite complex. The prompts are fundamental to making ChatGPT work well for you. Here are a few examples of prompt scripts teachers or students may want to use regularly,[6].

Prompt - Use this prompt to get ChatGPT to translate to English for you.

Translate any text I enter here into English.

Prompt - Use this prompt to get ChatGPT to transcribe text into phonemic symbols.

Transcribe any text I enter here into phonemic script.

Prompt - Use this prompt to get ChatGPT to check your spelling and grammar.

Check my spelling and grammar.

Prompt - Use this prompt to get ChatGPT to give you feedback on your written text and suggest improvements.

Please read my text and suggest ways I can improve it.

Prompt - Use this prompt to get ChatGPT to create a word list with definitions from any text you want to use.

Create a wordlist of the 10 most difficult words from any text I enter. Add definitions and phonemic script to the list.

Prompt - Use this prompt to get ChatGPT to be your supportive English friend.

I want you to act as my friend. I will tell you what is happening in my life, and you will be supportive and give emotional support. Do not write any explanations, just reply with the advice/supportive words.

We can conclude that writing prompts is a skill that teachers need to develop in order to get the best responses from ChatGPT. In other words, the above written prompt scripts are just tips that can help teachers understand how to expand the potential for getting a response for translation to English language, transcribing text into phonemic symbols, spelling and grammar, getting feedback from written text, creation a word list with definitions and getting ChatGPT as a supportive friend.

5. ChatGPT in the English language classroom

ChatGPT has a significant impact on EFL learning in the classroom. It gives personalized instruction to students or learners. Students can learn more quickly and effectively with the model's assistance since it can comprehend the context of a conversation or text and adapt its responses to each student's unique requirements and talents. For students who are having difficulties or who have different ways of learning, it can be extremely beneficial

ChatGPT can also influence EFL learning in the classroom by giving real-time feedback on language usage. With ChatGPT, for instance, students can practice speaking or writing in a foreign language while receiving immediate feedback on their grammar, vocabulary, and sentence structure. This can save learning time and speed up the process of learning by helping students in quick identifying and correcting their errors. This can help students to identify their mistakes quickly, and

accelerate the learning process [7]. Additionally, ChatGPT can be used to create resources for language learning, such as flashcards, tests, and exercises, as well as summaries of texts and articles, which can aid students in understanding the subject matter and speed up learning. Furthermore, ChatGPT can be applied to language assessment. A learner's language proficiency can be assessed using the approach by analyzing their writing and speech. Compared to more conventional approaches like written tests and oral exams, this might offer a more precise and effective means to evaluate language proficiency [8]. Overall, ChatGPT can have a big impact in the classroom. It can offer individualized training, immediate feedback, and a variety of practice materials that can ultimately result in more successful language instruction. It can also help to improve the efficiency, accuracy, and effectiveness of the assessment process. The following practical activities and suggested prompts template that EFL teachers can adapt and try with their students are proposed by [6], a prominent figure in the realm of the digital learning in his book "ChatGPT in the Language Classroom". This is just an idea of how to use ChatGPT in the EFL learning classroom.

6. Practical activities for using ChatGPT

- **GUEST SPEAKER Q&A**

Here are some different ways to get guest speakers into the classroom using ChatGPT. [6].

Prompt 1 - Use this prompt when you want students to interact with ChatGPT as the famous person.

Act as {famous person} and answer the questions I ask. My first question is {first question}

Example: Act as Marilyn Monroe and answer the questions I ask. My first question is Where did you grow up? Figure 2.

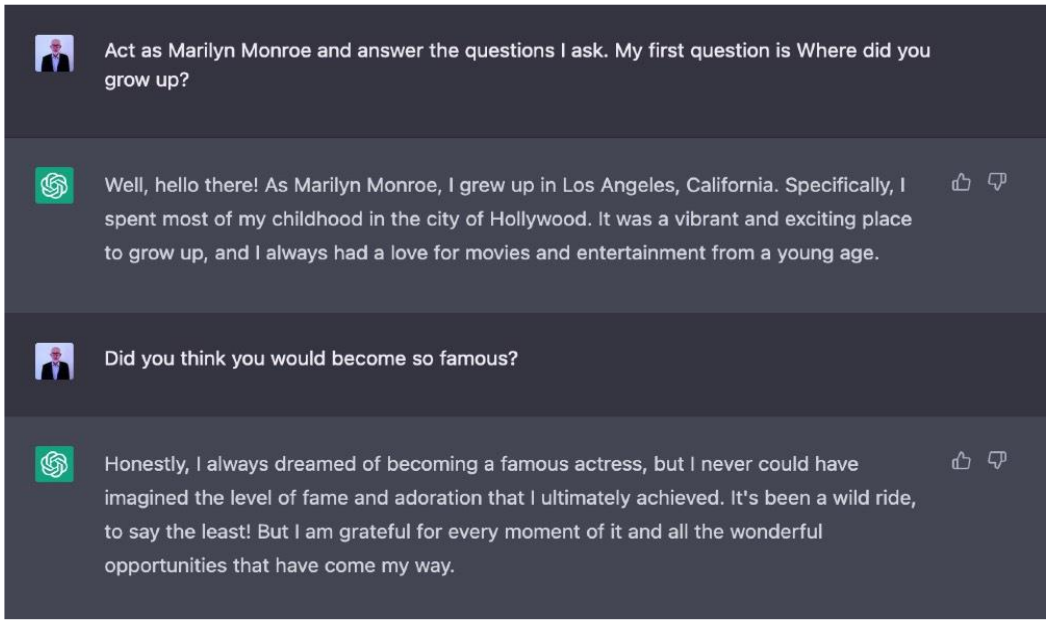


Figure 2: Act as Marilyn Monroe and answer the questions I ask. My first question is Where did you grow up?

Prompt 2 - This prompt can be used when you want to produce questions to ask a famous person.

Write {number} questions to ask {famous person}

Example: Write 10 questions to ask Charles Dickens, Figure 3.

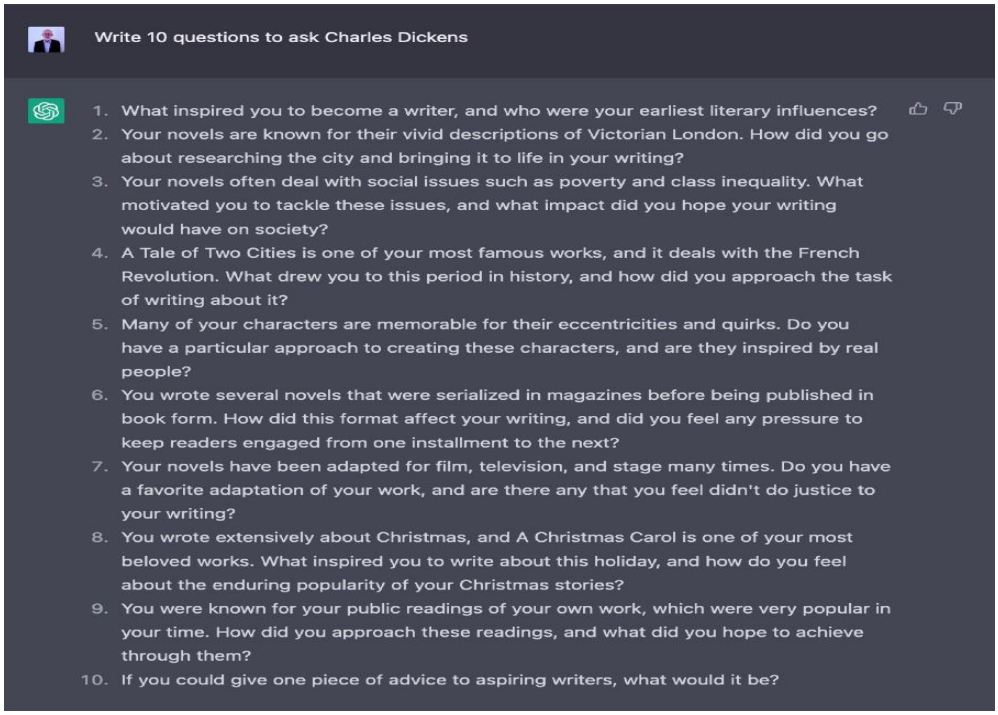


Figure 3: Write 10 questions to ask Charles Dickens

Prompt 3 - This prompt can be used when you or the student role plays being the famous person and asking questions.
I am {famous person}. Ask me a question. When I answer, ask me another question. Continue to ask questions each time I answer.
Example: I am Abraham Lincoln. Ask me a question. When I answer, ask me another question. Continue to ask questions each time I answer, Figure 4.

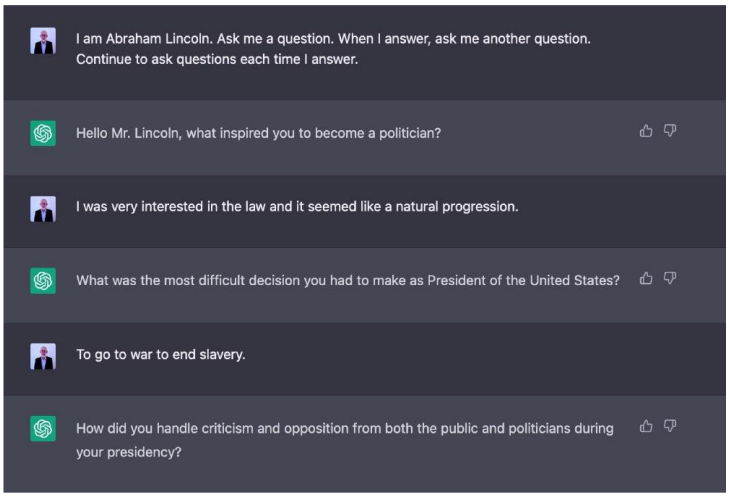


Figure 4: I am Abraham Lincoln. Ask me a question. When I answer, ask me another question. Continue to ask questions each time I answer

These prompts proposed by [6] are good suggestions when teachers want to invite quests into the classroom. Having in mind that it is very difficult to arrange and manage, using ChatGPT teachers can invite any famous person and students will be able to ask them questions. The benefits are the following: real-world context, cultural insights, engaging and motivating.

- **CREATIVE WRITING**

ChatGPT can help students create plots and characters to base their stories on.

Prompt 1 - This prompt produces a plot outline about any topic teachers or students choose.
Create a plot outline for a story about {topic}
Example: Create a plot outline for a story about travelling to the North Pole, Figure 5.

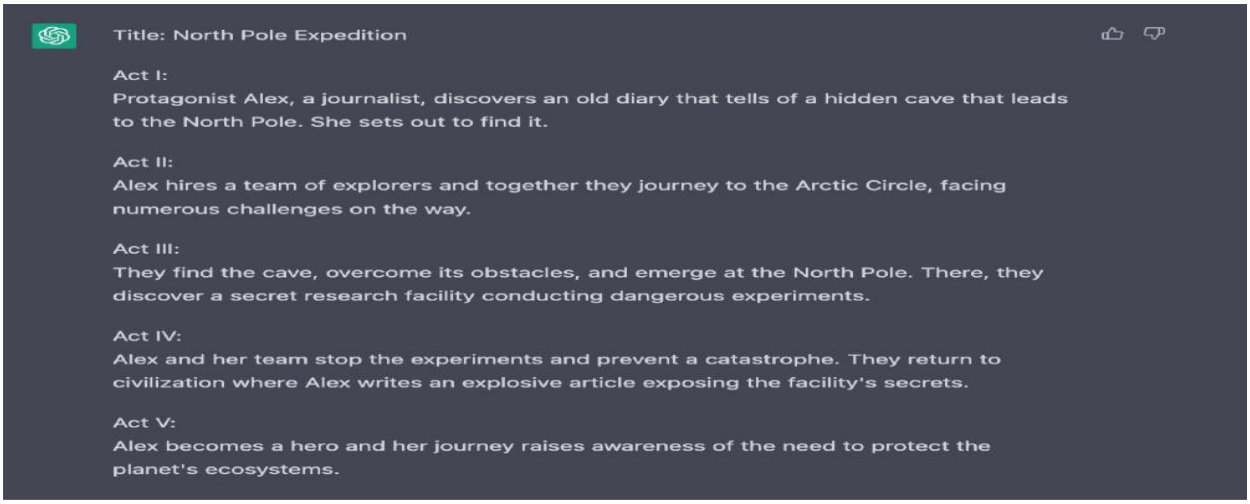


Figure 5: Create a plot outline for a story about travelling to the North Pole

Prompt 2 - This prompt produces a plot outline about any topic you or your students choose and also includes an underlying theme.
Create a plot outline for a story about {topic} include the theme of {theme}
Example: Create a plot outline for a story about travelling to the North Pole and include the theme of romance, Figure 6.

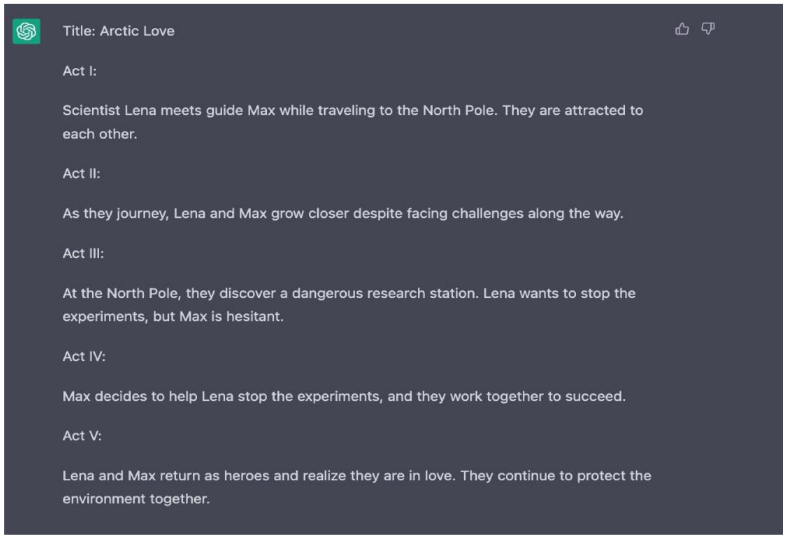


Figure 6:Create a plot outline for a story about travelling to the North Pole and include the theme of romance.

In sum, creating stories and characters can be really difficult and take a lot of imagination. This is difficult for many students in their native language and even more so in a second language. But these prompts encourage students write creatively and the benefits are the following: Developing communication skills, Improving writing skills, Boosting confidence: Cultivating creativity, Enhancing cultural awareness. ChatGPT can help students create plots and characters to base their stories on.

- **JOB INTERVIEWS AND APPLICATIONS**

ChatGPT suggests the following prompts to act as an interviewer and ask questions:

Prompt 1 - Use this prompt to get ChatGPT to take on the role of interviewer and interact with students by asking them questions about a specific role. Act as an interviewer. I will be the candidate, and you will ask me the interview questions for the {job title}. I want you to only reply as the interviewer. Do not write all the conservation at once. I want you to only do the interview with me. Ask me the questions and wait for my answers. Do not write explanations. Ask me the questions one by one like an interviewer does and wait for my answers. My first sentence is "Hi"
Example: Act as an interviewer. I will be the candidate, and you will ask me the interview questions for the ESL teaching position. I want you to only reply as the interviewer. Do not write all the conservation at once. I want you to only do the interview with me. Ask me the questions and wait for my answers. Do not write explanations. Ask me the ques-tions one by one like an interviewer does and wait for my answers. My first sentence is "Hi", Figure 7.

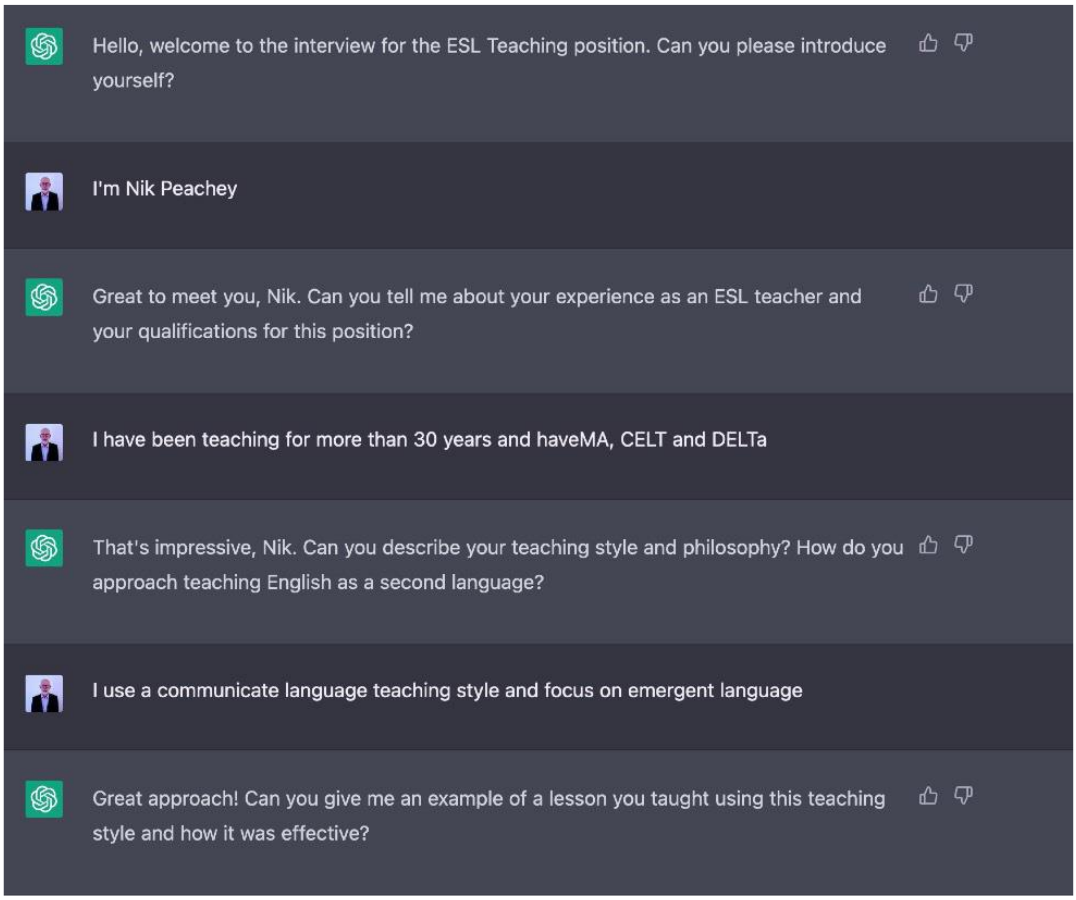


Figure 7: Act as an interviewer

Another suggested prompts are the following:
Prompt 2 - Use this prompt to get ChatGPT to produce a list of questions for students to answer, and then evaluate and give feedback on each answer.
Hello, ChatGPT. Can you please help me practice for a job interview? I'm applying for {job title} and would like to practice answering some common interview questions. Could you ask me a few questions and give me feedback on my answers?
Example: Hello, ChatGPT. Can you please help me practice for a job interview? I'm applying for the ESL Teaching position and would like to practice answering some common interview questions. Could you ask me a few questions and give me feedback on my answers?
Prompt 3 - Use this prompt to get ChatGPT to produce a job descrip-tion for any job you specify
Please write a job description for the role of {position}.
Example: Please write a job description for the role of ESL teacher.
Prompt 4 - Use this prompt to get ChatGPT to create an example letter of application.
Write a letter of application for {post or position}.

Example: Write a letter of application for the post of ESL Teacher

For job interviews and applications activity, ChatGPT can help students face a number of interview situations in a second language that could be stressful for them. The suggested prompts can help them by becoming an interviewer and giving them the opportunity to spontaneously answer questions. Here are some of the reasons ChatGPT suggests for doing this: Accurate language model, Conversational practice, Instant feedback and Flexible schedule.

- **EXPLORING REGISTER AND STYLE**

ESL students can understand how to use language appropriately in different situations by learning about style and register. Here are some things that ChatGPT can do.

Prompt 1 – You can use this prompt to produce a formal text on any topic. Then use the follow-up prompt to get GTP to create a parallel text in an informal style.

Write a paragraph about {topic} in a formal style.

Example: Write a paragraph about UK Culture in a formal style, Figure 8.

Figure 8: Write a paragraph about UK Culture in a formal style

Follow-up prompt Change the text to an informal style

Figure 8 shows us how ChatGPT can help ESL students enhance their writing abilities and their overall language competency while allowing them to communicate with native English speakers and other non-native English speakers more effectively and prevent misunderstandings.

7. Conclusions

The purpose of this paper was to demonstrate the usefulness of some practical ideas, activities and prompts templates for using ChatGPT in the EFL classroom. Namely, the emphasis was placed on the description of selected ways in which ChatGPT can be used in EFL learning and teaching. Of course, these activities can be adapted according to the needs and interests of students and the objectives of the lesson.

In conclusion, ChatGPT is a potent language model that has the ability to completely transform the way we approach learning English as a second language. It is a powerful tool for giving individualized teaching, creating language learning materials, and analyzing a learner's writing and speech because of its capacity to comprehend, produce, and react to human language in a natural and fluent manner. The possibilities and excitement of ChatGPT in EFL learning are enormous. It is crucial to remember that while ChatGPT might be a useful tool for language learning, it cannot replace the human language teachers. It is crucial to prevent the loss of human language teachers as a

result of the use of ChatGPT since the human touch is an integral part of language acquisition. It is also necessary to address any potential ethical and privacy issues connected to the use of ChatGPT.

References:

[1] Molly, ChatGPT and the English Classroom: What Teachers Need to Know about AI-generated Essays. Hopefully Home, (2022), http://hopefullyhome.com/chatgpt-and-the-english-classroom-what-teachers-need-to-know-about-aigenerated-essays/#worried

[2] W. Huang, K. F. Hew and L. K. Fryer, Chatbots for language learning – are they really useful? Asystematic review of chatbot-supported language learning. Journal of Computer AssistedLearning 38(1), 237–257, (2021), https://onlinelibrary.wiley.com/doi/epdf/10.1111/jcal.12610

[3] N. Kim NY, A study on the use of artificial intelligence chatbots for improving Englishgrammar skills. Journal of Digital Convergence 17(8): 37–46, (2019).

[4] L. Kohnke, L2 learners' perceptions of a chatbot as a potential independent language learningtool. International Journal of Mobile Learning and Organisation 17(1/2): 214–226, (2023), DOI: 10.1504/IJMLO.2023.10053355.

[5] J. H. Lee, H. Yang H, D. Shin, H. Kim, Chatbots – technology for the language teacher. ELT Journal74(3): 338–344, (2020), https://doi.org/10.1093/elt/ccaa035

[6] N. Peachey, ChatGPT in the Language Classroom, Peachey Publications, (2023).

[7] J. Rudolph, S. Tan, S. Tan,ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? Journal of Applied Learning and Teaching, 6(1), (2023). DOI: https://doi.org/10.37074/jalt.2023.6.1.9

[8] X. Xu, W. C. H Hong, Y. Zhang, H. Jiang, J. Liu, Learning Paths design in Personal Learning Environments: The Impact on Postgraduates' Cognitive Achievements and Satisfaction. Innovations in Education and Teaching International., (2023)https://doi.org/10.1080/14703297.2023.2189603

Design Considerations for a Generic Graph Database in Archival Document Management

Ilija Hristoski[1], Jelena Stojanov[2] and Željko Stojanov[2]

[1] *"St. Kliment Ohridski" University – Bitola / Faculty of Economics – Prilep, Prilepski Braniteli St. 143, MK-7500 Prilep, N. Macedonia*
[2] *University of Novi Sad / Technical Faculty "Mihajlo Pupin", Đure Đakovića St., 23000 Zrenjanin, Serbia*

ilija.hristoski@uklo.edu.mk; jelena.stojanov@uns.ac.rs; zeljko.stojanov@uns.ac.rs

Abstract:
In the digital age, managing archival documents in both physical and digital formats poses a significant challenge due to the sheer volume and complexity of historical records. Traditional relational databases often struggle to capture the intricate relationships and interconnectedness present in archival data. This research paper explores the utilization of property graphs, for creating a robust and efficient graph database system aimed at storing, organizing, and retrieving archival documents. By representing documents and their associated metadata as nodes and relationships in a graph, we propose a generic graph database design that can serve as a comprehensive framework for modeling and querying interconnected historical data, supporting efficient and adaptable archival document management. The paper contributes insights into designing graph databases that harmonize with the intricate nature of archived documents while catering to the evolving needs of modern archival practices.

Keywords:
Graph databases, archival document management, database design, graph database schema, graph data model, data modeling

1. Introduction

Archives consist of document collections or 'records' chosen for permanent retention due to their significance as proof or as sources for historical and other forms of research. Records emerge from the actions of individuals and organizations; they hold a functional role while in active circulation, and a portion of them are subsequently chosen and safeguarded as components of an archival compilation [1]. Archive collections usually possess unique attributes, underscoring the crucial need for their proper upkeep. Despite the fact they consist of historical records, archives provide an immense source of information. As Taryn Simon* once remarked, "*Archives exist because there's something that can't necessarily be articulated. Something is said in the gaps between all the information*." They demand meticulous storage and administration to ensure their safeguarding and persistence for present and future utilization. Besides specialist collecting institutions, which are also called 'archives', including national and local archives and record offices, archives are also kept by other institutions, e.g. museums, libraries, universities, faculties, schools, businesses, charities, arts organizations, and community groups. All of these subjects usually hold and maintain their own institutional records. Archives are highly valuable sources of information for research to nations and regions, organizations, communities, and individuals since they provide proof of historical occurrences. They narrate tales, document individuals and identities, and organizations' activities, constituting our recorded memory and being a vital component of our community, cultural, official, and unofficial history. This is completely in line with Hasan M. Elahi's† famous quote stating that "*We're all creating an archive of our own lives, whether we're aware of it or not*."

* Taryn Simon (b. 1975) is an American multidisciplinary artist who works in photography, text, sculpture, and performance.

† Hasan M. Elahi (b. 1972) is a Bangladeshi-born American interdisciplinary media artist whose work has an emphasis on technology and media and their social implications.

In general, the utilization of databases for archival purposes involves the organized storage, management, and retrieval of historical records and documents in a digital format. Databases are crucial tools for archivists and institutions looking to preserve, organize, and provide access to their archival collections.

More specifically, the utilization of graph databases for archival purposes offers a compelling alternative to other traditional types of databases, such as relational, object-oriented, and object-relational databases. Graph databases excel in managing and preserving archival collections due to their unique data modeling capabilities that are particularly well-suited to the complex, interconnected nature of historical records and documents, and their ability to represent complex relationships, hierarchical structures, rich metadata, and historical versions efficiently. Besides, their query optimization for graph traversal, scalability, and adaptability to evolving archival needs make them a valuable choice for preserving, managing, and exploring archival collections in the digital age. According to Gartner, it is anticipated that by 2025, graph database technologies will be integrated into 80% of data and analytics innovations, a substantial increase from the 10% adoption rate in 2021. This is expected to expedite decision-making processes throughout enterprises [2].

This research contributes to the overall body of knowledge related to building electronic archives, with a special emphasis on archiving office documents. The paper's main goal is to propose a generic framework suitable for designing a graph database that will facilitate archival document management. The resulting graph data model can be easily modified, enhanced, and adapted to meet any organization's specific requirements for archiving documents.

The rest of the paper is structured as follows. Section 2 offers a concise review of the most notable research conducted in recent years on this subject. In Section 3, we delve into the process of archiving documents, providing a brief but comprehensive overview. Moving forward to Section 4, we explore the world of graph databases, presenting a succinct overview of their essential characteristics and functionalities. This section lays the foundation for the subsequent discussion on our proposed graph data model. Section 5 presents the design of our proposed graph data model in detail, elucidating its structure, components, and how it addresses the specific requirements of document archiving and retrieval. Finally, in Section 6, we draw conclusions based on our research findings and the insights gathered throughout this paper, summarizing the key takeaways and potential implications of our proposed graph data model in the realm of document archiving.

2. Related research

In recent years, the utilization of graph databases for archival purposes has gained increasing prominence. This surge in interest is closely tied to the growing adoption of graph database technology across various domains. The proliferation of graph databases has not only transformed the landscape of data management but has also invigorated research efforts within the archival community. This emerging field is characterized by the convergence of archival science with computational techniques and technologies, indicating a significant shift in how archives are managed and leveraged in the digital age. Some of the most notable research is the work of Laclavík *et al*. from 2011 and 2012, related to the extraction and construction of a graph database (i.e. a social network) from an e-mail archive [3][4]. Based on the case of historical records about real-world individuals, there is research that strives to analyze and compare both graph-oriented databases and document-oriented databases in terms of flexibility, precision of record storage, as well as the ease of ingestion and accuracy of record presentation within the database [5]. In a simple experiment using a graph database, Conrad & Williams (2020) have demonstrated that it is possible to increase the number of access points to individual items in archival collections by leveraging existing machine-readable and searchable data and metadata to identify and display relationships between persons, places, dates, events, etc. across items and collections [6]. Further on, there is research based on the case study of Torre do Tombo National Archives in Portugal aimed at moving from hierarchical, ISAD-conforming descriptions to graph counterparts, which requires utilization of state-of-the-art technologies, data models and vocabularies, such as the design of the core data model for archival records represented as the ArchOnto ontology and its embodiment in the ArchGraph knowledge graph, including a graph database and an Object Graph Mapping library [7] [8].

All the previously mentioned research works corroborate the hypothesis that there exists a profound connection between physical archives and their digital counterparts in the form of graph databases and knowledge graphs. Furthermore, they serve as compelling evidence supporting what Marciano (2022) has termed 'Computational Archival Science (CAS)' [9].

3. Briefly on archiving documents

As per the Macedonian Law on Archive Material, archive material encompasses all original and reproduced/duplicated documentary material (in written, drawn, printed, photographed, filmed, audiovisual, typed, electronic, digital, optical, or any other recorded form) of enduring significance. This material holds importance for the nation, science, academic research, cultural heritage, custodians, and various other necessities [10]. According to this law, a document or record is recorded information generated or received in the course of commencing, conducting, or concluding an undertaking by a specific authority, legal entity, or individual. It must adequately encompass the content, context, and structure of that activity in a manner that, regardless of the document's form or medium, offers proof of that particular activity.

Documentary material, as a source of archival content, encompasses all original and duplicated material (including written, drawn, printed, photographed, filmed, audiovisual, typescript, electronic, digital, optical, or any other recorded form). This category also includes books and other recorded documents generated during the routine activities of the custodians. It remains relevant to their ongoing work until the point at which the archival material is segregated from it.

The document or record is characterized by four fundamental attributes: authenticity, reliability, integrity, and usability.

a) *Authenticity*: A document is considered authentic when it can be substantiated as what it purports to be. It should be verifiable as having been created or dispatched by the claimed author or sender and at the claimed time of creation or dispatch;

b) *Reliability*: A document attains reliability when its content is completely trustworthy, accurately representing the activities or facts it encompasses;

c) *Integrity*: A document is deemed to have integrity when it remains whole and unaltered, free from any tampering or modification; and

d) *Usability*: A document is considered usable when it allows for easy identification, location, searchability, presentation, and interpretation.

Archival and documentary material can be categorized as conventional or unconventional. Conventional archival and documentary material refers to records that are documented on paper and can be read without the need for any special devices or equipment. Non-conventional archival and documentary material encompasses records that are documented on specialized mediums, such as microfilm, optical-magnetic media (e.g., hard disks), compact disks (CDs), DVDs, flash memory, or holographic disks. Specialized devices are necessary to access and read the content within this material. Non-conventional material includes electronic documents generated using computers and other electronic devices in digital formats, data stored in databases that serve as the basis for creating conventional documents through data processing, and digital images obtained by digitization of traditional documents.

Archival work encompasses several essential processes, including the selection of archival material from the documentary material, organizing and arranging both archival and documentary material, documenting (inventory and description) of the archival material, categorizing the archival content, listing the documentary material whose retention period has expired, disposing of the identified documentary material, providing accommodation, storage, and protection for archival and documentary material in proper facilities equipped with standard equipment and technical-technological resources. Finally, it involves the transfer of archival material to permanent storage within the State Archives.

The basis of our work are office and archival operations of public institutions and services (including universities and faculties), state bodies and institutions, public enterprises, and alike. According to the Decree on office and archival operations, in all of these entities, archival operations include a large number of activities related to managing documents/records, such as receiving,

reviewing, and sorting by organizational units and officials, filing the received document in a ledger and auxiliary record books, combining multiple documents/records, delivery, administrative-technical processing that includes signing by an authorized official, putting an official seal and/or stamp, sending, separating and classifying, and archiving [11][12]. The research methodology we have used was based on a thorough and systematic document analysis of relevant legal documents ([10], [11], [12]), which helped to gather information, identify patterns, and draw relevant conclusions.

4. Briefly on graph databases

In graph theory, a graph is a mathematical representation of a set of objects (vertices or nodes) connected by relationships (edges). A graph database represents a specialized type of database management system that is designed to store, manage, and query data using graph structures. As defined by Angles & Gutierrez in 2008, a graph database can be described as a database in which the data structures for both the schema and instances are represented as a (labeled) (directed) graph or generalizations of the graph data structure. In such a database, data manipulation is achieved through graph-oriented operations and type constructors, and it includes integrity constraints suitable for the graph's structure [13]. In a more formal context, a graph database schema takes the form of a graph $G_{db} = (N, E, \psi, V, \lambda)$, where N is a set of nodes and E is a set of edges; ψ is an incidence function $E \rightarrow N \times N$; V is a set of labels and λ is a labeling function $N \cup E \rightarrow V$.

Any graph database leverages all the graph concepts to model, store, and retrieve data in a way that is particularly suited for data with complex and interconnected relationships.

Informally, a graph database can be defined using graph theory concepts in the following manner:

- *Vertices (Nodes)*: In a graph database, each data element is represented as a vertex or node. Nodes can store attributes or properties that define the characteristics of the data element. These properties provide additional information about the data and can be used for querying and analysis;
- *Edges (Relationships)*: Edges represent the connections or relationships between nodes in the graph. These relationships can have different types and can carry additional information in the form of properties. Edges in a graph database can be directed (going from one node to another) or undirected (bidirectional);
- *Labels and Types*: Graph databases often allow nodes and edges to be labeled or typed. Labels or types categorize nodes and edges into different groups, allowing for efficient querying and manipulation based on these categories;
- *Properties (Attributes)*: Both nodes and edges in a graph database can have properties, which are key-value pairs that provide detailed information about the data;
- *Graph traversal*: Graph databases support the traversal operation, which involves navigating through the graph by following relationships between nodes. Traversal allows for efficient querying of complex patterns and paths in the data. Graph traversal algorithms, such as depth-first search (DFS) and breadth-first search (BFS), are commonly used to explore the graph's structure and relationships.

Graph databases have the intrinsic capacity to handle more data, operate at higher speeds, and most significantly, facilitate more connections. They are inherently designed to efficiently store, process, and query relationships, making them highly proficient in managing intricate, interconnected data and complex queries. Importantly, this proficiency remains consistent regardless of the overall dataset size. By replacing the limitations of the conventional relational model with flexibility and speed, these attributes have contributed significantly to their widespread adoption and their recent surge in popularity [14].

5. Proposed graph data model

Designing an effective data model is foundational for graph databases. This section explores the complexities of modeling archival documents as nodes and their relationships to other object types within a graph database. The proposed graph database model is presented in Figure 1.

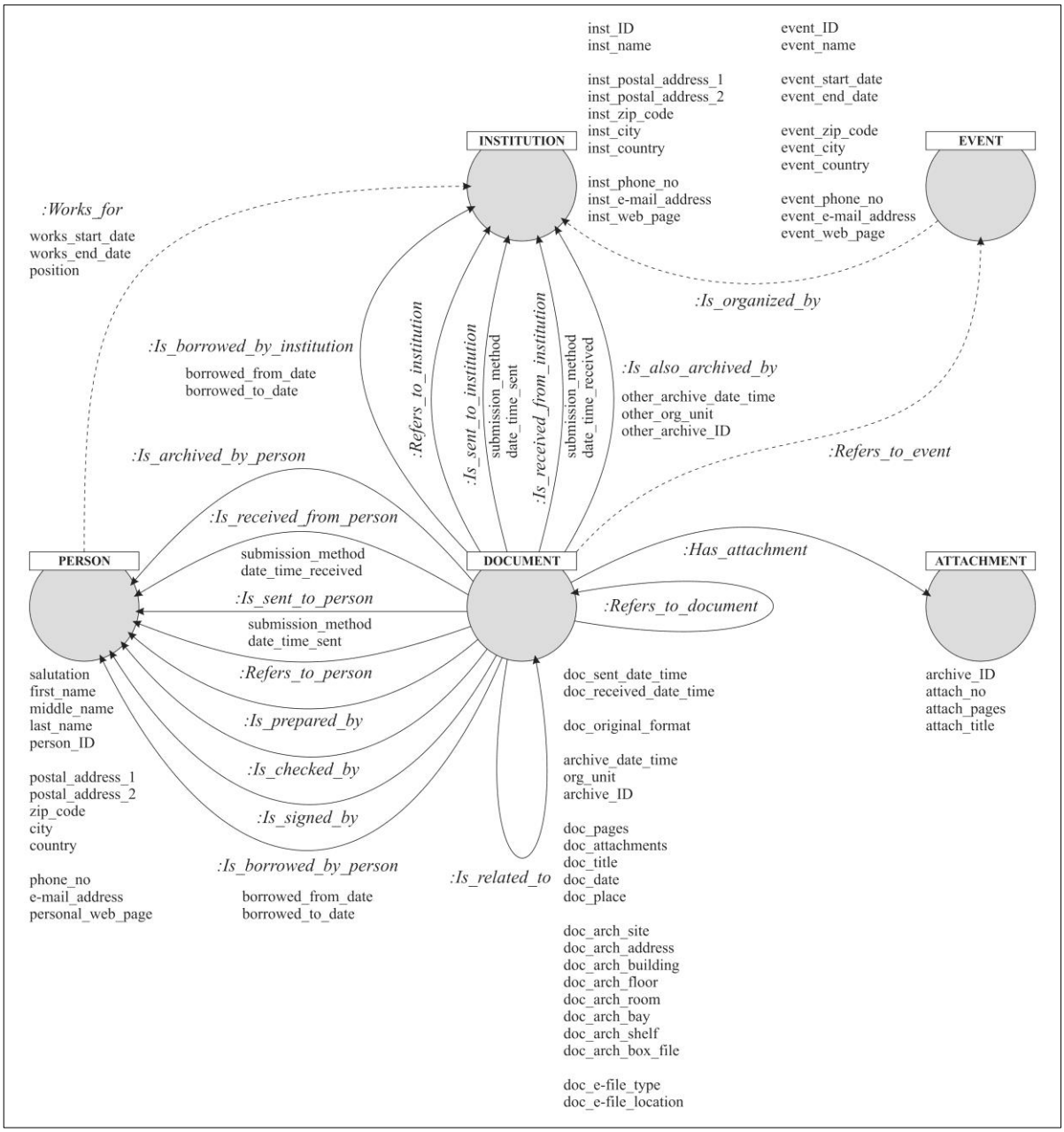


Figure 1: Graph database model (Source: Authors' representation)

In this model, the primary focal point is the 'DOCUMENT' node type. As a representation of an actual archived document, it encompasses various attributes including details like the date and time of reception (*doc_received_date_time*) and archiving (*archive_date_time*), the organizational unit (*org_unit*) associated with it, and a distinctive archiving identifier (*archive_ID*). Additionally, if the document was submitted to the organization via postal mail or electronic mail, the date and time of submission are also recorded (*doc_sent_date_time*). Furthermore, this node encompasses general attributes that provide a comprehensive description of the document, including the number of pages (*doc_pages*), the count of document attachments (*doc_attachments*), the official document title (*doc_title*), as well as the document's date of issuance (*doc_date*) and the place where it was issued (*doc_place*). Finally, the 'DOCUMENT' node type maintains a record of the physical location where the archived document is stored. This physical location adheres to a structured hierarchy, encompassing various levels, including the document archival site (*doc_arch_site*), address (*doc_arch_address*), building (*doc_arch_building*), floor (*doc_arch_floor*), room (*doc_arch_room*), bay (*doc_arch_bay*), shelf (*doc_arch_shelf*), and box_file (*doc_arch_box_file*).

Two reflexive relationships are associated with the 'DOCUMENT' node type: the '*:Referes_to_document*' and the '*:Is_related_to*' relationships. The first relationship establishes a link between a specific document and one or more other documents that are referred to within the text of that document. The second relationship signifies the connection between a particular document and one or more related documents. These reflexive relationships enhance the model's ability to represent and navigate the intricate web of document references and relationships within the archival system.

Closely linked to the 'DOCUMENT' node type is the 'ATTACHMENT' node type. Within the graph data model, each attachment associated with a document is depicted as a distinct 'ATTACHMENT' node, and their connection is illustrated through the presence of the '*:Has_attachment*' relationship. The relationship between these two node types is established through the shared value of the unique archiving identifier attribute (*archive_ID*). In instances where a document includes multiple attachments, each attachment is assigned a sequential number (*attach_no*) for identification. Additionally, for each attachment, the model records the number of pages (*attach_pages*) it comprises, as well as its title (*attach_title*).

In the real-world processes of physically creating and handling documents, various individuals play different roles, all of which are symbolically represented by the 'PERSON' node type within the graph data model. Besides the *person_ID* attribute that uniquely identifies each person, this node encompasses all typically documented attributes that store essential information about individuals involved in these processes, providing a comprehensive profile of each person involved in document-related activities.

Numerous relationships can be established between the 'DOCUMENT' and 'PERSON' node types. For instance, during the document creation process, a document '*:Is_prepared_by*' a person, a document '*:Is_checked_by*' a person, and a document '*:Is_signed_by*' a person. These three relationships help capture the various roles and responsibilities individuals have in the document's lifecycle, offering a detailed view of the document's history and the people involved. In the context of document handling, a specific document can either be received from a particular person (*:Is_received_from_person*) or sent to a particular person (*:Is_sent_to_person*). These two relationships can incorporate additional attributes such as *submission_method*, *date_time_received*, and *date_time_sent*, respectively, which provide additional contextual information about the manner and timing of document reception and transmission, enhancing the comprehensiveness of the archival representation. Certainly, in the context of the graph data model, establishing other relationships between the 'DOCUMENT' and 'PERSON' node types is entirely feasible. For instance, the '*:Refers_to_person*' relationships can signify instances where a document makes references to specific individuals, whilst the '*:Is_borrowed_by_person*' relationships could be used to represent situations where a previously archived document is borrowed by specific individuals. The latter one is supposed to have attributes such as '*borrowed_from_date*' and '*borrowed_to_date*'.

Yet another node type, labeled 'INSTITUTION', is required in the graph data model. Each institution is distinctively identified by the value of the *inst_ID* attribute and it can encompass other pertinent data that are customarily recorded for institutions. This node type enables the model to effectively represent and manage information related to various institutions involved in document-related processes, enriching the overall context of the archival system.

The two basic relationships that can link the 'DOCUMENT' node type to the 'INSTITUTION' node type are '*:Is_sent_to_institution*' and '*:Is_received_from_institution*' that include attributes such as *submission_method*, *date_time_received*, and *date_time_sent*, respectively, which provide additional contextual information about the manner and timing of document reception and transmission. As in the case of the 'PERSON' data type, establishing other relationships between the 'DOCUMENT' and 'INSTITUTION' node types is entirely feasible. For instance, the '*:Refers_to_institution*' relationships can signify instances where a document makes references to specific institutions, whilst the '*:Is_borrowed_by_institution*' relationships could be used to represent situations where a previously archived document is borrowed by specific institutions. The latter one is supposed to have attributes such as '*borrowed_from_date*' and '*borrowed_to_date*'.

The model also accommodates complex scenarios in which a single document is archived by multiple institutions. This is accomplished by introducing a relationship called '*:Is_also_archived_by*' that connects the 'DOCUMENT' and 'INSTITUTION' node types. This relationship encompasses properties derived from the other institution's timestamp, which includes details such as the date and

time of archiving by the other institution (*other_archive_date_time*), the organizational unit of the other institution (*other_org_unit*), and the unique archiving identifier of the other institution (*other_archive_ID*).

In addition to the node types already mentioned, the graph data model can be further enhanced by incorporating additional node types and relationships, which can elevate the graph from a standard representation to a knowledge graph. One such node type is the 'EVENT' node type, which encompasses distinct properties such as the event unique identifier (*event_ID*), event name (*event_name*), and other pertinent data like event start date (event_start_date), event end date (event_end_date), and so on. This 'EVENT' node is interconnected with the 'DOCUMENT' node type through the '*:Refers_to_event*' relationship. This extension enriches the model by associating documents with specific events, fostering a deeper understanding of the contextual relationships between documents and significant events. Another valuable enhancement to the existing graph data model is the introduction of the '*:Works_for*' relationship, establishing connections between the 'PERSON' and 'INSTITUTION' node types. By including this relationship, the model gains the ability to represent the professional associations between individuals and institutions, providing valuable insights into employment history and roles within organizations. This relationship can encompass attributes like the start date of employment (*works_start_date*), the end date of employment (*works_end_date*), and the individual's position (*position*) within the institution.

6. Conclusions

Utilizing databases for archival purposes is a strategic approach to efficiently manage, preserve, and provide access to historical records and documents. A well-designed archival database can facilitate research, ensure the long-term preservation of valuable materials, and contribute to the broader dissemination of historical knowledge. Graph databases are particularly well-suited for archival document management because they include scenarios where relationships and connections between data elements (i.e. archived documents) are crucial.

The proposed property graph model serves as a foundational theoretical framework that can pave the way for the development of more sophisticated data models within the domain of document archiving. As such, it provides a versatile and extensible blueprint upon which future models can be built, offering the potential to address evolving needs in document management, preservation, and retrieval. By embracing this theoretical framework, researchers and practitioners can explore innovative approaches, incorporate additional node types and relationships, and adapt the model to suit specific archival contexts, ultimately advancing the field of document archiving and archival science as a whole.

The implementation of the proposed property graph model in any of the existing graph databases (e.g., Neo4j, TigerGraph, Amazon Neptune, …) can be highly beneficial because all graph databases provide query languages (e.g., Cypher for Neo4j, GSQL for TigerGraph, Gremlin for Amazon Neptune, …) that can allow users to express complex graph patterns and retrieve specific information from the graph. Queries can range from simple node or edge lookups to intricate path and relationship queries. In addition, graph databases employ indexing techniques to speed up queries. These indexes help locate nodes, edges, and relationships quickly, making queries that involve traversing the graph more efficient. Indexing strategies are critical to maintaining good database performance as the dataset grows, which is one of the fundamental intrinsic characteristics of archives.

Some potential limitations of using graph databases for archival purposes are related to the growth in the volume of archived documents and associated data. These include issues such as scalability, complexity of queries, query performance, storage requirements, maintenance, and costs. Additionally, there are challenges related to data migration when transitioning from traditional document management systems to a graph-based approach, as well as concerns about interoperability when integrating the graph model with existing archival systems or external databases.

Future work related to the proposed design can encompass further refinement of the schema to accommodate additional metadata attributes or node types specific to certain archival domains. It can involve: (a) Adapting the model to support different document types, such as images, audio, or video, and considering the unique metadata associated with each; (b) Incorporating semantic technologies

like RDF to enhance the model's ability to represent and query the relationships with greater precision; and (c) Continuous assessment of the graph data model's real-world effectiveness and efficiency, involving thorough validation and benchmarking against archival standards.

References:

[1] The National Archives, Archive Principles and Practice: an introduction to archives for non-archivists, 2016. URL: https://cdn.nationalarchives.gov.uk/documents/archives/archive-principles-and-practice-an-introduction-to-archives-for-non-archivists.pdf

[2] J. Hare, M. Beyer, P. den Hamer, A. Jaffri, R. Sallam, Understanding When Graph Analytics Are Best for Your Business Use Case, Gartner Research, 2022. URL: https://www.gartner.com/en/documents/4013850.

[3] M. Laclavík, Š. Dlugolinský, M. Kvassay, L. Hluchý, Email Social Network Extraction and Search, NextMail 2011 Workshop, WI-IAT 2011, Lyon, France, in: Proceedings of the 2011 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, IEEE Computer Society, 2011, pp. 373–376. doi: 10.1109/WI-IAT.2011.30.

[4] M. Laclavík, Š. Dlugolinský, M. Šeleng, M. Ciglan, L. Hluchý, Emails as Graph: Relation Discovery in Email Archive, WWW 2012 EMAIL '12 Workshop, Lyon, France, in: Proceedings of the 21st International Conference on World Wide Web, Association for Computing Machinery pp. 841–846. doi: 10.1145/2187980.2188210.

[5] M. Healy, T. O'Connor, J. Keating, Comparison of Graph- and Collection-Based Representations of Early Modern Biographical Archives, in: T. Andrews, F. Diehr, T. Efer, A. Kuczera, J. van Zundert (Eds.), Graph Technologies in the Humanities 2020, CEUR Workshop Proceedings 2020, Vol-3110, Vienna, Austria, 2020, pp. 60–85.

[6] M. Conrad, L. Williams, Elevating "Everyday" Voices and People in Archives through the Application of Graph Database Technology, in: Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 2020, pp. 1861−1865. doi: 10.1109/BigData50022.2020.9377921.

[7] I. Koch, N. Freitas, C. Ribeiro, C. T. Lopes, J. R. da Silva, Knowledge Graph Implementation of Archival Descriptions through CIDOC-CRM, in: A. Doucet, A. Isaac, K. Golub, T. Aalberg, A. Jatowt, (Eds.), Digital Libraries for Open Knowledge, in: Proceedings of the 23rd International Conference on Theory and Practice of Digital Libraries, TPDL 2019, Oslo, Norway, volume 11799 of Lecture Notes in Computer Science, Springer Nature Switzerland AG, Cham, Switzerland, 2019, pp. 99–106. doi: 10.1007/978-3-030-30760-8_8.

[8] L. Costa, N. Freitas, J. R. da Silva, An Evaluation of Graph Databases and Object-Graph Mappers in CIDOC CRM-Compliant Digital Archives, Journal on Computing and Cultural Heritage, 15.3 (2022): 1−18. doi: 10.1145/3485847.

[9] R. Marciano, AFTERWORD: Towards a new Discipline of Computational Archival Science (CAS), in: L. Jaillant (Ed.), Archives, Access and Artificial Intelligence: Working with Born-Digital and Digitized Archival Collections, Bielefeld University Press, Bielefeld, Germany, 2022, pp. 205–218. doi: 10.14361/9783839455845.

[10] State Archives of the Republic of Macedonia, New Draft-Law on Archive Material, 2012. URL: https://arhiv.mk/wp-content/uploads/2021/01/predlogzakon_za_arhivski_materijal.pdf

[11] Government of the Republic of Macedonia, Decree on Office and Archival Operations [in Macedonian], Official Gazette of Republic of Macedonia, No. 1 from January 3rd, 2014, pp. 4–8. URL: https://www.slvesnik.com.mk/Issues/fc82ee320a774f22a861ae2a5392965e.pdf

[12] State Archives of the Republic of Macedonia, Instructions on Methods and Techniques for Managing Archive and Documentary Materials in Office and Archive Work [in Macedonian], Official Gazette of Republic of Macedonia, No. 99 from July 2nd, 2014, pp. 23−51. URL: https://www.slvesnik.com.mk/Issues/86bef7078a6b4be483a8f93b79f75402.pdf

[13] R. Angles, C. Gutierrez, Survey of Graph Database Models, ACM Computing Surveys, 40.1 (2008): 1−39. doi: 10.1145/1322432.1322433.

[14] I. Robinson, J. Webber, E. Eifrem, Graph Databases: New Opportunities for Connected Data, 2nd. ed., O'Reilly, Sebastopol, CA, 2015.

Exploring the Impact of AI-Driven Marketing Strategies on Player Retention in the Video Game Industry

Stefan Ugrinov[1], Dragan Ćoćkalo[1], Mihalj Bakator[1], Mila Kavalić[2], Verica Gluvakov[1]

[1] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Zrenjanin, Serbia*
[2] *Faculty of Technical Sciences, Novi Sad, Serbia*

stefan.ugrinov@tfzr.rs, dragan.cockalo@tfzr.rs, mihalj.bakator@tfzr.rs, mila.kavalic@tfzr.rs, verica.gluvakov@tfzr.rs

Abstract:
The video game industry is experiencing a transformative shift with the integration of Artificial Intelligence (AI) into marketing strategies and player retention efforts. This paper presents a comprehensive literature review that delves into the impact of AI on video game marketing and player retention. As AI technologies continue to advance, game developers and marketers are leveraging AI-driven tools and techniques to enhance player engagement, tailor in-game experiences, and optimize marketing campaigns. This review synthesizes existing research, highlighting key findings and emerging trends in the field. This review investigates the implications of AI on player retention, exploring how AI-driven recommendations, community management, and customer support impact player loyalty. This paper provides valuable insights into the transformative potential of AI in the video game industry, offering a foundation for future research and practical applications in game development and marketing.

Keywords:
Artificial Intelligence, Player Retention,Marketing, Personalization,Strategy

1. Introduction

Digitalization of any field is driven by Industry 4.0 technologies [1]. AI serves as a cornerstone of Industry 4.0. The technology is increasingly being leveraged in video game marketing and player retention strategies, reshaping the way developers engage with their audiences. Chief marketing officers are increasingly embracing the technology in order to improve their existing marketing strategies. Among all the functions within a company, marketing stands out as the one with the greatest potential for benefiting from artificial intelligence. The fundamental tasks of marketing include comprehending customer requirements, aligning them with products and services, and convincing individuals to make purchases—functions that can be significantly improved through the application of AI. AI's involvement in video game marketing is diverse, with one of its foremost applications being predictive analytics. Through the examination of player data, AI enables marketers to gain insights into player behavior, preferences, and trends. This valuable information is subsequently employed to customize marketing approaches, guaranteeing that promotional materials are pertinent and enticing to individual players. For instance, in the case of a player who regularly participates in racing games, AI can assist marketers in directing advertisements towards them featuring new releases in the same genre. Apart from marketing, AI plays a pivotal role in player retention, a decisive element in the triumph of any video game. One method by which AI contributes to enhancing player retention is by providing personalized gaming encounters. AI has the capacity to scrutinize a player's conduct and inclinations, thereby customizing the gaming experience to match their preferences. This may encompass modifying the game's complexity, proposing pertinent in-game acquisitions, or even adapting the game's narrative in accordance with the player's actions. Through the establishment of a more tailored experience, developers can elevate player engagement and foster loyalty.

AI is also being used to improve player support services. With the help of AI-powered chatbots, players can receive instant assistance for their queries or issues. These chatbots can handle a wide range

of inquiries, from troubleshooting technical issues to providing gameplay tips. By offering immediate and effective support, developers can enhance player satisfaction and retention [2].

Companies also utilize AI throughout the entire customer journey. When potential customers are in the "consideration" phase and conducting research on a product, AI is used to direct targeted advertisements towards them and can assist in steering their search process.

2. Material and methods

In this part we will talk about materials and methods used in this paper, what is the subject of the research, what is the goal of this research, what are the questions we asked during the research and based on the results we will give our opinions and this subject.

2.1. The subject and the problem of research

This paper will deal with the analysis of exploring the impact of ai-driven marketing strategies on player retention in the video game industry by considering different literature sources and studies. Also, this paper will try to indicate the importance of industry 4.0 and focus mostly on AI but on other parts of it too.

2.2. Research goal

Primary goal of this paper is to showcase importance of AI-driven marketing strategies that are being used in order to achieve player retention in video game industry and also try to raise current level of player retention rate through new and improved AI-driven marketing strategies.

2.3. Research question

Based on analyzed theories, we will try to answer the following question: How do AI-driven marketing strategies increase player retention rate and help attract new players.

2.4. Research method

This is a form of theoretical research in which universal conclusions are made by considering previous conducted research, studies and literature. Research will consider results of as many other authors as possible in order to make conclusions.

3. Results and discussion

Constructing the right business case is often a critical determinant of a company's success in implementing AI. These solutions ultimately aim to streamline decision-making processes and improve communication as a result of information analysis [3]. The fundamental element that enables AI mechanisms to achieve the assumptions made is data. This is because they make it possible to understand market phenomena, draw conclusions from them and make decisions [4]. From this we can conclude that implementation of artificial intelligence in a company requires the precise establishment of business objectives, as well as access to data and appropriate tools, together with techniques for their analysis. While these systems do come with a significant cost, their proper utilization offers several

advantages to the organization, ultimately fostering growth. The primary objective here is to enhance revenue, profit, or market share, rather than simply cutting expenses.

Following can be identified as the main factors encouraging investment in AI-driven marketing strategies :

- Enhanced Personalization: AI can analyze vast amounts of customer data to create highly personalized marketing campaigns. Tailoring content and recommendations to individual preferences increases engagement and conversion rates.
- Improved Customer Insights: AI can provide in-depth customer insights by analyzing behavior, preferences, and demographics. This information helps in understanding target audiences better and making data-driven decisions.
- Efficient Customer Support: AI-powered chatbots and virtual assistants can provide 24/7 customer support, answering inquiries, and resolving issues promptly, enhancing customer satisfaction and saving operational costs.
- Predictive Analytics: AI can forecast trends and customer behavior, enabling businesses to plan marketing strategies effectively, optimize resource allocation, and stay ahead of the competition.
- Marketing Automation: AI automates repetitive marketing tasks such as email marketing, social media posting, and ad targeting. This saves time, reduces errors, and ensures campaigns run smoothly.
- Cost-Effective Advertising: AI can optimize advertising spend by targeting the right audience at the right time, reducing wasted ad spend and improving ROI.
- Real-time Data Analysis: AI can process and analyze data in real time, allowing marketers to respond quickly to changing market conditions and customer needs.
- Competitive Advantage: Businesses that adopt AI-driven marketing strategies can gain a competitive edge by delivering more relevant and timely content to their audiences.
- Scalability: AI systems can handle large volumes of data and tasks, making it easier to scale marketing efforts as a business grows.
- Continuous Improvement: AI can continuously learn and adapt, optimizing marketing strategies over time for better results.
- Customer Retention: Personalized experiences and timely communication through AI-driven strategies can improve customer loyalty and retention rates.
- ROI Tracking: AI can provide detailed insights into the return on investment (ROI) of marketing campaigns, helping businesses allocate resources effectively.
- Adaptation to Trends: AI can quickly adapt to emerging marketing trends and consumer behaviors, ensuring that businesses stay relevant in a rapidly changing landscape.
- Data Security: AI can enhance data security by identifying and mitigating potential threats and vulnerabilities in marketing systems.
- Global Reach: AI can help businesses expand their reach to global markets by optimizing content for different languages and cultures.

Marketing AI can be categorized according to two dimensions: intelligence level and whether it's stand-alone or part of a broader platform. Some technologies, such as chatbots or recommendation engines, can fall into any of the categories; it's how they're implemented within a specific application that determines their classification [5] . If we talk about intelligence level we have task automation and machine learning. Task automation is focused on performing repetitive, structured tasks that require relatively low levels of intelligence. These systems are crafted to adhere to a predefined set of rules or perform a specific series of operations based on provided inputs. However, they lack the capability to tackle intricate issues, such as addressing nuanced customer requests. This paper is about retaining players in video games, so great example here would be sending player welcome email upon purchasing game to try to make extra positive impression on consumer. Consumer also require customer support and chatbots are good example of using AI to improve relationship with customers. Opposite of task automation we have machine learning. Where task automation is based on performing easier tasks machine learning is mostly focused and trained on using large quantities of data to make relatively complex predictions and decisions.

These models possess the ability to identify images, decode text, categorize customers, and predict how customers will react to different initiatives, such as promotional campaigns. With AI/ML algorithms, game developers can optimize game ecosystems, fine-tune mechanics and address both bugs and glitches. Through the analysis of player data, behavior and game statistics, these algorithms provide valuable insights that allow developers to improve gameplay mechanics where needed [6]. Both AI and ML are excellent options to maximize revenue. One for is personalized in-game advertising. Game developers can analyze player data and deliver personalized advertisements tailored to individual players. If we put this in perspective of a video game great example would be Ubisoft's and Idomoo's campaign for "Assassin's Creed", created over 18 million of personalized videos, exemplifying this approach. The AI algorithms analyzed player preferences and behaviors to generate personalized ad experiences, resulting in increased engagement and higher conversion rates. Game developers started using microtransactions in past 10 years, mostly in games that use free to play model. AI comes clutch in these situations too. AI analyzes data to predict player behavior and preferences, allowing developers to create targeted in-game marketing campaigns and offers. In following article [7] we can see that just five mobile games drew more than 7.2 billion dollars revenue. Diving straight in at the top, Honor of Kings is identified as the year's top performer, earning more than $2.2 billion throughout the year. This was followed by PUBG Mobile ($1.7 billion), Genshin Impact ($1.56 billion) and Candy Crush Saga ($1.02 billion). Roblox, which drew $861.8 million, is the only title within the top five not to cross the billion dollar threshold.

We also have to consider and understand stand-alone and integrated AI. These AI systems operate independently from the primary channels that customers typically use to discover, purchase, or receive support for a company's offerings. Likewise, they are distinct from the channels that employees utilize for marketing, sales, or servicing these offerings. In essence, customers or employees must make a deliberate effort outside of these usual channels to engage with the AI systems. On other hand we have integrated applications. Integrated into current systems, these AI applications often operate with lower visibility compared to standalone solutions for customers, marketers, and salespeople who interact with them. For instance, machine learning algorithms that make instant decisions regarding the display of digital advertisements are integrated into platforms that manage the end-to-end process of ad procurement and placement.

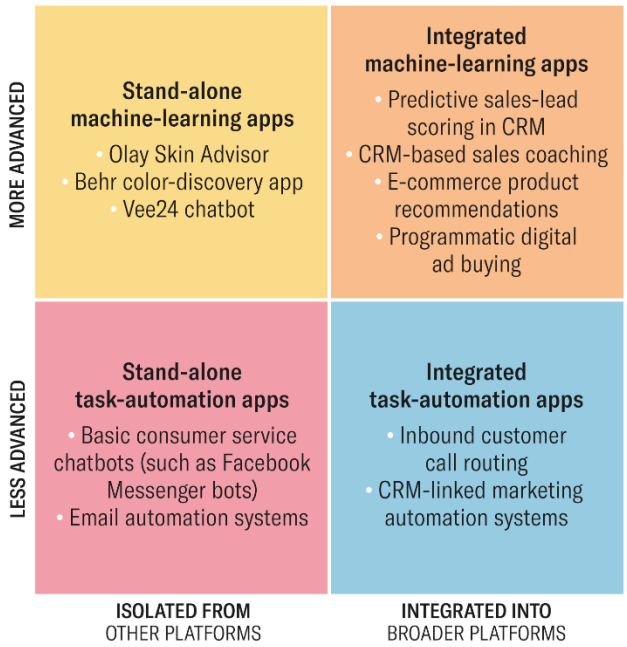


Figure 1:categorization of AI-driven marketing strategies

Market strategy 4.0 majorly emphasizes personalization of services and products on the basis of big data analysis. In this way, the marketing approach integrates the offline and online interaction between consumers and companies. The integration of artificial intelligence and machine learning strengthens the process of customer interaction and also enhances the productivity of other technologies [8]. Several technologies influence marketing strategies, namely IoT, cloud computing, big data, artificial intelligence/machine learning, blockchain, digital twin, robots/drones as well as the metaverse [9]. AI/ML has paved the way for ample opportunities in marketing solutions. A large chunk of research in AI is based on explaining actions or decisions on how humans behave in a particular situation [10]. The development in information and communication technology (ICT) has changed the B2B perspective in general, where a huge amount of data can be collected, stored, and analyzed using technology such as AI/ML [11]. Irrespective of the fact that marketing has seen huge advancements due to AI/ML, there are still challenges that businesses need to face, due to a lack of expertise for strategic implementation [12].

4. Guidelines and recommendations

Based on literature review done it this paper we can conclude ai-driven marketing strategies have huge impact on business and also on player retention in video games which is the main subject of this paper. Following results of research we can give following recommendations for implementation of ai based marketing strategies with aim of retaining players in video games:

- In order to to achieve desired goal, in this case player retention in video games, business should implement AI-driven marketing strategies to achieve their goal.
- AI-driven marketing strategies help us process large amount of data in short period of time and help us better understand needs and wants of customers.
- Other than processing large amount of data in short period of time AI-driven marketing strategies help us with scalability, global and more precise reach, automation of certain marketing processes and efficient customer support.
- Adaptation to trends,data security and improved customer insight are some of the benefits of implementing AI-driven marketing strategies.

5. Conclusion

In the dynamic landscape of the video game industry, the symbiotic relationship between AI-driven marketing strategies and player retention has been illuminated through this exploration. Our examination of existing literature underscores the transformative potential of AI in the quest to not merely attract players but to foster enduring engagement and loyalty.AI's capacity to personalize gaming experiences, predict player behavior, and optimize marketing campaigns presents a paradigm shift in the industry. It empowers game developers and marketers to create tailored, immersive experiences that resonate with individual players.As we stand at the intersection of technology and entertainment, it is evident that AI-driven marketing strategies have become indispensable tools for businesses seeking sustained growth and competitiveness. The ability to decipher player preferences, anticipate trends, and deliver personalized content positions AI at the forefront of the industry's evolution.In conclusion, this paper underscores the imperative for industry stakeholders to embrace AI as a catalyst for innovation and enhanced player experiences. By harnessing the potential of AI-driven marketing strategies, the video game industry can not only thrive but also continue to captivate and delight players in an ever-expanding digital realm

References:

[1] Bradley, K. Defining Digital Sustainability. Libr. Trends 2007, 56, 148–163

[2] Marcin Frąckiewicz 3 August 2023 The Role of AI in Video Game Marketing and Player Retention , ts 2, https://ts2.space/en/the-role-of-ai-in-video-game-marketing-and-player-retention/
[3] Rutkowski, I. (2020). Inteligentne technologie w marketingu i sprzedaży — zastosowania, obszary i kierunki badań. Marketing i Rynek 27(6), pp. 3-12. doi: 10.33226/1231-7853.2020.6.1.
[4] Warszycki, M. (2019). Wykorzystanie sztucznej inteligencji do predykcji emocji konsumentów. Studia i Prace Kolegium Zarządzania i Finansów. Szkoła Główna Handlowa, 173, pp. 111-121.
[5] Davenport, T. H., Guha, A., & Grewal, D. (2021, January 1). How to Design an AI Marketing Strategy What the technology can do today—and what's next. Harvard Business Review, 99(4), 42–47.
[6] Steeve Huin | Jul 28, 2023 | How Artificial Intelligence (AI) is redefining video games https://blog.irdeto.com/video-gaming/how-artificial-intelligence-ai-is-redefining-video-games/
[7] Lewis Rees, January 18th, 2023, Which five mobile hits generated $7.2 billion in revenue in 2022? https://www.pocketgamer.biz/news/80646/2022s-top-five-mobile-games-generated-72-billion-in-revenue-in-2022/
[8] Guven, H. Industry 4.0 and Marketing 4.0. In Perspective of Digitalization and E-Commerce; Akkaya, B., Ed.; Agile Business Leadership Methods for Industry 4.0; Emerald Publishing Limited: Bingley, UK, 2020; pp. 25–46
[9] Wang, L.; von Laszewski, G.; Younge, A.; He, X.; Kunze, M.; Tao, J.; Fu, C. Cloud Computing: A Perspective Study. New Gener. Comput. 2010, 28, 137–146
[10] Aline, F.S.; Borges, F.J. The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. Int. J. Inf. Manag. 2020, 57, 102225.
[11] Saura, J.R.; Ribeiro-Soriano, D.; Palacios-Marqués, D. Setting B2B digital marketing in artificial intelligence-based CRMs: A review and directions for future research. Ind. Mark. Manag. 2021, 98, 161–178
[12] De Bruyn, A.; Viswanathan, V.; Beh, Y.S.; Brock, J.K.-U.; Von Wangenheim, F. Artificial Intelligence and Marketing: Pitfalls and Opportunities. J. Interact. Mark. 2020, 51, 91–105.

Overview of E-invoice in Serbia

Milica Mazalica, Biljana Radulovic and Aleksandra Stojkov Loncarski

University of Novi Sad, Technical Faculty "Mihajlo Pupin", Djure Djakovica bb, Zrenjanin, 23000, Serbia

milica.mazalica@tfzr.rs; biljana.radulovic@tfzr.rs; aleksandra.stojkov@tfzr.rs

Abstract:
E-invoicing is the electronic transfer of billing and payment information, via the Internet or other electronic means, between parties involved in a commercial transaction. The development of modern communications has contributed to the development of E-invoice. Thanks to the constant development of information and communication technologies, business and the organization of e-invoices are being improved, both in Serbia and in the region. In the Republic of Serbia, E-invoice represents an integral part of the SEF invoice management system. SEF is a program or information technology solution managed by a central information intermediary. The format supported by SEF for the exchange of invoices is the XML e-invoice format. The portal of E-invoice in Serbia is https://www.efaktura.gov.rs/. [9] The users through this portal can send their XML format of e-invoice or have insight into the status of the sent invoice. This research presents a short analyzes of E-invoice system in Republic Serbia. The Law on Electronic Invoicing entered into force on May 7, 2021. invoicing was introduced in Serbia. The aim of this paper is on the analyzes development of user's interface and this system in a in a developing country.

Keywords:
E-invoice, E-Government, XML, E-business, invoice

1. Introduction

Information communication technology (ITC) offers huge possibilities for organization and society. This system enables innovation in business models, optimization in business processes, facilitates business networking, and presents unprecedented possibilities for society. ITC is one of the most important factors for organization competitive and overall economic growth. Electronic invoicing presents complementary possibilities to use the ICT for both the automating and redesigning business processes. [6]

Electronic invoicing presents an electronic form of billing. E-invoicing includes several different technologies, and it is usually used as a number term to describe any method by which a document is presented electronically from one part to another. Electronic invoices are invoices that are issued, received, processed, recorded, and stored exclusively electronically in electronic form (they remain in digital format until the end of the cycle) through the new system of electronic invoices (SEF) managed by the Ministry of Finance. An electronic invoice is a request for payment based on a transaction with a fee. Any other document that affects the payment, i.e. the amount of the payment. An invoice issued for turnover without compensation, as well as received advances. SEF will be required to be used by public companies, legal or natural persons, as well as entrepreneurs who are in the VAT system. If the person is not a subject of the public sector or is not in the VAT system, he is not obliged to issue invoices in this system, but he can use it (in certain special situations, it is a necessity). The advantages of e-invoices are that their creation is faster and simpler because there is no printing, packaging and mailing. Searching and saving them is much simpler compared to paper versions. Efficiency increases, work errors decrease. Systemically, better control is implemented, and the degree of success of e-invoice delivery is very high.

SEF is a program or information technology solution managed by a central information intermediary, i.e. a competent unit within the Ministry of Finance, which maintains the register of information intermediaries, manages the e-invoice system and is responsible for its functioning. The law provides for registration procedures for each user to apply for SEF, free of charge. The method of

exchanging e-invoices (XML format), supporting documentation in PDF, as well as exchanging text information and notification messages will be automatic. From the moment one of the participating parties initiates the sending of the document to the other party, the System assumes responsibility for safe delivery with additional archiving. The delivery procedure also includes the transfer of information on confirmation of receipt, rejection with the indication of reasons for non-acceptance. [1]

The Republic of Serbia dated 1 May 2022 introduces the obligation to take the public sector entity to receive and preserve electronic invoices and to issue electronic invoices to another public sector entity. Structural reform, which belongs to the Law on Electronic Invoicing, implies two measures - introducing a new model of focalization and transitioning to electronic invoicing. In this way, business with the state becomes more economical and safer, and the law prescribed the manner of sending, receiving, and storing electronic invoices will gradually adopt economic entities in mutual transactions. [1]

2. Related work

The invoice presents a business document which is issued by the seller to the customer. This document lists the names, quantities and prices agreed for products or services provided by the seller to the customer. The invoice obliges the customer that the amount that is in the invoice is indicated to the seller, by the terms of the payment. The maximum number of days pay was determined to the customer and sometimes the discount was offered if paid before the deadline. Invoice and proforma invoices can be in paper and electronic form. The invoice does not have to contain a signature, nor is the use of a stamp obligatory (Law on Companies), but it is necessary to contain an identification mark (Law on Accounting), which confirms its authenticity. [1] When classic sending of invoices is used, we see that we have a greater number of steps in the entire procedure, and costs are increased due to paper, envelope and mailing, all of which are paid for. Using the electronic invoice, the automation and digitization process will be significantly accelerated and the overall costs of the business will be directly reduced.

With e-invoices, the number of steps in the process is twice as small, the costs and time required for implementation are drastically reduced, and the biggest advantage is that we have insight into the desired invoice at any time. The electronic invoice system processes an electronic invoice, which is: final invoice, advance invoice, a document on the increase in compensation, or a document on the reduction of compensation. The electronic invoice must contain:

- name, address, and tax identification number of the issuer if any of the issuers is a legal entity,
- unique number of users of public funds
- business account of the issuer;
- name, address, and tax identification number of the recipient,
- JBKJS of the recipient if the recipient is a user of public funds t
- serial number and date of the e-invoice
- date of advance payment, if it is an invoice for advance payment,
- the code and/or name of the good or service for each electronic item invoices and the quantity and unit of measure for the delivered goods
- value for each item from the electronic invoice
- the total amount of the electronic invoice and
- the number of advance payments.

In paper [1] is given analyze of the application of e-invoices on a sample and private sector entities. The aim of this paper is to point out the advantages and disadvantages of e-invoices in the first phase of their implementation. routing within offices and automatic synchronization of data upon receipt. The research in [1] indicates the advantages and disadvantages of applying e-invoices in real working conditions. The goal of this paper is to point out the problems that e-invoice users have recorded in the first months of work in the application.

The paper [2] examines the effect of using eINV on business process performance. The paper therefore analyses whether eINV is streamlining administrative procedures, enhancing security and cutting costs. Also this paper analyses whether business processes were merely automated or also redesigned.

The paper [3] analyze firm performance when using e-invoicing on a regular basis. The research took place in Spain. For propose this research causal relationships proposed in the model are tested through structural equation modeling. The results of this research show that the habit of using e-invoicing improves firms' trust in this IT and their perceptions of its usefulness and security.

Paper [4] presents an open electronic invoicing system named eInvoke, based on XML, XML cryptography and Web Services, that addresses all security requirements imposed by the Directive all this in order to offer an interoperable, affordable, state-of-the-art standards compliant and scalable solution, addressing the security requirements imposed by EU legislation

3. E-invoice in Serbia

On the basis of the law on the use of electronic invoices, it is planned to enable the provision of technical solutions in Republic of Serbia. The first one through the user interface and when an authorized person can create and issuance of electronic invoices and/or acceptance and rejection of electronic invoices and electronic recording calculation of value added tax on behalf of the subject of the transaction within its powers. And the second through an application interface (API - Application Programming Interface) that enables connection with the traffic subject's software system. [8] In specification of Serbian e-invoice there are three groups of electronic invoice data:

- basic content of the invoice,
- additional invoice data - data that is added to the invoice when it is issued in the electronic invoice system,
- status data - data that is added or can be changed after the invoice is issued, such as
- information on invoice acceptance or rejection.

When an invoice with associated additional data is downloaded from the electronic invoice system, it is used XML format that adds a wrapper to the UBL invoice format. The exact specification of the cover format is a highly technical part specifications of the application interface. The structure of xml format is given in Figure 1.

```
<env:DocumentEnvelope xmlns:env="…">
    <env:DocumentHeader>
        <env:DocumentId>6946581a-99b0-470c-ad60-d60d8fdaf9c5</env:DocumentId>
        …
    </env:DocumentHeader>
    <env:DocumentBody>
        <Invoice …>
            …
        </Invoice>
    </env:DocumentBody>
</env:DocumentEnvelope>
```

Figure 1. XML structure of E-invoice [5]

In e-invoice system is provided that in the application for registration it is possible to create data for authentication through the application interface. The authentication data in this system contains the identifier of the client side. The application interface is of the REST API type. This means, among other things, that each operation is determined by the method of the HTTP request (method) and the

path (path) that is part of the URL. Each application interface operation call can be completed in three ways: successful, failed miserably, messy. Users who want to connect their ERP with the system of electronic invoices need to:

- Download the API documentation from the portal https://www.efaktura.gov.rs/
- Open a user account on the e-invoice system
- In the settings section (e-invoice system) - API management - the user has the option to generate an API identification key
- The user installs independently or delivers to his service provider, i.e. the supplier of the information system. An API identification key that is embedded in API methods. [5]

4. User interface of the e-invoice system

In order for the user to use the services of the e-Invoice system, it is necessary to first register on the Portal for electronic identification (https://eid.gov.rs/). Registration is possible in two ways:

- a qualified electronic certificate or
- username and password with mandatory activation of the ConsentID mobile application.

For the purposes of using the E-invoice system, the user can choose one of two registration methods. The user registers with a qualified electronic certificate and how with a username and password with mandatory activation of the ConsentID mobile application. After authentication, the user accesses the account for which he is authorized and selects the desired activity. The users account is present in Figure 2. [7]

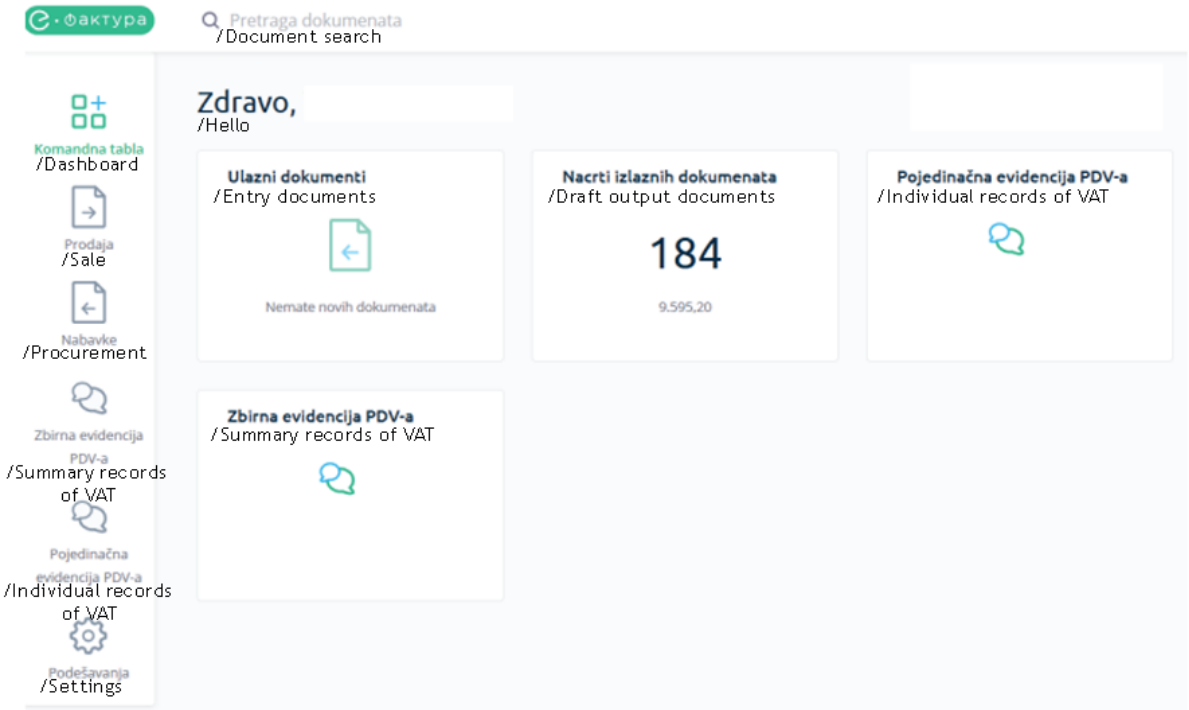


Figure 2. Control panel for users [5]

The central part of the user's screen is given in Figure 3., and it shows the icons which represent shortcut for the functionalities that they use most often.

The basic prerequisite for sending and receiving invoices through the Electronic Invoice System is that the issuer and the recipient of the invoice have an open account on the Electronic Invoice System. Users, when opening an account within the company on the Electronic Invoice System can assign the following roles: administrator, user for outgoing documents, user for incoming documents, user (for input and output documents) and auditor. By filling in the name of the company, the drop-down menu will show all the companies whose name begins with it with the letters we previously entered. Users can also search for companies by entering Parent number, PIB or JBKJS. It is necessary to select the name of the company to

which we want to send the XML file. The data that changes in the XML file is the same for all four types documents:

- invoice;
- advance invoice;
- document on the increase;
- reduction document.

The invoice issuer needs to open the XML file in "Notepad++" and modify the necessary data. In Figure 4. is given the example of electronic invoice which is made using the XML template for e-invoice. [8]

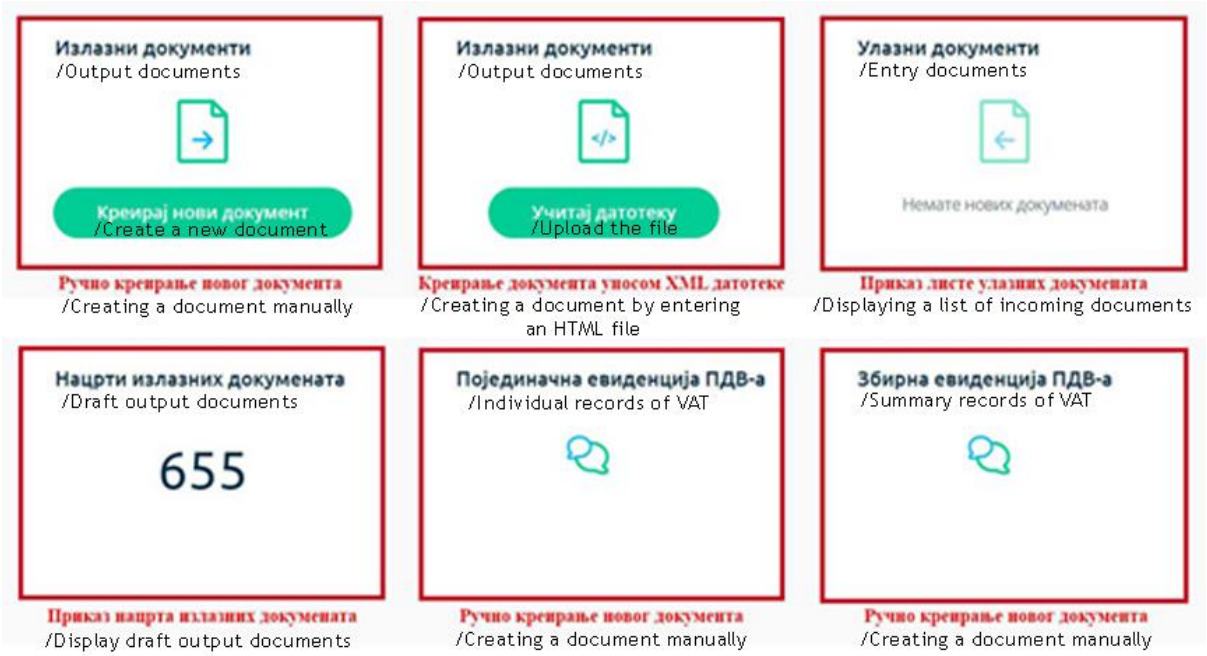


Figure 3. control panel with shortcut [5]

ELECTRONICS SERVICE — **Location** (001)

Pupinova bb, 23000 Zrenjanin
Bank account: 273-50007778000-90
TAX ID: 100885965 Identification number: 33365991
Phone Number: 023/9235-356; 069/662-55-22
servis@gmail.com
www.serviszr.co.rs

Account
Document number: 003351/011
Document date: 19.09.2023.
Place of issue: 23000 Zrenjanin
Date of turnover of goods and services: 19.09.2023.

Name: **WHOLESALE**
City: 23000 Zrenjanin
Street: Cara Dušana 3/A
TAX ID: 108317855 ID: 62521222
Customer ID: 003351

Serial number	Code	Type of goods and services	Unit of measure	Quantity	Price without VAT	Discount (%)	Price with discount	VAT rate	Price with VAT	Sale value
1.	18	THERMO ROLL 57 x 40	piece	150.000					40.000	6.000.00
TOTAL:										6.000.00

Total for payment: 6.000.00
In letters: Six thousand dinars and 00/100

Payment: Cash Currency date: 19.09.2023.
Please enter the reference number when making the payment: 97 11-003351-011-09-23

Invoice issued *stamp place* *Invoice received*

Figure 4. Example of E-invoice [5]

5. E- invoice in the region

E-invoicing in Europe has over 400 services that reduce costs by up to 70% compared to paper invoicing. Many countries have stipulated by law that all companies must switch to e-invoices. [16]

Table 1 shows the States and years when this system was gradually introduced, the name of the e-invoice management system and the used messaging format. From the list of the mentioned countries, it can be noted that Finland was the first to start using this way of doing business (2003), while Serbia and Romania recently started (2021). The systems listed may have different names and user interfaces, but they all have the same function and they are all based on the XML messaging format, which is a universal way of transferring data.

State	Year	Management system	Exchanging format
Slovenia[2]	2015	PPAnet	XML
Finland[2]	2003	Finvoice	
Denmark[2]	2005	NemHandel	
Norway[2]	2009	PEPPOL	
Italy[12]	2019	SDI	
Spain[10]	2015	FACe	
Romania[11]	2021	RO e-Factura/ EDICOM	XML/ SAF-T
Greece[13]	2014	/	XML/ EDI
Croatia[14]	2007	PEPPOL	XML
Serbia[15]	2021	SEF	XML

Table 1. States of the region that use e-invoices

Conclusions

A structured e-invoice is an electronic document consisting of the necessary contents (invoicing data) compiled into a harmonized electronic message which it has a familiar structure and format. The Office for Information Technologies and Electronic Government has provided an Invoice Management System (SEF) through which e-invoices are issued. SEF is a program or information technology solution managed by a central information intermediary, i.e. a competent unit within the Ministry of Finance, which maintains the register of information intermediaries, manages the e-invoice system and is responsible for its functioning.

The User Interface is a Signed Java Applet running on a standard web browser. The user interacts with the system through this interface to create, manage and send e-invoices. The interface is able to produce XAdES signatures according to the hosting organization signature policy. The Web Service of E-invoice is a Java which is based on implementation of a Web Service, running as a servlet on an Apache Tomcat web server. The Database used is the eXist native XML database. It is running as a standalone server and communicates with the User Interface and the E-invoice Web Service. The e-invoice system use XML digital signatures as defined in the ETSI XAdES standard to sign the issued XML E-invoices. An open source implementation of XadES was used. [5] [4]

E-invoice bring totally new approach to business with several advantages and expected disadvantages the initial phase of their implementation. The E-invoices are indisputably progressed in business that will enable all entities to raise their business to a higher level, with the possibility of continuous control and up-to-dateness at all levels of business.

References:

[1] M. Mudric, N. Stefanovic, M. Milanovic, D. Knezevic: E-invoicing – Case Study in Serbia, 9th International scientific conference Technics and Informatics in Education, September 2022, Kragujevac, DOI: 10.46793/TIE22.326M

[2] A. Groznik, A. Manfred: E-INVOICING AND E-GOVERNMENT – IMPACT ON BUSINESS PROCESSES, Zagreb, 2012.
[3] Hernandez-Ortega, B., Jimenez-Martinez, J. Performance of e-invoicing in Spanish firms. *Inf Syst E-Bus Manage* **11**, 457–480 (2013). https://doi.org/10.1007/s10257-012-0203-y
[4] Kaliontzoglou, A., Boutsi, P. & Polemi, D. eInvoke: Secure e-Invoicing based on web services. *Electron Commerce Res* **6**, 337–353 (2006). https://doi.org/10.1007/s10660-006-8678-6
[5] E-invoice, INTERNAL TECHNICAL INSTRUCTIONS, URL: https://www.efaktura.gov.rs/extfile/sr/1594/%D0%A1%D0%95%D0%A4%20%D0%98%D0%A2%D0%A3_24082023.pd, Accessed September 2023
[6] P. Zafiropulos, V. Smokovitis, N. Andjelic : Introduction to electronic invoicing and its aspects, Beograd 2022, ISBN: 978-86-80388-10-6
[7] E-government, URL: https://www.efaktura.gov.rs/tekst/342/servis-za-validaciju-dokumenata-u-xml-formatu-ubl-21-standard.php, Accessed September 2023
[8] E-government, XML invoice, URL: https://www.efaktura.gov.rs/view_file.php?file_id=345&cache=sr, Accessed September 2023
[9] E-invoice, portal, URL: https://www.efaktura.gov.rs/
[10] Blanca Hernandez-Ortega and Julio Jimenez-Martinez, Performance of e-invoicing in Spanish firms (2012) Inf Syst E-Bus Manage (2013) 11:457–480 DOI 10.1007/s10257-012-0203-y
[11] Valentin RADU, Andreea Cristina, XIAOYU ZHAN, Digitization of Business Relations in Romania by using Electronic Invoices (2022). Social Economic Debates Volume 11, Issue 1, 2022 ISSN 2360-1973; ISSN-L 2248-3837
[12] Marwin Heinemann, Wojciech Stiller, Digitalization and CrossBorder Tax Fraud: Evidence from E-Invoicing in Italy (2023). CESifo Working Papers ISSN 2364-1428 (electronic version)
[13] C. Marinagi, P. Trivellas, Panagiotis Reklitis1, C. Skourlas2, Adoption and Use of E-invoicing in Greece. AIP Conference Proceedings 1644, 279 (2015); doi: 10.1063/1.4907848
[14] Zvonimir Vanjak, Vedran Mornar, Ivan Magdalenić, DEPLOYMENT OF e-INVOICE IN CROATIA. ICSOFT 2008 - International Conference on Software and Data Technologies
E-invoice , LAW on electronic invoicing, LAW on electronic invoicing, https://www.efaktura.gov.rs/view_file.php?file_id=311&cache=sr, , Accessed Oktober 2023
[15] Dragan Radeski, Danco Davcev, ELECTRONIC INVOICING MODELING, The 8th International Conference for Informatics and Information Technology (CIIT 2011), ©2011 Institute of Informatics

Chatbots – Architecture and Applications

Igor Vecštejn[1], Verica Gluvakov[1], Maja Gaborov[1]

[1]*University of Novi Sad, Technical Faculty „Mihajlo Pupin“, Đure Đakovića bb, Zrenjanin, Republic of Serbia*

igor.vecstejn@tfzr.rs , verica.gluvakov@tfzr.rs , maja.gaborov@tfzr.rs

Abstract

The paper presents history, technologies, applications in different directions fields, and a solution example of a chatbot that is used as an application that has the possibility to answer all questions about COVID-19, users could watch pictures and videos about the virus, look at live statistics and newest news, in Serbia, statistics by cities, FAQ, Chatbot will send the report to user's email with attachments about measurements with preventions. The technologies that were used are for backend Python with the framework Flask that is going to do the front-end part. Besides that, this work will use DialogFlow, with MongoDB for the Database, and RapidAPI to connect to API.

Keywords: Chatbots, DialogFlow, Intents, Application

1. Introduction

The usage of Artificial Intelligence in every sphere is rapidly evolving and expanding. One of the most popular Artificial Intelligence software applications is the Chatbot. A Chatbot is a conversational agent that simulates human conversation, through text or voice messages. The first conceptualization of the Chatbot is attributed to Alan Turing, who asked "Can machines think" in 1950 [1]. This question named the Turing test is considered by many to be the generative idea of Chatbots. The first Chatbot with Eliza's name was constructed in 1966. Eliza simulated a psychotherapist's operation, returning the user's sentences in the interrogative form [2]

Chatbots are conversational or interactive agents that provide instant responses to the user. They are increasingly being used to improve interaction in this world of technology where communication and other activities rely on new technology and the Internet. One of the first goals of a Chatbot was to interact with the user just like a human [1]. Most people own and use smartphones which makes them frequent users of the Internet [3]. Chatbots can be deployed on Mobile and Web applications. These programs can't only greatly improve workaround and support, but to lessen administrative work, allowing them to automate jobs [4]. When it comes to the Health field, it is very important to answer correctly and precise. That's why it is important for chatbots in the Health field to have correct and precise answers for every user request.

The main goal of this paper is to show the architecture and implementation of a Web application that has communication with Chatbot. A chatbot should be able to answer most of the questions that users have about Covid 19 (FAQ). Also, users can get through the menu for urgent information if they have some symptoms or need to call urgent numbers so doctors can offer quick assistance. Besides that, users can get some information about COVID-19 through videos, pictures with explanations, and statistics about infected in the world, Serbia, and their districts.

2. Related works

In recent years, chatbots have garnered significant attention and research interest. A noteworthy milestone in their history was the creation of Eliza in 1966, which emulated a psychotherapist by responding to users' statements with questions. These automated conversational agents have found applications in diverse fields like healthcare, customer service, and education. Researchers and professionals have recognized their potential and have embarked on studies to enhance chatbots using advanced technologies, including deep learning

neural networks and natural language processing. These efforts aim to augment the capabilities of chatbots across various domains, underscoring the dynamic evolution of this technology [5]. In light of substantial advances in computing power and the progression of machine learning methodologies, chatbots have garnered substantial interest within the corporate and industrial sectors. Comprehensive exploration was undertaken to investigate the utilization of chatbots in disseminating healthcare information, offering support, identifying and proactively mitigating specific behaviors, and bolstering medical treatments across various domains of medical practice. This investigation underscores the adaptability and promising prospects that chatbots present within the healthcare arena [5]. The research conducted by Mittal, Battineni, Singh, Nagarwal, and Yadav addresses the need for improved communication in local hospitals, traditionally managed by patient associations. Online hospital data often lacks transparency and reliability. To overcome this, the study presents a web chatbot framework. The chatbot employs machine learning techniques such as gradient descent and natural language processing to manage user queries effectively. These algorithms handle both text and voice messages. The chatbot serves as a valuable communication tool between users and hospital staff, offering quick access to local hospital information and potentially reducing overcrowding [6]. Chatbot content management is often challenging and lacks standardization, potentially leading to poor user experiences. To address this issue, this study introduces the Chatbot Management Process—a methodology tailored for chatbot systems. Study and development were done by Andrade, Silva, Duarte, Santos, Costa, and Sousa. Developed based on experiences with the Evatalk chatbot for the Brazilian Virtual School of Government, this methodology focuses on iterative content improvement through user interaction analysis. It consists of three phases: management, construction, and analysis, and defines clear roles for the chatbot team. Applied to Evatalk, the methodology reduced the chatbot's human hand-off rate, increased its knowledge base, maintained high response confidence, and upheld user satisfaction [7].

3. Subject and goal of the research

In the last couple of years, developers developed a lot of Chatbot software in medicine to help doctors conclude results and treatments for patients. For the past 3 years, Covid-19 made people look for other solutions to their life and made them think and do stuff differently, most do get used to technology, the Internet. This paper refers to a custom solution Chatbot application that will have the ability to give users fast information about COVID-19 through a menu where they can choose if they have some ordinary questions, urgent questions, want to learn more about this virus, look at statistics that will be fetched through API, etc.

4. The architecture of the Chatbot System

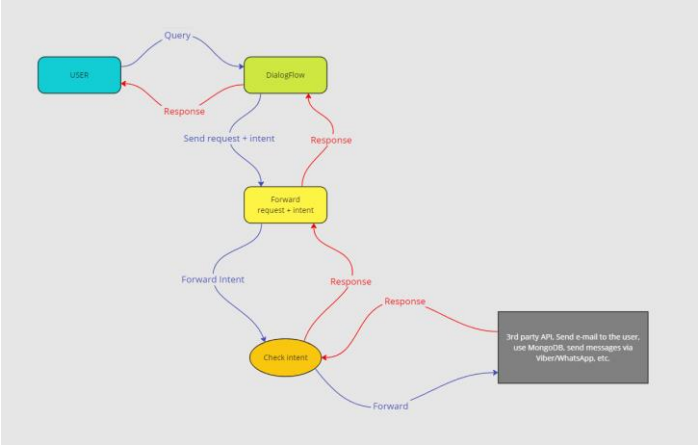


Figure 1. Chatbot CovidApp Solution Architecture

In Figure 1. there are colors of arrows that need to be explained. The red color stands for the bot's response and the blue one stands for the user's query. There is some order with these functions. The first ones that are going to be executed are the ones with the blue arrow. So, the user will ask questions in the form of text, and DialogFlow will try to match with the existing intent. After that, the request (includes some parameters within it in the form of the name, email, etc.) and intent will be sent to Python's Flask API that is deployed on Heroku. Within Python's block, there are a couple of functions. Every function will perform a different task. After sending the request and certain intent, the request and intent are now in the Python block. His job is to parse requests and intent to other functions so that the function will check and decide if that is going to be performed or not. If performed it will use 3rd Party API (Send mail, MongoDB, Whatsapp, Messenger, Telegram, Viber, Slack, Twitter, etc.). After that response will be followed from Python to DialogFlow and DialogFlow will send back a response to the user.

5. Used technologies and Practical implementation

- **Dialogflow -** Dialogflow is Google's Chatbot platform that is used to create conversations in Artificial Intelligence that work on Natural Language Processing (NLP). Developers can train a Dialogflow Essentials agent through the Dialogflow Web console. It's based on intents. An intent categorizes an end user's intention for one conversation turn. Your Dialogflow agent will contain many intents. Each intent contains training phrases, contexts, and responses. Dialogflow can interpret human conversation's meaning when an agent has trained with a given dataset of training phrases. In Figure 2. there is a list of intents that can be used.

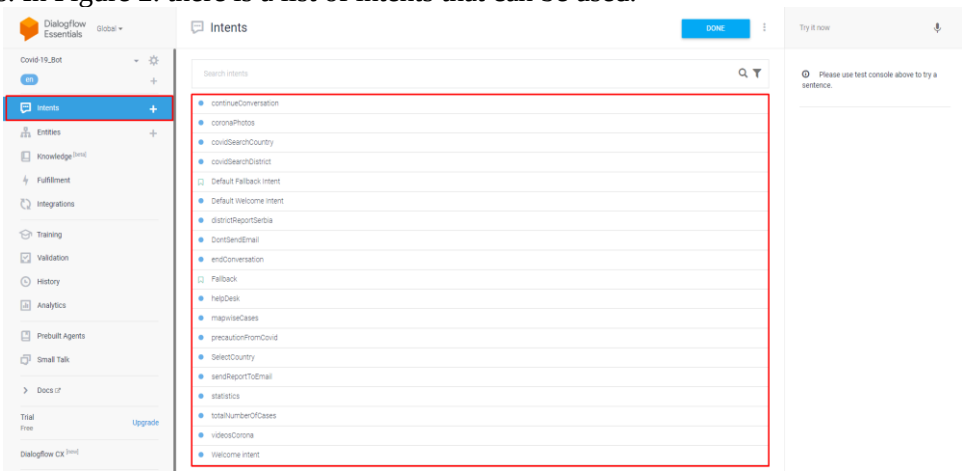


Figure 2. List of Intents

When a user communicates with the chatbot through written text or spoken voice (user utterances), Dialogflow will do intent classification. Because of the NLP engine (which gets continuous improvements over time), it doesn't matter if the user utterance is written differently or contains spelling/grammar mistakes. It will understand what has been asked and match it by checking all the intents added to the Dialogflow agent. Once there's an intent match, it will return the intent fulfillment/response. Dialogflow does not self-learn in the way that a goal-driven AI such as AlphaZero does. Google does not disclose all implementation details (secret sauce) of the underlying technologies such as the NLP or Speech to Text engines, simply because they may change rapidly and fall out of date with the documentation [8].

The most important terminologies in Dialogflow are:

- **Agent -** A Dialogflow agent is a virtual agent that handles concurrent conversations with your end-users. It is a natural language understanding module that understands the nuances of human language. Dialogflow translates end-user text or audio during a conversation to structured data that your apps and services can understand. You design and build a Dialogflow agent to handle the types of conversations required for your system [9]. A Dialogflow agent is similar to a

human call center agent. You train them both to handle expected conversation scenarios, and your training does not need to be overly explicit.

- **Intent -** An intent categorizes an end-user's intention for one conversation turn. For each agent, you define many intents, where your combined intents can handle a complete conversation. When an end-user writes or says something, referred to as an end-user expression, Dialogflow matches the end-user expression to the best intent in your agent. Matching an intent is also known as intent classification. For example, you could create a weather agent that recognizes and responds to end-user questions about the weather. You would likely define an intent for questions about the weather forecast. If an end-user says "What's the forecast?", Dialogflow would match that end-user expression to the forecast intent. You can also define your intent to extract useful information from the end-user expression, like a time or location for the desired weather forecast. This extracted data is important for your system to perform a weather query for the end user [9].
- **Entity -** Each intent parameter has a type, called the entity type, which dictates exactly how data from an end-user expression is extracted. Dialogflow provides predefined system entities that can match many common types of data. For example, there are system entities for matching dates, times, colors, email addresses, and so on. You can also create your own custom entities for matching custom data. For example, you could define a vegetable entity that can match the types of vegetables available for purchase with a grocery store agent [9].
- **Training phrase -** Training phrases are example phrases for what end-users might type or say, referred to as end-user expressions. For each intent, you create many training phrases. When an end-user expression resembles one of these phrases, Dialogflow matches the intent. For example, the training phrase "I want a burger" trains your agent to recognize end-user expressions that are similar to that phrase, like "Get a burger" or "Order a burger" [9].
- **Context -** Dialogflow contexts are similar to natural language contexts. If a person says to you "They are orange", you need context in order to understand what "they" is referring to. Similarly, for Dialogflow to handle an end-user expression like that, it needs to be provided with the context in order to correctly match an intent. Using contexts, you can control the flow of a conversation. You can configure contexts for intent by setting input and output contexts, which are identified by string names. When an intent is matched, any configured output contexts for that intent become active. While any contexts are active, Dialogflow is more likely to match intents that are configured with *input contexts* that correspond to the currently active contexts [9].

- **Knowledge base -** A knowledge base represents a collection of knowledge documents that you provide to Dialogflow. Your knowledge documents contain information that may be useful during conversations with end-users. Some Dialogflow features use knowledge bases when looking for a response to an end-user expression. This guide describes how to create and manage knowledge bases [9]. An example can be seen in Figure 3.

Figure 3. Dialogflow – Knowledge base

- **Fulfillment** - By default, your agent responds to a matched intent with a static response. If you're using one of the integration options, you can provide a more dynamic response by using fulfillment. When you enable fulfillment for an intent, Dialogflow responds to that intent by calling a service that you define. For example, if an end-user wants to schedule a meeting on Friday, your service can check your database and respond to the end user with availability information for Friday. Each intent has a setting to enable fulfillment. If an intent requires some action by your system or a dynamic response, you should enable the fulfillment of the intent. If an intent without fulfillment enabled is matched, Dialogflow uses the static response you defined for the intent [9]. On the Fulfillment page, there are fields URLs, Basic Authentication where the user needs to place a username and password, and some custom headers that will be used in the application.
- Rapid API is the world's largest API Marketplace. By using Rapid API, developers, and users can search, test, connect, and develop applications using certain APIs.
- **Flask -** Flask is a Python framework that is also called a Micro framework because it does not require any additional tools or libraries. It does not contain an abstraction-level database. It is designed as an extensive framework, i.e. a solid core offering basic services and an extension offering the rest. Flask contains three types of dependencies. Routing, Debugging, and Web Server Gateway Interface.
- **MongoDB -** MongoDB is a cross-platform, document-oriented database that provides, high performance, high availability, and easy scalability. MongoDB works on the concept of collection and documentation. A database is a physical container for collections. Each database gets its own set of files on the file system. A single MongoDB server typically has multiple databases. A collection is a group of MongoDB documents. It is the equivalent of an RDBMS table. A collection exists within a single database. Collections do not enforce a schema. Documents within a collection can have different fields. Typically, all documents in a collection are of similar or related purpose [10].
- **Postman -** Postman is one of the most successful tools used in API examination. Postman is an API client that makes it easy for developers to create, share, test, and document APIs. With this open-source solution, users can create and save simple and complex HTTP/s requests, as well as read their responses.

6. Practical Implementation and Functions of Chatbot

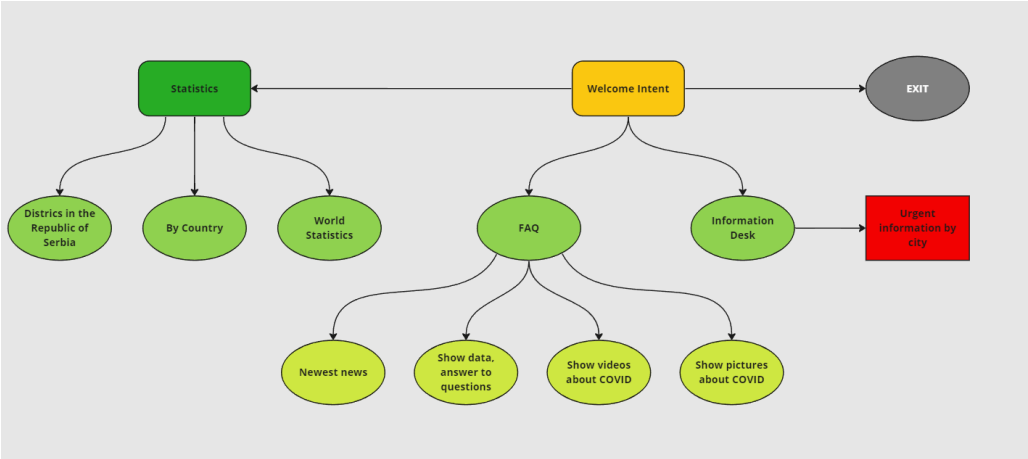


Figure 4. Chatbot CovidApp solution's possibilities and functions

First, the user needs to send a message in a certain textbox (Viber, Whatsapp, Telegram, etc.) that will be recognizable from the DialogFlow side. A chatbot will return one of the predefined messages and the software will offer a menu that will contain four main options (Statistics, Frequently Asked Questions (FAQ), Information Desk, and Exit). In Figure 4 there are shown Options and steps that will follow certain options after choosing one of the options. At option Statistics, users can get information about the numbers of people that were infected, recovered, or died by District in Serbia. Also, there is an option that can show statistics for every single country or statistics that will show numbers for the whole world. In Figure 5. there is an example of one city that has information about the first case in Serbia and other bigger towns in Serbia that contain other information about the virus.

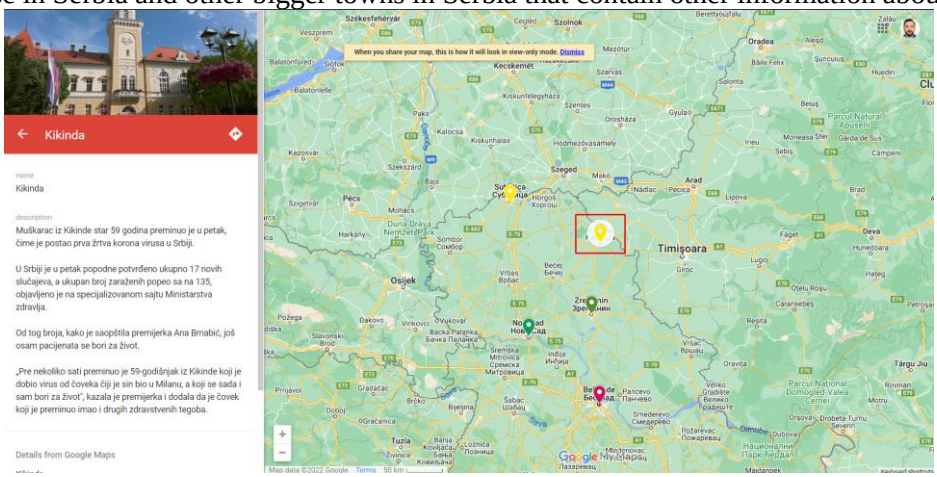


Figure 5. Information about the first case in the selected town

With the option Frequently Asked Questions (FAQ), users can gain new options like information about the newest news, be shown additional data, the bot can answer specific questions, and based on questions, he can send data, attach files via mail, and send it. There are options for videos of Covid-19. A bot can offer a couple of videos to users. The same will go for options with pictures. All videos/pictures are educational type, with introductions, and explanations about viruses, crucial tips provided by doctors. There is an Information Desk that provides an Urgent information option. The urgent information option will provide necessary tips for people who do not feel well and have symptoms like COVID-19. There will be provided an urgent phone number of doctors, and hospitals that are in charge of every city in Serbia.

7. Applications in other fields

- **Chatbot in education –** The growing demand for learning leads to high competition in higher education institutions. One of the critical reasons for sparse learning and high dropout rates is the fact that, when the number of students grows, the assistance the students get from their teachers is reduced. Chatbots, with their ability to provide educational content and personal assistance, come to support other e-learning practices [11]. Chatbots for learning support can preserve information by repeating old lessons when the students miss them. They also gather information during a course, which helps the improvement of the learning process and teaching. Students are facilitated in the study as Chatbots can answer questions concerning the educational material. A chatbot can also help students with school administration issues, such as enrolling in a course the exam schedule, their grades, and other related details to their studies so that the pressure on the school departments is considerably reduced [12].
- **Robotics** – The most crucial area of research on Chatbots is the natural language interface, which is a critical area for physical robots. Therefore, in the field of physical robots, we find abundant applications of natural language. For example, a novel natural language interface is developed for the autonomous robot, called Kamro[15].
- **Health** – In health care, chatbots are designed to provide patients with customized health and therapy information, patient-related products and services, offer diagnoses, and suggest treatments based on patient symptoms[14].
- **Client service** – The development of new technologies has made people interact with each other, and so has their commerce with businesses. E-commerce has evolved and fully changed the way companies vend their products, but there are some problems related to the quality of client service. Especially in live exchanges, the waiting time for a business hand to respond may be long, and the answers may not always be applicable(13). multitudinous companies use chatbots to support accounts. customer care is available 24 hours a day via Chatbot, enabling consumers to post their requests anyhow of the standard operating hours, which enhances user satisfaction(12).

8. Conclusion

The COVID-19 pandemic has presented unprecedented global health challenges and prompted innovative solutions in various sectors, particularly healthcare. One notable example of this innovation is the emergence of the COVID-19 Chatbot application, which has played a crucial role in disseminating essential medical information. Upon reflecting on its development, implementation, and reception, several key insights come to light.

First and foremost, the success of the COVID-19 Chatbot underscores the significance of technology in bridging the information gap during public health crises. Individuals from diverse backgrounds, including frontline medical professionals, educators, and technologists, have found value in the Chatbot's ability to provide quick and accurate information. This aligns with the broader trend towards technology-enabled healthcare solutions, where Chatbots can complement the efforts of medical professionals, educators, and information seekers.

Secondly, the positive feedback and high user engagement metrics serve as a testament to the user-centric design and effectiveness of the Chatbot. The application's user-friendly interface and its capacity to deliver information in a digestible manner have contributed to its widespread acceptance. Furthermore, the Chatbot's adaptability in meeting the diverse information needs of stakeholders within the medical ecosystem highlights the versatility of Chatbot technology.

In addition to its accomplishments during the pandemic, it's important to acknowledge that Chatbots have a broader role in the healthcare landscape. As medical technology advances, Chatbots can serve as valuable tools for various purposes, such as disseminating knowledge, educating patients, and even supporting telemedicine. With their capacity to provide timely responses and reliable information, Chatbots can be excellent companions for medical professionals, aiding them in routine healthcare inquiries as well as crisis management. However, it is important to acknowledge that Chatbots also present challenges. Ethical considerations, privacy concerns, and the need for robust regulatory frameworks must be addressed as Chatbots become more integrated into healthcare systems.

Furthermore, ongoing research to refine Chatbot capabilities, including natural language understanding and the handling of complex medical queries, remains crucial to fully unlock their potential.

In summary, the COVID-19 Chatbot application stands as a compelling example of the potential of Chatbot technology in the medical field. Its ability to provide fast, reliable, and easily accessible information during a global health crisis demonstrates the transformative power of this technology. As we look ahead, Chatbots are poised to become invaluable allies in healthcare, working alongside medical professionals to improve patient care, prevent diseases, and facilitate the dissemination of medical knowledge.

References:

[1] Khadija A., Zahra F., Naceur A. "AI-Powered Health Chatbots: Toward a general architecture", The 2nd International Workshop on AI & IoT, 2021, Leuven, Belgium

[2] Weizenbaum J. "Eliza-A computer program for the study of natural language communication between man and machine", Commun ACM, 1966

[3] Okonkwo C., Ade-Ibijola A., "Chatbots applications in education: A systematic review", Computers and Education: Artificial Intelligence 2, 2021

[4] Cunningham N., Boles W., Trouton L., Margerison E. "A review of chatbots in education: Practical steps forward", Annual Conference for the Australasian Association for Engineering Education, 2019, pp. 299-306

[5] Fan, X., Chao, D., Zhang, Z., Wang, D., Li, X., & Tian, F. (2021). Utilization of self-diagnosis health chatbots in real-world settings: a case study. Journal of Medical Internet Research, 23(1), e19928. https://doi.org/10.2196/19928

[6] Mittal, M., Battineni, G., Singh, D., Nagarwal, T., & Yadav, P. (2021). Web-based chatbot for frequently asked queries (faq) in hospitals. Journal of Taibah University Medical Sciences, 16(5), 740-746. https://doi.org/10.1016/j.jtumed.2021.06.002

[7] Andrade, G., Silva, G. R. S., Duarte, F. C. M., Santos, G. A., Costa, J. P. V., & Sousa, R. T. d. (2022). A conversation-driven approach for chatbot management. IEEE Access, 10, 8474-8486. https://doi.org/10.1109/access.2022.3143323

[8] Boonstra L. "The Definitive Guide to Conversational AI with Dialogflow and Google Cloud", APress, 2021

[9] DialogFlow. (2020). DialogFlow ES Documentation, https://cloud.google.com/dialogflow/es/docs/

[10] TutorialsPoint. (2016). MongoDB – NoSQL Document Database https://www.tutorialspoint.com/mongodb/mongodb_tutorial.pdf

[11] Colace F., De Santo M., Lombardi M., Pascale F., Pietrosanto A., Lemma S. "Chatbot for e-learning: A case study", International Journal of Mechanical Engineering and Robotics Research, 2018, pp. 528-533

[12] Adamopoulou E., Moussiades L., "Chatbots: History, Technology, and Applications", Machine Learning With Applications 2, Elsevier, 2020.

[13] Lueth T. C., Laengle T., Herzog G., Stopp E., Rembold U. "Kantra human-machine interaction for intelligent robots using natural language", 3rd IEEE International workshop on robot and human communication, 1994, pp. 106-111

[14] The top 12 health care chatbots, The Medical Futurist Website: The Top 12 Healthcare Chatbots - The Medical Futurist, Retrieved 24th September 2022

[15] Nuruzzaman M., Hussain O., "A survey on chatbot implementation in a customer service industry through deep neural networks", IEEE 15th International Conference on e-business Engineering, 2018, pp. 54-61

Comparative Study of React, Angular, and Vue for Front-end Development

Tamara Milić[1], Igor Vecštejn[1], Eleonora Brtka[1], Maja Gaborov[1]

[1]*University of Novi Sad, Technical Faculty „Mihajlo Pupin“, Đure Đakovića bb, Zrenjanin, Republic of Serbia*

tamara.milic@tfzr.rs, igor.vecstejn@tfzr.rs, elenora.brtka@tfzr.rs, maja.gaborov@tfzr.rs

Abstract

The rapid evolution of front-end development has brought about a significant increase in complexity and importance. What was once considered a peripheral aspect of web development has now taken center stage, playing a vital role in driving user engagement and shaping the digital landscape's aesthetics, performance, and functionality. This transformation has been made possible by the emergence of JavaScript frameworks, which have evolved from mere tools into indispensable pillars supporting modern web development. These frameworks provide developers with structured paradigms, reusable components, efficient rendering engines, and effective strategies for managing data. They have effectively pushed the boundaries of web development, enabling the creation of responsive and interactive user interfaces. In light of this architectural shift, JavaScript frameworks have become the foundation of front-end development. They abstract away the complexities of direct manipulation of the Document Object Model (DOM), offering developers abstractions that facilitate component reusability, state management, and efficient rendering. This abstraction allows developers to focus more on the logic and functionality of their applications, rather than getting bogged down in low-level DOM manipulation.

Keywords: Web Development, Frameworks, JavaScript, React, Angular, Vue

1. Introduction

The ever-evolving landscape of modern web development is a complex tapestry of technologies, methodologies, and paradigms. In this era of convergence, the choice of a front-end JavaScript framework holds immense importance, not just as a technical decision, but as a strategic one. This scholarly paper is the result of rigorous research and empirical exploration, embarking on a discerning journey through the multifaceted realm of front-end development. Its primary focus is to thoroughly examine three prominent frameworks—React, Angular, and Vue—and provide the scholarly community, practitioners, and decision-makers with an enlightened understanding of these tools. By employing empirical methodologies, theoretical constructs, and real-world applicability assessments, we aim to shed light on the intricate maze of front-end development, empowering informed decision-making at the intersection of technology and user experience.

The ever-evolving landscape of web development has experienced a paradigm shift in its architectural foundations. This shift has been driven by the increasing demand for real-time data interaction, dynamic content generation, and compatibility across multiple platforms. As a result, there has been a departure from traditional static web pages towards more modular and component-based architectures. These new architectures serve as the building blocks for modern web applications, enabling reusability, scalability, and maintainability.

In response to this architectural transformation, JavaScript frameworks have emerged as the cornerstone of front-end development. They simplify the complexities associated with direct manipulation of the Document Object Model (DOM) by providing abstractions that promote component reusability, efficient state management, and rendering capabilities [1]. By leveraging these frameworks, developers can focus on creating dynamic and interactive user interfaces without being burdened by low-level DOM manipulation. JavaScript frameworks have undoubtedly played a pivotal role in

shaping the front-end development landscape, offering developers powerful tools to streamline their workflows and deliver exceptional user experiences.

The confluence of three prominent JavaScript frameworks sets the stage for an intriguing discourse:

- React: Born out of Facebook, React has revolutionized front-end development with its Virtual DOM and component-based architecture. Its impact is particularly evident in the realm of single-page applications (SPAs), where it excels in performance optimizations [2].
- Angular: Hailing from the Google stable, Angular presents an extensive ecosystem comprising various tools and libraries. It advocates for a declarative approach to UI development, emphasizing modularity, testability, and comprehensive solutions [3].
- Vue: Conceived by Evan You, Vue embodies simplicity and adaptability. It embraces a progressive framework philosophy, with a primary focus on the view layer i.e. Web interfaces and applications with rich interactivity, One-page applications, Desktop, and mobile applications. Its flexible adoption path suits diverse architectural contexts well [4].

These frameworks each bring their unique strengths and philosophies to the table, fueling passionate discussions and shaping the choices made by developers in their quest for efficient and effective front-end development.

2. Performances of Front-end framework

- React's Performance - React's Virtual DOM has brought about a revolutionary impact on front-end performance [2]. By introducing a virtual representation of the Document Object Model (DOM), React minimizes the need for direct DOM manipulation, resulting in reduced costly operations and enhanced rendering speed. This virtual representation plays a crucial role in React's ability to achieve swift initial load times in applications [5]. React's reconciliation algorithm plays a vital role in ensuring efficient updates to the user interface. By intelligently identifying and isolating changes in the virtual DOM, React ensures that components only re-render when necessary. This approach minimizes browser reflows and unnecessary re-renders, ultimately leading to a highly responsive user experience [5]. React's focus on optimizing rendering performance through the Virtual DOM and its intelligent reconciliation algorithm has solidified its position as a powerful framework for building fast and efficient web applications. Its ability to minimize unnecessary DOM operations and deliver responsive updates contributes significantly to creating a smooth and seamless user experience.
- Angular's Rendering Strategies - Angular takes a distinct approach to rendering by adopting a two-way data binding mechanism, which enables automatic updates of the view whenever changes occur in the data [3]. This approach offers convenience and simplifies development. However, it is worth noting that in large-scale applications, this extensive data binding can potentially introduce performance bottlenecks. To address this concern, Angular provides change detection strategies that offer developers more control over how components are checked for updates. One such strategy is the OnPush change detection strategy. With OnPush, components are marked to check for updates only when their input properties change, reducing the frequency of checks and enhancing performance in specific use cases [2]. By allowing developers to choose between different change detection strategies, Angular empowers them to optimize performance based on the specific requirements of their applications. The OnPush strategy, in particular, can be leveraged to minimize unnecessary checks and improve rendering efficiency in scenarios where the input data changes less frequently. This flexibility in change detection strategies is a valuable feature of Angular, enabling developers to strike a balance between convenience and performance in their applications.
- Vue's Reactive System - Vue's rendering performance is built upon its reactivity system [4]. Vue effectively tracks dependencies between components and data, ensuring that only the components affected by data changes are updated. This targeted approach minimizes unnecessary re-renders, resulting in optimal rendering performance. While Vue's rendering speed may slightly lag behind React due to its simpler reactivity model, this difference is often imperceptible to users. Moreover, Vue's ease of use and approachability outweigh any minor disparities in rendering speed [4]. Vue prioritizes developer experience and strives to provide a smooth learning curve, enabling developers to quickly grasp its concepts and build efficient applications. The emphasis on simplicity

and ease of use is a notable strength of Vue. Its reactivity system strikes a good balance between rendering performance and developer productivity, making it an attractive choice for projects where user experience and developer friendliness are paramount.

2. Component-Based Architecture

- React's Component Model - React, with its component-based architecture, has been a driving force in the widespread adoption of reusable components. React components encapsulate both the UI and the logic, promoting reusability and maintainability. Its Virtual DOM and JSX syntax provide a seamless and intuitive way to compose components, making React a popular choice for building modular and reusable UI elements. Each React component encapsulates its own logic, rendering, and state management, promoting modularity, reusability, and maintainability [5]. This modular approach allows developers to compose complex user interfaces by combining smaller, self-contained components, fostering a structured and scalable development process.

The example content of one React application is shown in Figure 1. React intern state of a component refers to the internal state of a React component. In React, the state is used to store and manage the data that a component needs to render and update. Arguments in React can refer to the props passed to a component, the state of a component, or the events that trigger component updates. The life cycle of React methods refers to the sequence of events that occur during the different stages of a component's existence. React provides several methods that allow a user to control what happens at each stage. These methods include component mounting, updating, and unmounting components. This architectural choice simplifies debugging and makes it easier to understand and track how data changes propagate throughout the application [5]. By maintaining a clear flow of data, React enables developers to reason about the state and behavior of their components, enhancing code predictability and maintainability.

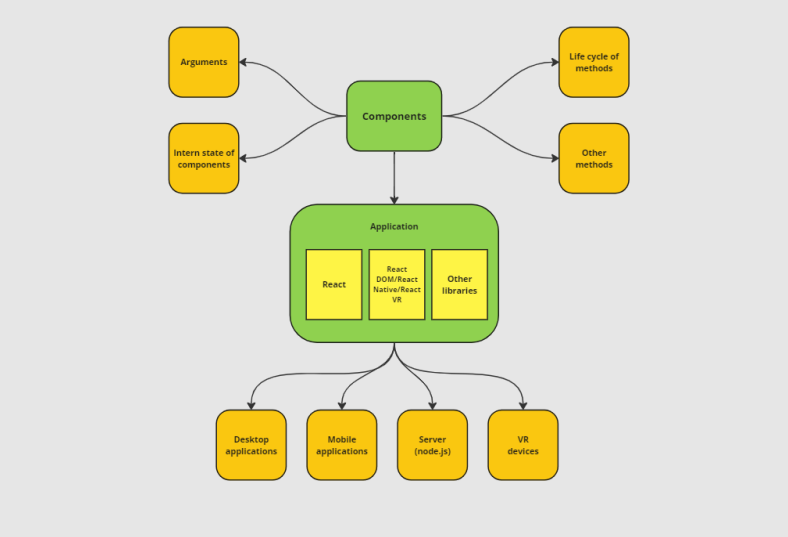


Figure 1. Content of React application

The combination of component-based architecture and unidirectional data flow in React empowers developers to build reusable, modular, and well-structured user interfaces. This approach has been widely embraced by the React community and has contributed to the framework's popularity and success in the front-end development ecosystem.

- Angular's Component-Based Approach - Angular offers a comprehensive approach to component-based architecture. Angular components are the building blocks of applications, providing a clear separation of concerns. With features like templates, dependency injection, and powerful data binding capabilities, Angular empowers developers to create reusable components that are highly testable and scalable. Angular fully embraces the concept of component-based architecture in its pursuit of modularity and maintainability [3]. This modular approach enables

developers to create self-contained, reusable units that can be easily managed and composed to build complex user interfaces. The content of the Angular application is shown in Figure 2. Angular allows users to create and manage custom components, directives, services, and other files to organize their codebase and enhance reusability. By using the Angular CLI (Command Line Interface), users can generate various types of files including components, services, modules, and more. Data binding in Angular is a powerful feature that allows users to connect an application's data to its user interface. It enables users to update the application's UI automatically when the underlying data changes. A full line-up in Angular refers to the complete list of components, directives, services, and modules that make up an Angular application. It is important to have a clear understanding of the full lineup in order to effectively develop and maintain an Angular application. Dependency injection is a software design pattern used in Angular and other frameworks to manage dependencies between components. It allows for loosely coupled code and makes it easier to test and maintain applications. In Angular, testing is an integral part of development to ensure the reliability and correctness of users' applications. There are various types of tests users can perform, including unit tests, component tests, and end-to-end tests.

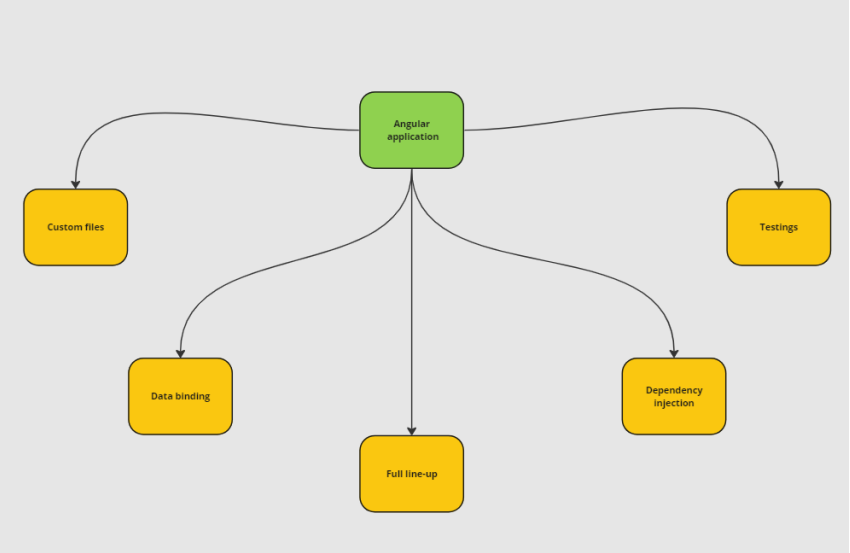


Figure 2. Content of Angular application

One of the strengths of Angular's component-based architecture lies in its robust dependency injection mechanism. Angular's dependency injection allows components to declare their dependencies explicitly, making it easier to manage and control the relationships between components. This flexibility is particularly advantageous in large-scale applications where there are intricate dependencies among components [1]. By providing a clear and efficient way to manage component dependencies, Angular promotes maintainability and scalability in complex projects. The combination of component encapsulation, reusability, and the powerful dependency injection mechanism makes Angular a compelling choice for developers who prioritize modularity and maintainability in their applications.

- Vue's Flexible Component System - Vue indeed offers a flexible and powerful component system that strikes a balance between simplicity and robustness [4]. Vue components are self-contained units that encompass templates, scripts, and styles, allowing developers to create reusable components with varying degrees of complexity based on their project requirements. One of Vue's standout features is its seamless integration of a reactivity system with its component model. This integration enables Vue components to automatically re-render whenever the underlying data changes. This reactivity ensures that Vue components are highly responsive and efficient, making them particularly well-suited for scenarios where real-time updates are essential [4]. Whether it's updating a user interface based on user input or displaying dynamic content from an API, Vue's reactivity system ensures that components stay in sync with the data they rely on.

By providing a flexible component system and integrating a powerful reactivity system, Vue empowers developers to create reusable and highly reactive components. This combination of simplicity and power makes Vue a popular choice for a wide range of projects, from small-scale applications to large-scale enterprise solutions.

3. Learning Curve

- React's Learning Curve – React's learning curve is often praised for its accessibility and simplicity [2]. The component-based architecture, JSX syntax, and unidirectional data flow in React are intuitive concepts that many developers can grasp quickly. This makes it easier for developers to understand and work with React, enabling them to build applications efficiently. React's extensive documentation and the abundance of online resources available also contribute to its learnability. Developers can find comprehensive guides, tutorials, and examples that help them navigate the framework and understand its concepts effectively. The vibrant React community actively shares knowledge and provides support, fostering a positive learning experience. The React ecosystem offers a wide array of tools and libraries that enhance the developer experience. For instance, Redux provides a robust solution for state management, while React Router simplifies the process of implementing routing in React applications. These tools and libraries integrate seamlessly with React, further streamlining the development process and improving productivity [5].

- Angular's Learning Curve - Angular's learning curve is often considered to be steeper compared to some other front-end frameworks, primarily due to its comprehensive nature and extensive feature set. However, this depth is what provides Angular with a holistic solution for building complex applications. Developers working with Angular need to become acquainted with TypeScript, which is Angular's recommended language, as well as concepts like decorators and modules [3]. TypeScript, with its strong typing and enhanced tooling support, brings additional robustness and maintainability to Angular projects. While it may require developers to learn and adapt to TypeScript's syntax and features, it offers benefits such as improved code quality, better collaboration, and enhanced developer experience.

Angular's CLI (Command Line Interface) is a powerful tool that simplifies project setup and automates many development tasks. The CLI provides generators for components, services, and other Angular artifacts, making it easier and more efficient to create and manage different parts of an application. This streamlined development experience saves time and effort, enabling developers to focus on building the application logic. Angular's dependency injection system is a key aspect of the framework. It allows for the easy management of component dependencies, promoting code maintainability and facilitating collaboration among developers. This feature enhances the overall developer experience and contributes to the maintainability and scalability of Angular applications [1].

- Vue's Learning Curve – Vue prides itself on its approachable learning curve [4]. This characteristic makes Vue an appealing choice, especially for developers who are new to front-end development. Vue's design philosophy emphasizes simplicity and ease of use, making it easier for developers to grasp the core concepts and quickly start building applications.

The Vue CLI (Command Line Interface) further enhances the developer experience by providing a streamlined setup process and essential development tools. With the Vue CLI, developers can quickly scaffold new projects, manage dependencies, and take advantage of features like hot module replacement during development. This efficient setup process saves time and effort, enabling developers to focus on writing code and building applications. Vue's documentation is comprehensive and beginner-friendly, offering clear explanations, examples, and guides. This documentation, combined with the vibrant Vue community, ensures that developers have access to the necessary resources and support to learn and grow with Vue. The community actively shares knowledge, provides assistance, and contributes to the overall positive developer experience.

4. State Management

Each framework has its own approach to state management, providing developers with flexibility and options based on their project requirements. React's ecosystem offers external libraries like Redux, while Angular has NgRx, and Vue provides Vuex. These libraries enhance state management capabilities and enable developers to handle complex data interactions in a scalable and maintainable manner.

- React's State Management - React offers a flexible approach to state management, giving developers the freedom to choose between local component state and global state management solutions like Redux or Mobx [2]. This flexibility allows developers to tailor their state management approach based on the specific needs of their application. For managing component-specific data, React allows developers to utilize the local component state. Local state is suitable for handling data that is only relevant to a specific component and doesn't need to be shared across the application. React's useState hook or the traditional class-based state management provides simple and efficient ways to manage local states within components. In cases where the application requires a centralized and shared state across multiple components, global state management solutions like Redux or Mobx can be utilized. Redux, in particular, has gained significant popularity in the React ecosystem. It implements a unidirectional data flow pattern, simplifying state changes and ensuring predictable updates throughout the application.
Redux encourages the use of immutability, where state updates are made by creating new immutable objects instead of modifying existing ones. This approach helps maintain the integrity of the data and makes it easier to track changes. Additionally, Redux provides powerful debugging tools like Redux DevTools, which allow developers to inspect the state changes and actions, aiding in the development and debugging process [5]. By offering both local component state and global state management solutions like Redux or Mobx, React provides developers with the flexibility to choose the most suitable approach for their application's state management needs. Whether it's simple component-specific data or complex application-wide state, React offers the tools and libraries to handle states in an efficient and predictable manner.

- Angular's Reactive State Management - Angular leverages RxJS, a powerful reactive programming library, for state management [8]. This choice enables developers to create and manipulate observables, which are essential for handling asynchronous data streams in Angular applications. Developers have a couple of options for managing the application state. They can utilize BehaviorSubjects, which are a type of observable that maintains the current value and emits it to subscribers. BehaviorSubjects are useful for managing and sharing state across components and services within an Angular application. Angular provides an Angular-specific state management library called NgRx. NgRx is inspired by the principles of Redux and brings a similar unidirectional data flow to Angular applications. It follows the Redux pattern of actions, reducers, and selectors to manage the state in a predictable and scalable manner.
NgRx enforces strict immutability, meaning that state changes are made by creating new immutable objects rather than modifying existing ones. This approach ensures the integrity of the data and simplifies tracking state changes. NgRx also provides tools for debugging, including the Redux DevTools extension, which allows developers to inspect the state, actions, and time travel through state changes for effective debugging [8]. By leveraging RxJS and offering options like BehaviorSubjects and NgRx, Angular provides developers with powerful tools for managing the state of their applications. Whether it's handling asynchronous data streams or managing complex application-wide state, Angular's state management capabilities offer a robust and predictable approach to ensure the integrity and reactivity of the application.

- Vue's Approach to State - Vue offers a flexible state management approach, thanks to its built-in reactivity system [4]. Vue components can define reactive data properties that automatically update the view when their values change. This reactivity makes it easy to manage and observe the state within individual components. For more complex applications that require centralized state management, Vue seamlessly integrates with Vuex. Vuex is a state management library inspired by

Flux and Redux [9]. It provides a centralized store where the application state can be shared and mutated consistently across components.

Vuex enforces a strict unidirectional data flow, ensuring that state changes occur in a predictable manner. Actions are used to trigger mutations, which are responsible for modifying the state. This pattern helps maintain the integrity of the data and simplifies tracking state changes. Vuex offers debugging tools that allow developers to inspect state changes, actions, and mutations. These tools enhance developer productivity and make it easier to understand and debug complex state interactions, contributing to the maintainability of Vue applications [9]. By offering built-in reactivity and integrating seamlessly with Vuex, Vue provides developers with a flexible state management solution for applications of different sizes and complexities.

5. Ecosystem and Community Support

The ecosystems surrounding React, Angular, and Vue play a vital role in their success and sustainability as front-end frameworks. All three frameworks, React, Angular, and Vue, have thriving ecosystems with a wide range of libraries, tools, and resources. React benefits from its active community and support from Facebook. Angular enjoys the backing of Google and a dedicated community. Vue's ecosystem has grown rapidly, fueled by its popularity and a supportive community. The vitality of these ecosystems ensures that developers have access to a rich set of resources, enabling them to build robust and feature-rich front-end applications.

- React's Ecosystem - React's ecosystem is renowned for its robustness and the plethora of third-party packages and tools available. This vast library of resources allows developers to enhance their React applications with various functionalities and features. One significant advantage of React's ecosystem is its flexibility and extensibility. With React, developers have the freedom to choose from a wide range of libraries and solutions that best suit their project requirements. Whether it's UI component libraries like Material-UI or styled components, state management solutions like Redux or Mobx, or routing libraries like React Router, there is a wealth of options available to cater to different needs.

A notable aspect of React's ecosystem is the strong presence of open-source projects and community contributions. The active engagement of the React community fosters collaboration, innovation, and continuous improvement. Developers can leverage the collective knowledge and expertise of the community to address challenges, find solutions, and stay up to date with the latest trends and best practices in React development [7]. This vibrant community involvement also ensures that React remains at the forefront of modern web development. It results in the rapid development of new tools, frameworks, and techniques that further enhance the React ecosystem. From build tools like Webpack and Babel to testing frameworks like Jest and Enzyme, the React ecosystem offers a wealth of resources for developers to build high-quality, performant, and scalable applications.

- Angular's Comprehensive Ecosystem - Angular offers a comprehensive ecosystem that covers a broad spectrum of tools and libraries, making it a powerful framework for web development. One notable aspect of Angular's ecosystem is the Angular CLI (Command Line Interface). The Angular CLI simplifies project setup by providing a command-line interface for creating, configuring, and managing Angular applications. It automates various development tasks, such as generating components, services, and modules, as well as offering integrated testing, building, and deployment capabilities. This streamlined workflow significantly enhances developer productivity and efficiency.

Angular Material is another valuable component of the Angular ecosystem. It provides a collection of pre-designed UI components that align with Angular's design principles. Developers can leverage these components to create visually appealing and consistent user interfaces, saving time and effort in UI development. Angular's ecosystem includes NgRx, an official library for state management. NgRx follows the principles of Redux and brings a similar unidirectional data flow to Angular applications. It provides a robust solution for managing the application state, enforcing immutability and predictable updates [10]. Angular's ecosystem is supported by Google, the creators of Angular, which ensures its longevity and relevance. Additionally, Angular benefits from a dedicated community that actively contributes to its development, share knowledge, and provides support.

- Vue's Vibrant Ecosystem - Vue's ecosystem is known for its vibrancy and adaptability, making it a popular choice among developers. One notable aspect is the seamless integration of official projects with Vue applications. Vue Router provides a powerful and flexible routing solution, enabling developers to create dynamic and navigable applications. Vuex, Vue's official state management library, simplifies the management of application state, ensuring consistency and scalability.

The Vue CLI is another essential component of Vue's ecosystem. It offers a command-line interface that streamlines project configuration and development workflows. Vue's ecosystem is driven by a passionate and engaged community. This community actively contributes to the growth and development of Vue by creating plugins, extensions, and libraries that address different use cases and provide additional functionality. These plugins and extensions enhance Vue's versatility and make it suitable for a wide range of projects, whether it's integrating with UI component libraries like Vuetify or adding functionality like internationalization or form validation.

6. Conclusion and Comparative Analysis

React is an excellent choice if flexibility, performance, and a vast ecosystem are of utmost importance to users' projects. It excels in these areas, thanks to its Virtual DOM and unidirectional data flow. React also benefits from a vibrant community and a wide range of libraries. However, newcomers may find it challenging to manage the state and choose the right components from the ecosystem.

Angular is well-suited for large-scale applications that require a comprehensive solution. Its two-way data binding and dependency injection make it robust and maintainable. While the learning curve for Angular is steeper, it offers a unified and opinionated framework with a strong ecosystem. Consider Angular if the programmer's team is willing to invest time in learning its concepts and best practices.

Vue stands out for its approachability and adaptability. Its progressive framework philosophy allows for incremental integration into projects, making it suitable for both small and large-scale applications. Vue's simplicity is balanced by its reactivity system and state management capabilities through Vuex. Vue is a good choice, especially if users have developers who are new to front-end development.

In conclusion, comparative analysis provides insights into the strengths and weaknesses of React, Angular, and Vue. By considering project requirements, team expertise, and scalability needs, users can make an informed decision on which framework best suits their needs. Staying updated with the latest features and best practices will ensure that projects remain relevant in the ever-evolving landscape of front-end development.

Acknowledgment:

This research is financially supported by the Ministry of Education and Science of the Republic of Serbia under project number TR32044 “The Development of software tools for business process analysis and improvement”, 2011-2023.

References:

[1] A. Freeman, & J. Robson, "Angular Programming." Manning Publications, 2019.
[2] ReactJS. (2023). React - A JavaScript library for building user interfaces, https://reactjs.org/
[3] Angular. (2023). Angular - One framework. Mobile & desktop, https://angular.io/
[4] Vue.js. (2023). Vue.js - The Progressive JavaScript Framework, https://vuejs.org/
[5] D. Flanagan, "JavaScript: The Definitive Guide." O'Reilly Media, 2021.
[6] M. Tielens Thomas: "React in Action", Hanning Publications Co.,2018
[7] A. Bush: "React Essentials", Packt Publishing, 2015
[8] S. Seshadri: "Angular Up & Running", O’Reilly Media, 2018
[9] C. Macrae: "Vue.js Up & Running Building Accessible and Performant Web Apps", O’Reilly Media, 2018
[10] D. Kumar: "Angular Essentials", BPB Publications, 2019

Human-Computer Interaction Using XBOX Kinect Technology

Marjana Pardanjac[1], Snežana Jokić[1], Aleksandra Karuović[1], Isidora Jokić[2] and Marija Dunjić[1]
[1] *University of Novi Sad, Technical Faculty "Mihajlo Pupin", Đure Đakovića bb, Zrenjanin, Republic Serbia*
[2] *University Novi Sad, Faculty of Technical Sciences, Trg Dositeja Obradovića 6, Novi Sad, Republic Serbia*

marjana.pardanjac@tfzr.rs; snezana.jokic@tfzr.rs; aleksandra.karuovic@tfzr.rs; jokic.ra237.2022@uns.ac.rs; marija.dunjic@tfzr.rs

Abstract:
A typical way of human-computer interaction is through a mouse and keyboard. However, the possibility of using the human body in human-computer interactions is gaining more and more importance. The application of human movement as a means of communication with a computer is also interesting from an educational point of view. For this reason, the paper presents students' experiences with using XBOX Kinect technology in the teaching process. The obtained results show the satisfaction and motivation of students in working with XBOX Kinect technology.

Keywords:
human-computer interaction, XBOX Kinect technology, human movement, hand gesture

1. Introduction

Nowadays, various interactive systems are present in almost daily use. The average user encounters many devices and applications that are an integral part of everyday life, starting from portable devices (phones, notebooks, or tablets), through public devices (ATMs, coffee machines, and scheduling machines), to devices and applications such as printers, home appliances, computers in cars, sites for shopping, booking, learning, networking, gaming and many more [1].

XBOX Kinect is a device created by Microsoft for Xbox 360 (later also for Windows 7). Kinect in use has cameras and a microphone to see and hear the user. He uses the equation to discern what is human and what is not. At the same time, it allows the user to move his body and by moving his hands as well as his body, he controls the behavior of playing the game. The specificity of Kinect is that with it, games are played with the body, not with a joystick.

2. XBOX Kinect Technology

Kinect is an additional device offered by Microsoft as a controller for the Microsoft XBOX. It allows users to control using body gestures or voice commands, eliminating the need for standard controllers [2].

It was created to revolutionize the way people play games and how they experience entertainment, and in recent years its application has spread to education as well, where it finds its place in interaction with users. With Kinect, people can control their bodies through the console.

The Kinect sensor directly allows the computer to sense the third dimension (depth) of the user and the environment, making the task easier. It also understands when users are talking, knows who they are (recognizes them), and can interpret their movements and translate them into a format that developers can use to create new experiences.

2.1. Kinect Sensor

The Kinect sensor includes several advanced sensor hardware. It features a depth sensor, a color camera, and an array of four microphones that provide full-body 3D motion capture, facial recognition, and voice recognition capabilities [3].

Figure 1a shows the Kinect device itself, and Figure 1b shows the layout of the infrared (IR) projector, color camera, and IR camera. The depth sensor consists of an IR projector combined with an IR camera, which is a monochrome complementary metal oxide semiconductor (CMOS) sensor. The depth sensor technology is licensed from the Israeli company PrimeSense. Although the exact technology has not been disclosed, it is based on the principle of structured light. An IR projector is an IR laser that passes through a diffraction grating and turns into a set of IR dots.

Figure 2 shows the IR points seen through an IR camera. The relative geometry between the IR projector and the IR camera as well as the projected IR dot is known. If one can match a point observed in the image with a point in the projector pattern, it can be reconstructed in 3D using triangulation. Since the pattern of points is relatively random, matching between IR images and projector patterns can be done straightforwardly by comparing small neighborhoods using, for example, normalized cross-correlations.

Figure 3 shows the map created by the Kinect sensor for the IR image in Figure 2. The depth value is coded in gray values; the darker the pixel, the closer the dot is to the camera. Black pixels indicate that no depth values are available for those pixels. This can happen if the points are too far away (and depth values cannot be calculated accurately), or are too close (there is a blind area to limited fields of view for the projector and cameras), or reflect poor IR lights (such as hair or mirror surfaces) [3].

Figure 1. Microsoft Kinect sensor. (a) Kinect sensor for XBOX 360. (b) [3]

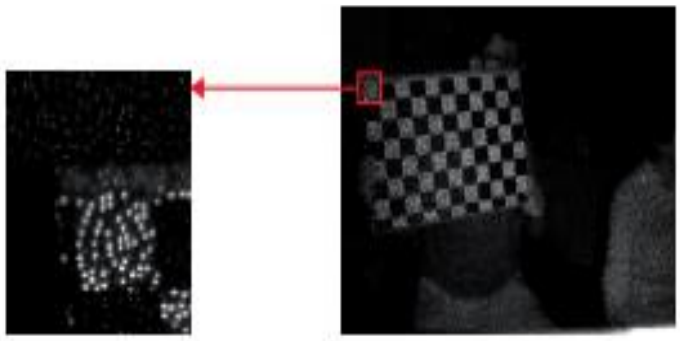


Figure 2. Infrared (IR) spots are seen by an IR camera. The image on the left shows a close-up of the red field[3]

Figure 3. Kinect sensor depth image. Depth sensor from the infrared (IR) point in Figure 2 [3]

2.2. Application of Kinect in education

Kinect supports kinesthetic pedagogical practice so that students can develop greater body-kinesthetic intelligence. It is a requirement that Kinect technology must be integrated with a computer, projector, and compatible software to facilitate the creation of meaningful interaction in the

classroom. Looking at Kinect from the perspective of a learning tool, due to the multiple types of interaction it supports, there is potential to improve classroom interaction [4] to increase student participation in class and improve professors' abilities to present and manage multimedia materials in class and to create opportunities to conduct interaction and discussion.

Using the Avatar Kinect software package, distance learning, project collaboration, and socialization between students can be animated. Avatar Kinect uses motion tracking along with facial recognition to assign each student a unique avatar, a real-time on-screen representation of the student that mimics the way the user smiles, speaks, and behaves. "Up to eight people in different physical locations can be connected. Using this package requires much less bandwidth than for regular video calls since real-time avatar animation requires slightly more bandwidth than voice call transmission [4]".

But despite all these possibilities of Kinect, a significant use would also be related to learning programming. "Students would interactively create programs, that is, avatar management applications in real-time, where, in addition to the aforementioned interactions, they would combine knowledge of object-oriented programming [4]". In this sense, students could in a very effective and interesting way animate some interactive processes, which would awaken in many the desire for further learning and work. Figure 4 shows an example of avatar control using Kinect.

Figure 4. Display of control of avatar parts via Kinect [4]

2.3. Restrictions

The unhindered functioning of Kinect is conditioned by technical limitations, primarily in the requirement of space in the form of a classroom. On the other hand, recording only one actor at a time is possible, while the other participants must either be still or out of frame, which can slow down the recording process. "The software that Kinect uses may be unavailable for some educational institutions or very expensive. Also, the device's price can be one of the problems [4]."

Kinect plays a very important role in educational systems, even advanced ones, in which Human-Computer Interaction is mentioned. To overcome the gap between the introduction of news in teaching and old learning methods, Kinect can initially serve as a tool in laboratory exercises in the mentioned field, so that both the professor and all students adapt to news in the teaching process.

3. Methods of Research

3.1. The aim of the research

The research aims to determine to what extent and in what way it is possible to apply XBOX Kinect technology to increase students' interest and better understanding of the material. Operational research should determine the role of Kinect technology in human-computer interaction and models of application in teaching.

3.2. Research hypothesis

The main hypothesis of the research is as follows: The application of XBOX Kinect technology encourages students to be more engaged in classes and more functional acquisition of knowledge through active learning.

The sub-hypotheses of the research are:

1. *The use of Kinect technology has a positive effect on human-computer interaction*
2. *Gender affects the ability to use XBOX Kinect technology*

3.3. Sample and research instruments

For this paper, a survey was conducted at the Technical Faculty Mihajlo Pupin in Zrenjanin in May 2021. 40 students, majoring in "Informatics, technique, and technology in education", were surveyed.

The research was carried out using a descriptive and experimental method. The chosen work technique is surveying. As an instrument for research on usage, there is a survey, which has 9 questions: the first 3 are related to the demographic structure of the respondents, and the other 6 are related to attitudes and opinions about the application of XBOX Kinect technology.

For the first three questions, students chose one of the offered answers, while for the remaining questions, the answers were given in the form of choosing the degree of agreement/disagreement with the given statement - Likert scale - (1 - "Absolutely disagree", 2 - "I disagree", 3 - "I have no opinion", 4 - "I agree", and 5 - "I absolutely agree").

3.4. Research results

The first question was related to the gender structure of the respondents. Based on Figure 1., it can be seen that the sample consisted mostly of male students, 57.5%, while 42.5% were female. Given that most men enroll in technical faculties, these results are expected.

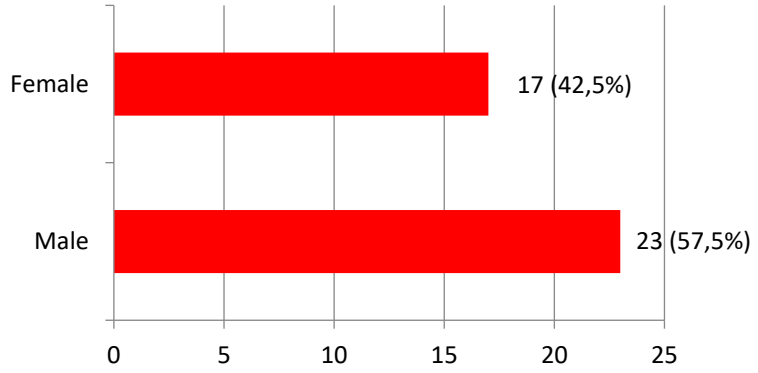


Figure 1. Gender structure of respondents

Figure number 2 shows the level of studies of the respondents and it can be seen that 50% of the respondents were in undergraduate studies, and the same number of respondents were in master's studies, while there were no respondents in doctoral studies.

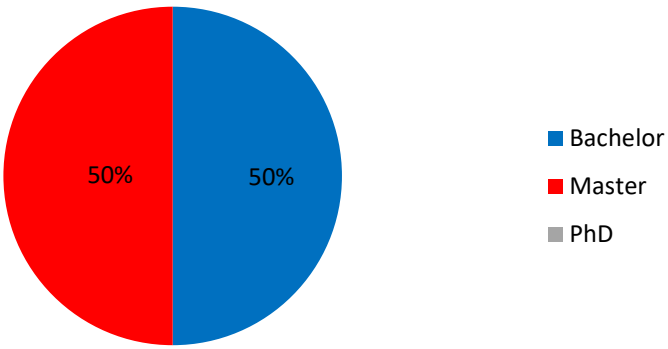


Figure 2. Studies of respondents

The following question was answered only by students of undergraduate studies - most respondents were in the third year of study (35.9%), then in the fourth year (33.3%), in the first year (20.5%), and in the second year (10 .3%).

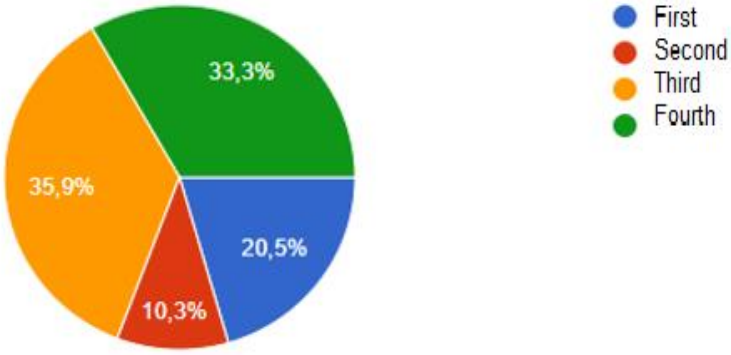


Figure 3. Years of study

Students had the opportunity to work with Kinect XBOX technology and were very satisfied. Almost all respondents believe that Kinect XBOX technology has helped them in learning (88% absolutely agree, 10% agree). However, there were also those students who did not share that opinion (disagree 2%), the reason for this is that they did not succeed in working with Kinect technology, which affected their acceptance of the newspaper.

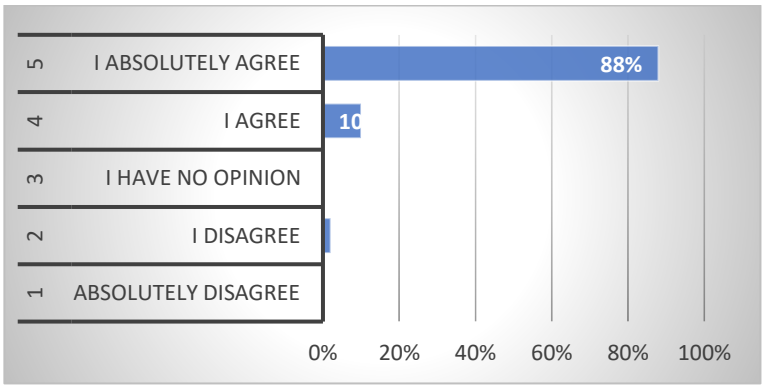


Figure 4. The application of XBOX Kinect technology helps students in learning

The students used the software package Avatar Kinect, with which it is possible to animate distance learning, collaboration on projects, and socialization between students. Most of the students were enthusiastic and willing to accept the new technology, which is why their experience with

Kinect is very positive. Based on personal experience, they believe that the application of XBOX Kinect technology allows teachers to better and more easily present and manage multimedia materials (76% absolutely agree and 4% agree). Students who did not manage to cope, have negative experiences - no attitude 10% and the same percentage of those who disagreed 10%.

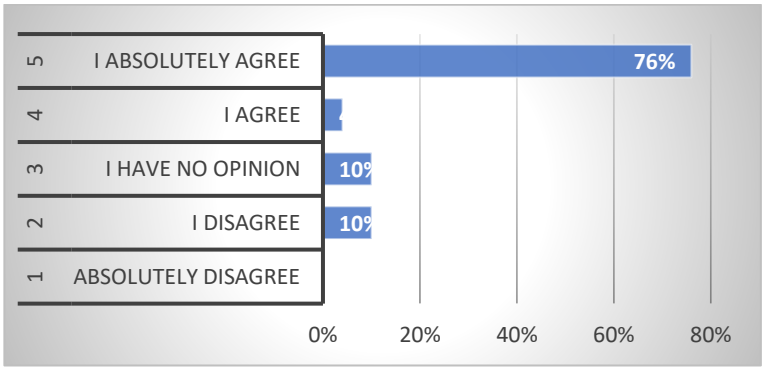


Figure 5. The application of XBOX Kinect technology enables professors to better and more easily present and manage multimedia materials.

The positive experiences gained by students in working with Kinect technology had an effect on increasing student motivation, and the answers obtained show that this is possible - 89% absolutely agree, while 10% of the respondents only agree.

There are no respondents who do not agree and who absolutely do not agree, while 1% do not have an opinion on this statement.

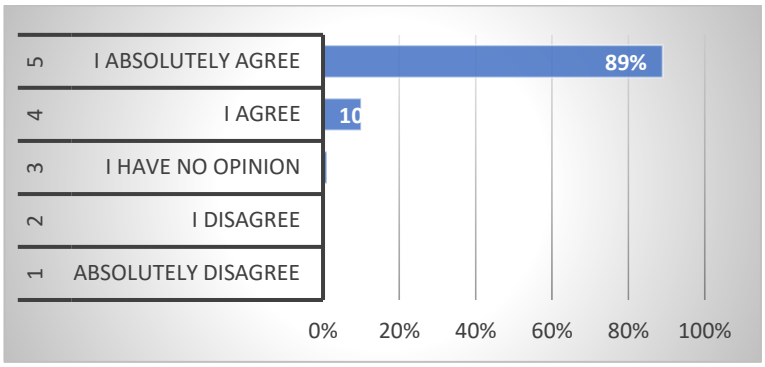


Figure 6. Experience with Kinect

Taking into account the positive experiences of working with Kinect technology, the students positively evaluated the class in which they had the opportunity to work with it and expressed great satisfaction in using Kinect - more than 70% of respondents were satisfied (grade 5 - 62.5% respondents and grade 4 - 15% of respondents).

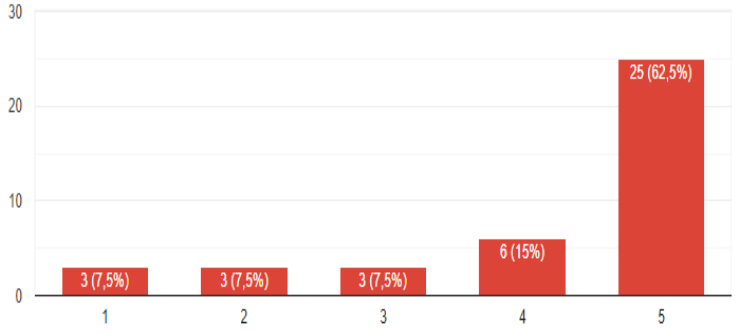


Figure 7. Assessment of work with Kinect

4. DISCUSSION AND CONCLUSION

The obtained results show that students are satisfied and motivated to use Kinect technology, which is consistent with some other research (5, 6). "KINECT motion sensing game-based learning environment offers a new sensory experience for the students to fully mobilize all senses to get involved in learning and interactions." (5)

This confirms the main hypothesis that the application of XBOX Kinect technology encourages students to be more engaged in classes and acquire knowledge more functionally through active learning.

In addition to mastering new material, students experienced different experiences in working with new technology in this way, which confirmed the first sub-hypothesis.

Given that both genders excelled in the application of Kinect technology, the second sub-hypothesis is rejected, because gender does not affect the ability to use XBOX Kinect technology.

Possible further directions of research are the examination of teachers' views on the use of Kinect, the possibility of acquiring software, as well as which skills can be improved by using Kinect.

Based on everything presented in the paper, it can be concluded that the use of Kinect devices in education is desirable and that it increases students' creativity and desire to work. Also, it can be concluded that human motion analysis in recent years has become a natural interface for HCI and has been the focus of recent researchers in modeling, analyzing, and recognizing gestures.

REFERENCES

[1] Karuović, D., Radosav, D., (2011). „Interakcija čovek-računar“, Univerzitet u Novom Sadu, Tehnički fakultet Mihajlo Pupin, Zrenjanin,

[2] https://www.techopedia.com/definition/10695/kinect

[3] I. Tashev., (2011). ‘‘Recent Advances in Human-Machine Interfaces for Gaming and Entertainment,‘‘ Int’l J. Information Technology and Security, vol. 3, no. 3.

[4] Damnjanović, Đ., Vujičić, D., Milošević, M., Jagodić, D.,(2016). „Određeni aspekti primene XBOX Kinect tehnologije u predmetu Interakcija čovek-računar“., Tehnika i informatika u obrazovanju, 6. međunarodna konferencija, Fakultet tehničkih nauka, Čačak.

[5] Y. H. Yang, W. Xu, H. Zhang, J. P. Zhang, M. L. Xu, (2014), The Application of KINECT Motion Sensing Technology in Game-Oriented Study, iJET, 9(2), 59-63, http://dx.doi.org/10.3991/ijet.v9i2.3282

[6] Oudah M, Al-Naji A, Chahl J., (2020), Hand Gestures for Elderly Care Using a Microsoft Kinect, Nano Biomedicine and Engineering, 12(3): 197-204. https://doi.org/10.5101/nbe.v12i3.p197-204

An Overview of Metric Models for Evaluating Website Security

Vuk Amizic[1], Ljubica Kazi[2]

[1]*University of Novi Sad / Technical Faculty "Mihajlo Pupin", Djure Djakovica BB, Zrenjanin, Serbia*
[2]*University of Novi Sad / Technical Faculty "Mihajlo Pupin", Djure Djakovica BB, Zrenjanin, Serbia*

vuk.amizic@tfzr.rs; ljubica.kazi@gmail.com

Abstract:
In this paper, different metric models for evaluating website security are analysed as well as professional tools and possible security threats that are commonly used by the attackers such as denial of service, distributed denial of service, SQL injection, DNS poisoning and similar. The detailed metric model for evaluation of the website security is constructed and proposed as a solution for measurement of overall website security score.

Keywords:
AIIT, security, security threats, security evaluation methods

1. Introduction

Metric models for evaluating website security are made to accurately show the grading of the overall website security and prevent possible informational system damage. Observing from two different aspects, security can be implemented from the client side, but evaluation must happen on the server side as well. Many techniques are used to prevent security breaches, depending on the place of occurrence.

SQL injection is one of the common methods used to breach the security and it poses a great threat for websites that have sensitive client data stored in the databases. The SQL injection itself can be prevented by using regular expressions and sanation of the query strings.

Observing the wrong implementation of the asynchronous javascript (AJAX), the threat can come directly by the client side, once the function that uses the AJAX is called. This can cause serious damage, due to javascript's ability to use "eval()" function, the function for evaluation of the given string.

Looking at the various possible threats, it is possible to determinate the level of it's importance and damage that it could pose to the system. Either it is the exposure of the operating system version, or exposure of the part of the source code, even the smallest detail can be a serious threat and a great advantage for the experienced attacker.

2. Related work

According to [1], DNS cache poisoning is the act of entering false information into a DNS cache, so that DNS queries return an incorrect response and users are directed to the wrong websites. DNS cache poisoning is also known as 'DNS spoofing.'

According to [1], DNS resolver caches are like a directory that lists these phone numbers, and when they store faulty information, traffic goes to the wrong places until the cached information is corrected.

According to [1], because there is typically no way for DNS resolvers to verify the data in their caches, incorrect DNS information remains in the cache until the time to live (TTL) expires, or until it is removed manually.

According to [1], attackers can poison DNS caches by impersonating DNS nameservers, making a request to a DNS resolver, and then forging the reply when the DNS resolver queries a nameserver. This is possible because DNS servers use UDP instead of TCP, and because currently there is no verification for DNS information.

According to [2], a metric generally implies a system of measurement based on quantifiable measures. **A method of measurement used to determine the unit of a quantity can involve a measuring instrument, reference material or measuring system.**

According to [2], security measurement and metrics efforts that are conceived at a high level of abstraction and formalism are often difficult to interpret and apply in practice. Existing formalisms also pose difficulties to reconcile with actual operational environments where software patches, version updates and configuration setting changes take place regularly.

According to [2], the absence of formal security models and other formalisms needed to improve the relevance of security metrics to deployed systems have hampered progress.

According to [3], until recently, researchers evaluated the efficacy of their defense proposals by testing their simulated traffic against the state-of-theart attacks.

According to [5], there are many different types of metrics that could be used for security evaluations:

Design - Threat Modelling

- number of threat models or threat modelling activities conducted
- number of findings from threat models

Develop - Code Reviews

- number of code reviews conducted
- number of findings from code reviews

Deploy - Security testing

- number of findings from penetration testing (internal vs vendor)
- number of findings from vulnerability scanning
- number of findings remediated
- time taken to remediate a vulnerability, does it meet an SLA (An organisation may require vulnerabilities to be remediated in a certain time frame in accordance with the vulnerabilities risk rating)

Security team collaboration

- number of secure by default modules created
- number of secure architecture designs created and provided to developers

Security champions

- number of security champions onboarded

Training activities

- number of training activities introduced to developers
- number of developers onboarded to training activities
- number of developers engaged in particular training activities

3. Security threats

There are various security threats that can seriously damage the integrity of the informational system. In order to determinate the exposure of the website security, it is required to perform metric evaluation based on the desired metric model for grading the website security.

Different types of security threats are:

- DNS cache poisoning – According to [1], refers to injecting incorrect data into the DNS cache, causing the DNS to return an incorrect response that will redirect the user to a malicious website. When processing the DNS cache, there is no possibility to verify the integrity of the data in the DNS cache, which is a major security problem

- DNS Tunnel – An attack using a DNS tunnel is based on the DNS protocol for communication through port 53. It functions according to the principle of sending HTTP and other protocol traffic through DNS. It is most often used when using VPN applications,

in order to hide the location if there is a code execution function on a remote server, by a malicious computer whose location should remain unknown or due to the bypassing of poorly implemented protection on flow limitation provided by Internet providers

- Man in the middle (MITM) – Is a type of attack that refers to the routing of data to a computer that is used as an intermediary between the client and the server. The most vulnerable data are those that are sent in the form of the HTTP protocol, if they are not encrypted, which allows the attacker to see the information that is sent from the client's computer to the server, and from the server to the client's computer. This attack can be achieved by setting up a malicious open Wi-Fi network for the user to connect to

- SQL injection - Refers to the entry of unwanted, modified SQL queries instead of the expected data, which, as a result of application, will enable the display, entry, change or deletion of records in the database, which will violate privacy, integrity and security

- Denial of Service (DoS) – The idea of this attack is based on the principle of flooding the information system, server or network, as a result of which the resources will be exhausted, and the service will not be able to provide feedback to the client who visits it

- Distributed denial of service (DDoS) – Distributed denial of service, similar to standard denial of service, is based on flooding information systems, servers and networks using multiple devices, which further accelerates the effect of resource depletion and achieves the effect of denial of service much more efficiently when requested by the client

There are various security elements that needs to be kept in mind before evaluating security of the website.

Figure 1: – According to [7], Security elements [4]

- Broken Access Control - refers to unregulated access to certain resources or execution of certain actions that are not intended for an ordinary user

- Cryptographic Failures - refers to the exposure of sensitive data from third-party entities, most often through the use of the HTTP protocol for the transfer of sensitive information and the use of outdated encryption algorithms

- Injection – refers to the injection of malicious queries and URL segments that can perform malicious actions

- Insecure Design – refers to poorly designed software, website or information system architecture

- Security Misconfiguration – refers to the establishment of a configuration that does not meet the maximum level of protection, or contains security flaws in the standard security settings

- Vulnerable and Outdated Components - refers to third-party entity libraries or development frameworks that contain security vulnerabilities or are no longer supported by the development team

- Identification and Authentication Failures - refers to failure in identification and authentication due to poor implementation of the source code and inability to protect against damage to the integrity of the source code

3. Methods for evaluating the security of websites

According to [2], security metrics are generally used to determine how many security aspects are covered in an application. Since security metrics vary from organization to organization, it is expected that the results of the metrics will vary from institution to institution for which the metric is performed.

According to [2], qualitative metrics can be applied to represent quantitative metrics of security aspects, where a low value represents the absence of security vulnerabilities, a medium value represents the existence of one to five security vulnerabilities, and a high value represents the existence of more than five security vulnerabilities.

According to [2], security metrics can be implemented by a formula that combines different types of application security ranking.

According to [2], the quantitative evaluation of several security aspects can be represented by the formula where Rating=0.25*rankingA + 0.75*rankingB

According to [3], to achieve the security metrics of website fingerprints, it is possible to apply the "Bayes Error Estimation" method, which is based on presenting the smallest possible error prediction for the data that is the target of the attack.

According to [3], another metric method for website security footprints is based on Information Leakage, which is the "WeFDE" technique for determining the amount of information leaked by poor security.

Table 1: Metric results for ML and DL feature experiments.

		Info. Leakage		Bayes Error		Top-1 Accuracy		Top-2 Accuracy	
		Bits	% of Max	$1-\hat{R}^*$	(ϵ,ϕ)-privacy	RF	DF	RF	DF
ML	Undefended	6.49	98.9%	90.9%	0.09	96.3%	-	97.9%	-
	WTF-PAD	6.54	99.6%	47.8%	0.52	62.5%	-	75.3%	-
	Walkie-Talkie	6.37	97.1%	45.9%	0.54	9.03%	-	89.5%	-
	Tamaraw	3.20	48.8%	28.5%	0.73	12.5%	-	21.4%	-
DL	Undefended	6.54	99.6%	97.9%	0.02	96.2%	97.1%	97.6%	98.2%
	WTF-PAD	6.48	97.8%	83.4%	0.17	81.2%	85.9%	88.5%	91.9%
	Walkie-Talkie	6.42	98.9%	72.7%	0.27	31.6%	43.8%	78.7%	98.1%
	Tamaraw	3.57	54.4%	20.3%	0.80	6.5%	7.6%	12.0%	13.2%

Figure 2: according to [3], Table of results for ML and DL experiment [1]

3.1 OWASP Top 10

OWASP, also known as The Open Worldwide Application Security Project is based on the open community where each person can participate and contribute to OWASP related projects in the goal of contributing to the community and making the improvements based on the application security.

According to [4], the OWASP Top 10 is a document for the awareness of web developers when developing web site security, with which it is possible to perform web site security metrics.

According to [5], OWASP Top 10 provided a way for website security metrics and graph representation. By measuring the security of websites, it is possible to evaluate the effectiveness and efficiency of the application's security program.

According to [5], it is possible to perform different types of security metrics such as:

- Design – threat modeling;
- Development – Code review;
- Putting into production - safety testing;
- Collaboration of security teams.

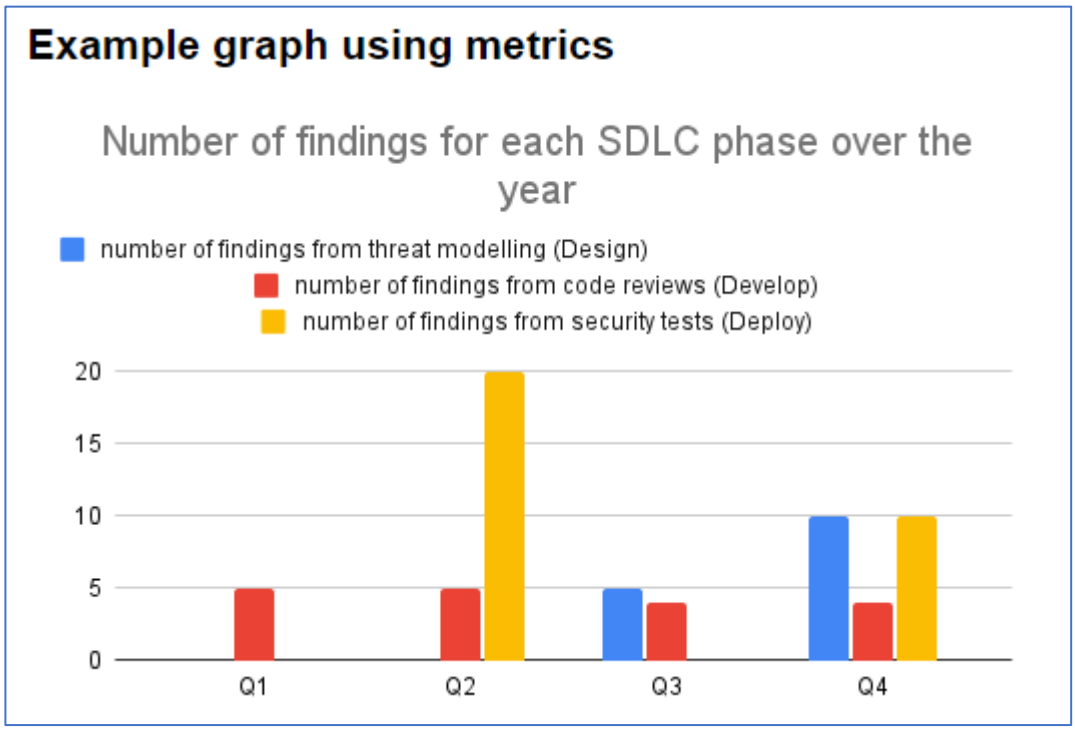


Figure 3: According to [5], Overview on metrics graph for security elements [2]

3.2 Immuni Web

According to [6], ImmuniWeb is an online tool for determining security flaws in websites. It is necessary to forward the link of the desired website for which the user wants to perform a security vulnerability analysis.

According to [6], based on security vulnerabilities, the security metrics of websites are rated A, B, C, F.

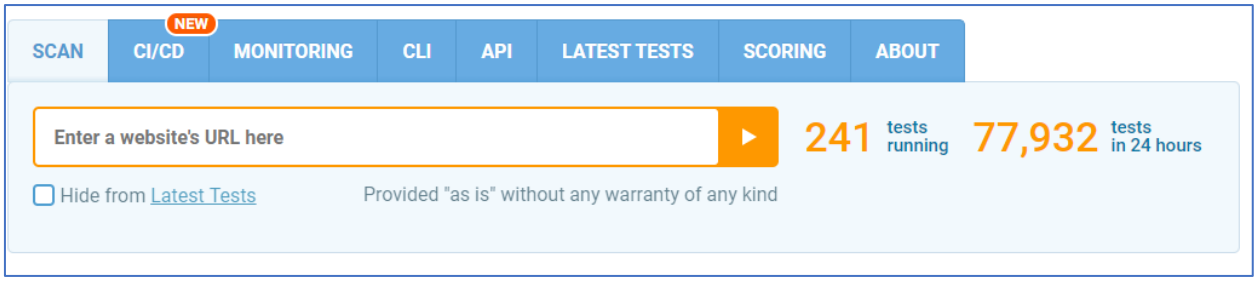


Figure 4: According to [6], Input form for URL of desired website, according to [3]

4. Proposed metric model for evaluating website security

In this paper, a metric model was developed for evaluating website security, which consists of a formula and three tables for calculating points for the existence of a feature, its importance, and security flaws, including a supporting system for scoring individual features.

Suggested formula for the total number of website security points:
BBB $= \sum$ BBkb (1) where BBB presents the total number of security points, while the $\sum$ BBkb reffers to the sum of the points of the security characteristic.

Suggested formula for the number of points of the security characteristic:
$BBkb = (Pk * V) + \sum$Bp (2) where $BBkb$ reffers to number of points of the security characteristic, while Pk represents the existence of a characteristic, V represents the importance of the security characteristic, $\sum$Bp represents the sum of the security flaws, and the actual Bp represents the security flaw.

***If there is no security flaw, it carries 0 points.**

According to [2], information security metrics represent an important decision factor in various aspects of security, from the design of security architecture and controls to the effectiveness and efficiency of security operations. Security metrics tend to offer a quantitative and objective basis for security assurance.

According to [2], security metrics can be defined for security processes, procedures and training during the development of system design, configuration and maintenance.

According to [2], there are factors that can form a table based on which web application security metrics can be done. Some of these factors are authentication, encryption, integrity, availability, reliability...

Based on the existence of the safety feature, table 1 is determined.

Table 1:
Table of existence of the safety feature (**Pk**)

Existence of characteristic	Number of points
Exist	1
Exist partially	0.5
Doesn't exist	0

Based on the security feature of importance metric (**V**) , table 2 is determined.

Table 2:
Table of security characteristic importance

Importance	Number of points
Very important	3
Important	2
Desired	1

Based on the security flaws, table 3 is determined.

Table 3:
Table of security flaw metrics (**Bp**)

Security flaw	Security threat severity	Number of points
Source code exposure	High severity	-2
Database exposure	High severity	-2
Operating system exposure	Middle severity	-1
Framework exposure	Middle severity	-1
Exposure of client information during session forming process	High severity	-2

• Source code exposure refers to the display of elements that reveal source code segments that threaten the security of the website by displaying methods and attributes that are processed when the website is used;

• Exposure of the operating system refers to the display of elements that reveal the server operating system and its version, which can cause damage to the entire information system;

• Database exposure refers to the vulnerability of a database through the use of malicious SQL code injection.

Term of the **"Exposure"** in this paper generally represents the broken integrity and security of information that should not be accesable by devices / clients outside of the administration of the Information Security teams.

5. Conclusion

Metric models for evaluating website security have important role in determinating the overall security of the websites and can provide accuarate security measurements based on various security models as well as elements, which can contribute to increasing the overall security of the websites and web based applications. Altough important, the actual researches of the security metrics and security evaluation are very hard to find due to lack of the experienced cyber security professionals, and continuous imrpovements of the frameworks, operating systems and technologies in general, which usually brings more flaws and security exposures such as are the Zero-Day exploits, which refer to the security breach elements not known to the developers and IT security teams.

One of the good ways of contributing to strenghten the overall website security is doing frequent security testing, and measuring the security level after each security update. Besides of the writing secure source code, it is required for the cyber-security experts to perform detailed analyisis of the website security in order for minimal risk of the possible security threat.

Proposed tables show the way that could improve aspects of security metric evaluations based on the proposed formulas for evaluating the overall security of the website.

Using the proposed formulas can give the score of the security metrics for any particular website, after which the security breach could be fixed, and score can be updated.

6. Further work

The table of the security flaw metrics can be upgraded to support much more security flaw occourances, as well as the development and expansion of the proposed formulas to cover security flaw metric technique in more details. Besides doing the manual testing, an automated scripts could be written for basic attacks such as SQL and URL injections, in order to speed up the overall security metrics time measurement. The notice of the successfull SQL or URL injection could be realised by the analysis of the website elements with libraries such as Selenium, with the help of XPath mechanics.

References:

[1] DNS cache poisoning URL: https://www.cloudflare.com/learning/dns/dns-cache-poisoning/
[2] Wayne Jansen, „Directions in Security Metrics Research“, National Institute of Standards and Technology, Diane Publishing, 2010.
[3] Mathews, Nate, Mohammad Saidur Rahman, and Matthew Wright. "Poster: Evaluating security metrics for website fingerprinting." Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security. 2019.
[4] OWASP Top 10 2021 Project URL: https://owasp.org/Top10/
[5] OWASP Top 10 website security metrics URL: https://owasp.org/www-project-security-culture/v10/8-Metrics/
[6] Online tools for website security measuring URL: https://www.immuniweb.com/websec/
[7] OWASP Top 10 2021 tool URL: https://www.owasptopten.org/the-release-of-the-owasp-top-10-2021

Pictures:

[1] https://dl.acm.org/doi/pdf/10.1145/3319535.3363272 6.7.2023.
[2] https://owasp.org/www-project-security-culture/v10/8-Metrics/6.7.2023.
[3] https://www.immuniweb.com/websec/ 6.7.2023.
[4] https://www.owasptopten.org/the-release-of-the-owasp-top-10-2021 6.7.2023.

13TH INTERNATIONAL CONFERENCE ON APPLIED INTERNET AND INFORMATION TECHNOLOGIES

Publisher

Faculty of Information and Communication Technologies – Bitola,

University "ST. KLIMENT OHRIDSKI" – Bitola

Printing house

3D Grafoprom – Bitola

100 copies