UNIVERSITY OF ZAGREB
FACULTY OF ORGANIZATION AND INFORMATICS
VARAŽDIN

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

Central European Conference on Information and Intelligent Systems

CECIIS

29th International Conference
2018

Organized by:
University of Zagreb
Faculty of Organization and Informatics

September 19th – 21st, 2018
Varaždin, Croatia

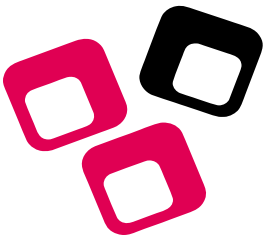
ISSN 1847-2001 (Print)
ISSN 1848-2295 (Online)

UDC 004:007

University of Zagreb
Faculty of Organization and Informatics

Central European Conference on Information and Intelligent Systems

CECIIS



29th International Conference

September 19th - 21st, 2018
Varaždin, Croatia

Varaždin, 2018.

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

Contact
e-mail: ceciis@foi.hr
phone: +385 42 390 861
fax: +385 42 213 413
web: www.ceciis.foi.hr
fb: www.facebook.com/ceciis.foi

CECIIS Proceedings have been indexed in the following databases: INSPEC, ProQuest (SciTech Premium Collection, ProQuest Technology Collection and ProQuest Adavanced Technologies & Aerospace Database).

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

Contents

Computer Games

Data and Knowledge Bases

Education and Learning Analytics

Emerging trends in ICT

ICT Entrepreneurship and Innovation

Intelligent Information Systems

Quality of Software and Services

Software Engineering

Strategic Planning and Decision Making

Index of Authors

List of Reviewers

Balogh, Aniko, Hungary
Banek Zorica, Mihaela, Croatia
Boban, Marija, Croatia
Ćosić, Jasmin, Bosnia and Herzegovina
Črepinšek, Matej, Slovenia
Crookall, David, France
Ćukušić, Maja, Croatia
Cvetkoska, Violeta, Macedonia
Delibašić, Boris, Serbia
De-Marcos, Luis, Spain
Detelj, Kristina, Croatia
Dmitrović, Veljko, Serbia
Dobša, Jasminka, Croatia
Faganel, Armand, Slovenia
Ferenc, Szani, Hungary
Galland, Stéphane, France
Gams, Matjaž, Slovenia
Guardia, Lourdes, Spain
Gutiérrez Martínez, José María, Spain
Hunjak, Tihomir, Croatia
Ivković, Nikola, Croatia
Jadrić, Mario, Croatia
Janeš, Aleksander, Serbia
Jereb, Eva, Slovenia
Kirinić, Valentina, Croatia
Kovačić, Bozidar, Croatia
Krašna, Marjan, Slovenia
Lang, Michael, Ireland
Lovrenčić, Alen, Croatia
Mekovec, Renata, Croatia
Meštrović, Ana, Croatia
Mladenić, Dunja, Slovenia
Okreša Đurić, Bogdan, Croatia
Orehovački, Tihomir, Croatia
Oreški, Dijana, Croatia
Palanca, Javier, Spain
Paralič, Ján, Slovakia
Pažur Aničić, Katarina, Croatia
Pejić-Bach, Mirjana, Croatia
Perišić, Branko, Serbia
Plantak Vukovac, Dijana, Croatia
Ristić, Sonja, Serbia
Rodič, Blaž, Slovenia
Schatten, Markus, Croatia
Schloegl, Christian, Austria
Seljan, Sanja, Croatia
Sobodić, Aleksandra, Croatia
Spremić, Mario, Croatia
Steingartner, William, Slovakia
Terna, Pietro, Italy
Tomičić Furjan, Martina, Croatia
Tomičić, Igor, Croatia
Umek, Lan, Slovenia
Virkus, Sirje, Estonia
Zekić-Sušac, Marijana, Croatia
Žgela, Mario, Croatia
Zlatović, Miran, Croatia

Preface

The Central European Conference on Information and Intelligent Systems (CECIIS) has been annually held since 1989. International participants, program committee and reviewers have promoted CECIIS to the most important conference in the area of applied information science in Central Europe. These are also prerequisites for referring the CECIIS Book of Proceedings to prominent databases. The Conference and its accompanying events are certainly the most significant scientific event organized by the Faculty of Organization and Informatics of the University of Zagreb. From the early beginnings to the present, the main aim of CECIIS is to foster the cooperation among researchers involved in the development and application of methods and techniques in the field of information and intelligent systems. The special topic of this year's conference is digital transformation, a comprehensive business transformation that seeks out all the advantages and opportunities of contemporary digital technologies and their impact on society.

The Proceedings of CECIIS 2018 include 28 scientific papers, each peer-reviewed by two independent reviewers and approved by the international Program Committee within the following sections: Computer Games, Data and Knowledge Bases, Education and Learning Analytics, Emerging Trends in ICT, ICT Entrepreneurship and Innovation, Intelligent Information Systems, Quality of Software and Services, Software Engineering, Strategic Planning and Decision Making. After the conference, the selected papers are further reviewed by two independent international reviewers with the aim to include the best papers in the Journal of Organizational and Information Sciences (JIOS), the international scientific journal published by the Faculty of Organization and Informatics.

The poster and presentation section for students organized within CECIIS gathered students from Croatia (Varaždin), USA (Cincinnati and Pittsburgh) and Albania (Shkodër). The poster abstracts are published separately on-line. The special value of this year's conference were the outstanding invited lectures held by two esteemed researchers:

- Dimitris Karagiannis (University of Vienna, Faculty of Computer Science, Vienna, Austria) - The impact of digitization on industry: a research perspective
- Dieter Hertweck (Reutlingen University, Faculty of Computer Science, Herman Hollerith Research Centre, Reutlingen, Germany) - The Digital Transformation in the manufacturing sector - Challenges and Chances for SMEs

Furthermore, several parallel events were held:

- 2nd Workshop on Data Analysis
- professional tracks: Digital Transformation of Educational Institutions
  Computer Games - Professional Development
- e-Schools Day - presentations and workshops of the national project "e-Schools: Establishing a System for Developing Digitally Mature Schools (pilot project)"
- project presentations - e-Schools, HigherDecision, CRISS, DIGITRANS and ISSES
- CECIIS 2018 conference sponsors' events.

We would like to express our kind gratitude to the invited speakers, authors, reviewers, session chairs, attendees, organizers of parallel events, as well as Program and Organizing Committee members for their contributions. We acknowledge that the organizational and supporting assistance and hosting of the Faculty of Organization and Informatics was crucially important for the overall success of the Conference. Finally, we are most grateful to all our business partners and sponsors for their support.

We believe that the following pages will provide you with relevant and interesting papers in the field of information and intelligent systems and we look forward to your participation in one of many CECIIS conferences to come.

September, 2018

On the behalf of the
Program and Organizing Committees

Vjeran Strahonja and Valentina Kirinić

Computer Games

How SMEs Can Use Games to Assess the Innovation Potential of New Technologies

Martin Kinitzki, Dieter Hertweck, Peter Kühfuß, Valeria Kinitzki
Reutlingen University
Faculty of Informatics
Alteburgstr. 150, 72762 Reutlingen, Germany
martin.kinitzki@reutlingen-university.de
dieter.hertweck@reutlingen-university.de
peter.kuehfuss@student.reutlingen-university.de
valeria.kinitzki@student.reutlingen-university.de

Abstract. *With the digital transformation being one of the most discussed topics in the business world today, many enterprises – especially small and medium sized ones – find themselves struggling with the understanding of new digital technologies and thus the potential benefits and risks for their companies. New technologies like the Internet of Things, Blockchain or Machine Learning have great potential for businesses. However, carefully evaluating and selecting purposeful technologies – aligned to the digital strategy – is the key to success. Technologies appear, change and also vanish so rapidly in the digital age, that a proper understanding is crucial for a sustainable technological foundation. Focusing on the characteristic features of technologies, the presented approach promises to create a better technological understanding for decision makers in small and medium-sized enterprises (SMEs) in a playful manner: With a serious game that fosters insight and allays fears of digitalization.*

Keywords. digital innovation, technologies, digital skills, serious game, SMEs

1 Introduction

The increasing digitalization driven by paradigm shifts like the Industry 4.0 or the Internet of Things puts companies under great pressure and forces them to develop new business models to remain competitive on global markets (Briken 2015; Schallmo et al. 2017; Wolan 2013). The literature states that the digital transformation should be considered the basis of a technical and economical paradigm shift influencing economy and society as a whole (Blättel-Mink and Menez 2015; Picot et al. 2017). Being innovative and transforming into the digital age is often reduced to the implementation of new technologies. But there is no digital transformation without appropriate business transformation and the consequences of ignoring or escaping the trend become very clear with examples like Kodak, where the company was unable to make a transition into digital thinking, causing its own downfall due to the disruptive character of digital photography. But it is not the global players or startups that struggle the most, especially decision makers in SMEs suffer from the digital revolution as they still lack competences to cope with the challenges arising from the digital transformation and find themselves in a completely different situation compared to big companies who employ technology experts or startups that found their digital business models on top of digital natives' competencies.

Proper instruments on how to cope with the digitalization for SMEs are provided by the European transnational research project Digitrans. The project's main objective is to enable SMEs to create competitive and sustainable innovative business models. It offers a transformation method divided into an innovation phase that is linked to an innovation space design and an organizational implementation phase supported by an online platform that provides training content and self-assessment tools. The method is provided in innovation trainings and workshops to more than 300 SMEs.

After the first trainings, three main deviations from transformation processes in global companies and investor financed startups can be stated: (a) different level of technological awareness of decision makers before they enter business model development procedures; (b) hierarchical B2B customer value chains with little to no customer integration opportunities into the innovation process; and (c) the need of cost benefit calculations and feasibility analysis in early stages of the business model development process. These characteristics derive from the fact that – compared to startups – SMEs do not start their innovations on a green field but have to cope with change management and the actual transformation of existing business models. Unlike large global companies SMEs are often lacking dedicated human and monetary resources for trial and

error approaches, not to mention the dedicated personnel like Chief Digital Officers (CDOs) that global companies employ. For most of the SMEs that are not in the position of bringing new know-how into the company by hiring digital experts or integrating startups into their structures, it is up to the company owners or CEOs to acquire the crucial knowledge of digital strategies and technologies.

That is why the goal set for the innovation method development was to make the drivers of digital transformation more comprehensive to the owners and managers of SMEs. A feasible innovation method has to support them to discover innovation potentials for their businesses by raising their technological awareness and allaying their fears of becoming the "digital prey" by means of a serious game concept. Serious games have been tested in the past as an interesting alternative to bring business knowledge to decision makers in SMEs (PLAYITIL, 2018).

Working on the nature of serious games to educate decision makers, the chosen research approach consists of two methodological dimensions: (a) Gaining basic features of digital technologies by decomposing them to an atomic level; and (b) Designing a serious game to foster technology understanding as well as identify innovation potentials for the digital transformation of SMEs.

Therefore, the focus was set on answering the following research questions: (a) What influence does technology have on the transformation process of SMEs?; (b) How can a serious game be used to raise the technological awareness of SME decision makers?; and (c) Is it instructive to decompose technologies to their characteristic features for the purpose of knowledge transfer?

In preparation of this study, a systematic literature review was conducted to examine the influence of technologies and technology understanding on digital transformation processes.

2 Methodology of Literature Review

The underlying method of literature research as shown in Figure 1 was adapted from Webster and Watson (2002). Literature was captured from (a) GoogleScholar; (b) Elsevier and (c) SpringerLink since these databases provided the most relevant results during an initial explorative search phase. Search and analysis were conducted in February 2018 using the following search keywords: *SME*, *serious*, *game*, *technologies*, *innovation*, *digital*, *skills*, *learning, disruptive, disruption, impact, influence*.

To set the focus on the latest findings, only literature published since 2005 was considered, when the definition of serious games was established by Michael Zyda in the IEEE Computer journal. Whenever possible, filter criteria restricting results to scientific papers were applied. Due to the large amount of results, the search had to be narrowed down considering titles only.

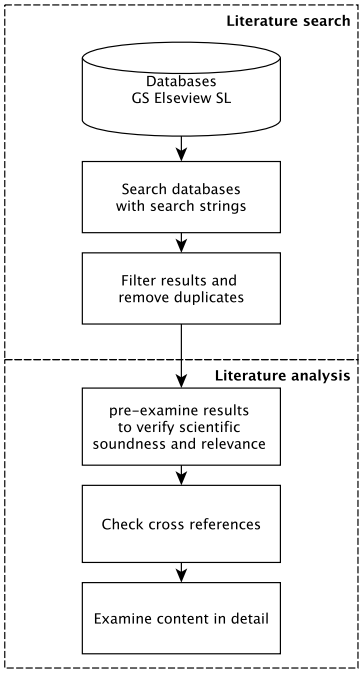


Figure 1: Process of literature research

After removing duplicates from the respective search results, a first content-related evaluation based on the documents' titles and abstracts was carried out. Results without any relevance to the object of this research had to be excluded from further analysis, as well as results which lacked scientific standards. The remaining results were analyzed in full detail. Finally, the bibliographies of highly relevant results were examined to determine further literature contributing to answer the raised research questions. During this process, 132 relevant results were identified.

3 Influence of Technologies and Technology Understanding on Digital Transformation

As early as 1982, Curran and Mitchell described our fast-changing world and its impact on managers: They are obliged to foresee the impact of technological developments on their companies and identify opportunities and threats in time. The authors describe "a basic understanding of technology's far-reaching scope" as crucial to managers in order keep up with the transformation of the business world. (Curran and Mitchell 1982) Since then, the pace of change has only increased further and the perception of information technology shifted from being purely supportive to being an enabler of new business models (Châlons and Dufft 2017). Despite the realization that the digital transformation does not spare anyone, most companies

still struggle to grasp the benefits from technologies due to the lack of proper understanding (Fitzgerald et al. 2013). Naskali et al. (2018) highlight the significance of this condition for SMEs, where functions are less segregated and decision makers are often left alone with the digitalization as they usually lack digital experts compared to global companies or startups. This leads to the requirement of a different transformation approach for SMEs.

The existing digital transformation process models found in literature assume a certain degree of digital maturity and expertise. Schallmo and Williams (2018) for instance, present a roadmap for the development of digital business models. It synthesizes existing transformation approaches and consists of five phases: (1) Digital Reality, (2) Digital Ambition, (3) Digital Potential, (4) Digital Fit and (5) Digital Implementation. None of these phases covers the aforementioned potential lack of technology understanding as seen in SMEs. For that reason, SMEs require a different entry point into the transformation process that begins with imparting knowledge about technology as an important driver of the digitalization. The transformation model presented in this paper (Figure 2) is a work in progress. It is codeveloped by empirical experiences of 7 universities in the Danube Region. As part of the Digitrans Project, it addresses this condition by providing a technology entry point into the transformation process. It is solely dedicated to fostering technology understanding of SME decision makers.

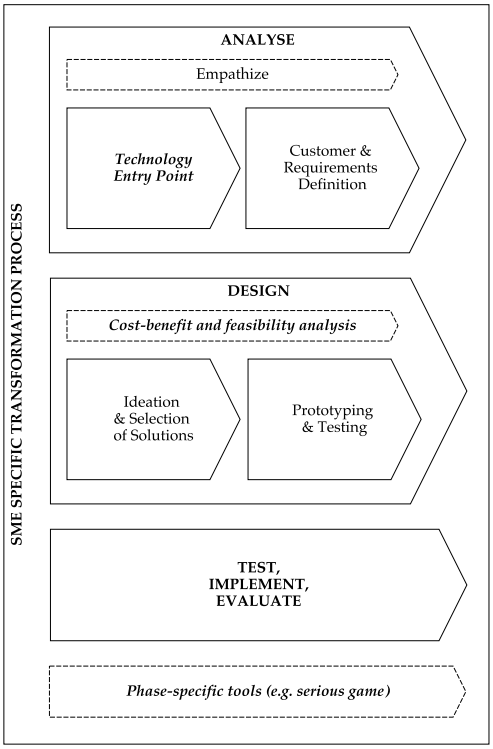


Figure 2: SME specific transformation process

Technology understanding is prerequisite for a successful transformation. The chosen serious game approach delivers the advantages of learning by playing and has the potential to tackle the existing knowledge gap.

4 Gaining Basic Technological Features by Decomposing Digital Technologies

Taking a look at the digital businesses of the last decade and their disruptive influence on economic and social levels, they all have one thing in common: their business model is mainly based on digital technologies. Whether it is Amazon or eBay that have ousted classic mail-order businesses, WhatsApp replacing SMS or Airbnb and Uber with their impact on the traditional hotel and taxi industry (Châlons and Dufft 2017). The reason for their success was the versatility of digital technologies next to lean processes due to the lack of organizational and system legacy (Fleisch et al. 2014).

In their study about Service Systems Engineering, Böhmann et al. (2014) point out the importance of technologies in cyber-physical systems – which can easily be transferred to the context of this study. The progress resulting from the connection of human and technological resources enables new forms of business models, products and services capable of developing disruptive characteristics that also affect the society as a whole (Böhmann et al. 2014; Kempf 2013). Social revolutions like crowdfunding and the current form of shareconomy would not have been possible without the technologies and know-how available at that time. It becomes clear that the digital transformation is far more than just "transferring analog information to digital medium" (Hamidian and Kraijo 2013).

Researchers and scientific literature covered in the conducted literature research concur that technology has an immense influence on life and business and missing the chance to innovate, might decide whether a company will survive or not. The discovery and further understanding of the impact that diverse technologies have on businesses and society marks an interesting object of research. Such studies usually take place over a period of several years in which success stories are analyzed in detail and similarities are determined. An excellent example is the St. Gallen Business Model Navigator based on a five-year study in which 250 business models were analyzed and patterns derived (Gassmann et al. 2013).

Unlike business model types, the amount of digital revolutions or technology induced paradigm shifts is much smaller and attempting to detect innovation patterns on a higher level has proven to be less expedient. Therefore, the attempt on breaking through the technologies' surfaces by dismantling them to their atomic artifacts was chosen. This marks the top-down approach of the proposed research: Collecting

technologies, structuring, cataloging and decomposing them. The foundation of this heavily knowledge-driven concept is the catalog of technologies dismantled into their essential features. Identifying these characteristic features is considered the main research objective of the approach proposed in the research study. Understanding the "nature" of a specific technology on an atomic level might be fruitful in two ways: Improve the understanding by focusing on the main features and hiding the technical frame; and determining the innovation potential for different business contexts.

An example for the decomposition is given with the blockchain technology. It describes an algorithm and distributed data structures for secure data transfers without a central administration (e.g. electronic cash transfers). It was originally designed for the crypto-currency Bitcoin and the blockchain concept was mainly driven by the rejection movement against money and bank-controlled payments regulated by governments. The original vision of the Bitcoin developers was to enable people to spend money without friction, intermediaries, regulations or the necessity to know or trust third parties (Lewis 2015; Nofer et al. 2017).

Examples for essential features that can be derived are the following:

- Secure transactions
- Trusted transactions
- Decentralized transactions
- Frictionless transactions

With these characteristics in mind – completely detached from the technology itself – one can start thinking about how businesses might enrich their business models, products or processes or even create new ones using these features. It becomes quite evident that a technology like blockchain can be transferred to other fields as its capabilities are entirely independent of its original use case – Bitcoin. From a technological perspective, blockchain allows for decentralized, secure and trusted data transactions that can be used in completely different scenarios like voting systems, signatures and legal proofs of existence or possession. Complemented by the serious game that will be used on innovation workshops with SMEs across Europe, not only an expedient way to transfer knowledge to decision makers is expected. But also interesting results examining the impact of these essential features on the organizational, cultural and technological levels of SMEs. Figure 3 depicts this proposed research approach with the combination of desk and empirical research.

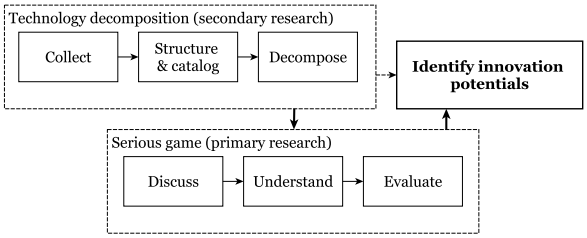


Figure 3: Proposed research approach

Besides the empirical study of intrinsic innovation potentials of technologies, the serious game concept aims at raising the players' understanding of technologies and the relation to business contexts in a playful manner. It is also considered as a proof-of-concept instrument aiming at answering the research question whether decomposing technologies is beneficial for the knowledge transfer in the first place.

5 Learning Innovation by Combination of Technological Features and Business Context - The Serious Game Approach

Playing is closely associated to learning processes and therefore paves the way to achieve learning success more easily with the acquired knowledge being transferable onto other contexts. This is the idea of serious games. (Kerres et al. 2009) A serious game refers to a game that follows traditional game mechanics but is designed to fit a different purpose than pure entertainment, like education or exploration (Breuer and Bente 2010; Djaouti et al. 2011; Göbel et al. 2014).

In terms of digital education, there is a certain demarcation between serious gaming and concepts like gamification, E-learning and game-based learning. Gamification deals with game mechanics in non-gaming situations in order to enhance motivation or influence a certain behavior (Groh 2012). E-Learning (or educational technology) refers to the utilization of technologies in education to improve the learning process and its results without the integration of playful aspects (Dichanz and Ernst 2001). Game-based learning uses games for motivation to achieve learning effects and reduce performance pressure on learners by concealing the educational aspects (Pivec et al. 2003).

The presented serious game concept is meant to educate the users about technologies by hiding the technical frame and focusing on their characteristic features. The reduction of complexity does not only prevent the players from being overwhelmed but it also allows for the validation of the aforementioned decomposition process by empirically discovering innovation potentials and problem solutions based on features rather than technological buzz terms.

As mentioned before, there is no digital transformation without appropriate business transformation and therefore technologies should always be set in relation to business contexts. The main motivation of the round-based game lies in fostering discussions about technologies and their potential as innovation drivers in these contexts. By bringing together people from different positions and departments – each with certain skills, experience and knowledge – multi-perspective debates on the subject can be conducted. It should be noted, that the tracking of resulting technological solutions with respect to a specific context is an important key aspect of the concept, as it poses the empirical part of discovering innovation potentials by contemplating technologies' essential features.

The structure of the serious card game is fairly simple as it consists of three card types only: *Technology Cards*, *Context Cards* and *Joker Cards*.

Each technology card contains a certain technology or technological concept and its set of essential features identified in the previously mentioned decomposition phase of the study.

Context cards represent specific challenges that are to be solved through the use of technologies. They mark the impulse on the discussion of technologies. Besides the predefined contexts that are meant to demonstrate the serious game approach, organization specific challenges and problems can be tackled and added to the library as well.

Joker cards represent a special kind of cards, that can be played instead of ordinary technology cards. By playing a joker card, the player can choose whether to skip the current round, take another card from the stack or replace it with a technology of choice. Joker cards are supposed to increase the dynamics of the game.

There are two roles intended in the game: a moderator and the players. The moderator creates the game by selecting from predefined sets of contexts and technologies or defining new ones, moderating the discussion and accepting or rejecting solutions. The players simply join the game, play their cards, debate and vote for the best solutions.

At the beginning, the moderator puts a context card on the table and each player receives a defined amount of technology cards. In turn, every player plays a technology card or set, justifying why it serves (as part of) a potential solution. Throughout each round, the technologies are discussed, and the best technology or combination of technologies is chosen and checked in by the moderator. If there is no suitable solution found, the context is rejected, and the next round starts. The players' card stacks are refilled and every player whose card was a part of the solution, is rewarded with game points. The solutions checked in by the moderator are stored in the system.

The game structure is based on a console-like architecture that is easy to set up with no additional hardware. The console itself is represented by an ordinary web browser and the gamepads through smart phones or tablets. With the web browser illustrating the game board managed by a moderator, the players can use their devices to interact on the board.

The underlying system architecture follows the proven three-tier architecture, strictly separating the presentation, application logic and data storage as shown in Fig. 3. With a web service as the central element of the system, the object-oriented communication between the clients and the server is established and the computation is running in the cloud, so that neither the players nor the moderator have to cope with the application's system environment. Being accessible in web browsers across all devices renders it fairly easy to play the game without any installation or configuration needed.

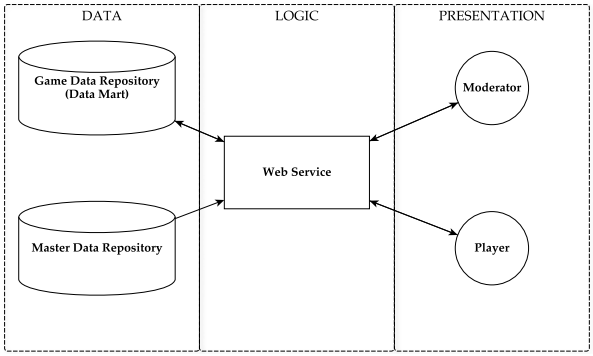


Figure 4: Game architecture

The several components of the system have to provide or enable the following main functions:

- Digital game structure
- Game moderation dashboard
- Access control
- Tracking solutions
- Recording new contexts and technologies
- Technology lookup

The presentation layer contains all views that are presented to the moderator as well as the players. This includes the game table, cards, dashboards and all information shown. It is stacked on the other two layers and provides the visuals for the underlying logic and data storage components. The communication is bidirectional, so that through the web services, operations in form of requests are received and computed and responses are provided by rendering them in the presentation layer. An important role of this layer lies in the presentation of information about technologies in order to foster the understanding by briefly explaining it to the player. This can be implemented in form of texts, images or videos showing sample scenarios of the technologies in use. Without doubt, storing all relevant data is utterly important and therefore a mature database concept is inevitable, but the representation of the data in a form that is understandable and reliable is just as important. To facilitate the accessibility, each technology card is

equipped with a question mark symbol that provides the demanded information on-click.

The web services are core of the logic layer which computes all requests sent by the moderator and players. All requests as well as the responses are visualized by the presentation layer in form of user interfaces. The required data is fetched from the databases and new data is stored in it. In the data layer, two separate databases are envisaged: One containing the master data, like predefined contexts and technologies and the other for the game instances with the data required per game.

For the future prospects, a recommender system providing the most relevant (combinations of) technologies to similar contexts is desirable – derived from the growing collection of empirical data.

6 Conclusion

With all the new opportunities and challenges emerging from the digital transformation, ignoring it puts organizations at risk of falling behind the competition. However, properly evaluating and selecting the right technologies as drivers of the transformation presents great challenges to businesses. Especially small and medium-sized enterprises find themselves struggling due to the lack of financial and human resources limiting the possibility of trial-and-error approaches on being innovative. Decision makers are often left alone with the digitalization as result of these conditions. It is inevitable to understand that the digital transformation is not going to stop, and companies will find themselves on an everlasting journey with the need to adapt, rethink and reshape constantly. A proper technological understanding and awareness is crucial to remain competitive.

The proposed research study tackles that condition by offering a way on making digital technologies more comprehensive. It helps SMEs to discover innovations for their businesses and solutions to existing problems by raising their technological awareness with a serious game. By exploring technologies and helping SME decision makers to grasp the benefits of technology in a playful manner, the prospects of this research study appear very bright.

Acknowledgement

The conducted research is part of the European transnational research project Digitrans which focuses on the digital transformation of SMEs in the Danube region. This project is part of the Interreg Danube Transnational Programme and is co-funded by the European Union funds (ERDF).

References

Baum, G. 2013. “Innovationen Als Basis Der Nächsten Industrierevolution,” in Industrie 4.0, U. Sendler (ed.), Heidelberg: Xpert.press. Springer Vieweg, pp. 37–53.

Blättel-Mink, B., and Menez, R. 2015. “Technischer Wandel Und Innovation,” in Kompendium Der Innovationsforschung, pp. 99–110. (https://doi.org/10.1007/978-3-531-19971-9_6).

Böhmann, T., Marco Leimeister, J., Möslein, K., Leimeister, J., and Möslein, K. 2014. “Service Systems Engineering,” Business & Information Systems Engineering (6:2), pp. 73–79. (https://doi.org/10.1007/s12599-014-0314-8).

Boughzala, I., and Michel, H. 2016. “Introduction to the Serious Games, Gamification and Innovation Minitrack,” Proceedings of the Annual Hawaii International Conference on System Sciences (2016–March), p. 817. (https://doi.org/10.1109/HICSS.2016.105).

Breuer, J., and Bente, G. 2010. “Why so Serious? On the Relation of Serious Games and Learning,” Journal for Computer Game Culture (Vol. 4).

Briken, K. 2015. “Gesellschaftliche (Be-)Deutung von Innovation,” in Kompendium Der Innovationsforschung, pp. 21–31. (https://doi.org/10.1007/978-3-531-19971-9_2).

Châlons, C., and Dufft, N. 2017. “The Role of IT as an Enabler of Digital Transformation,” in The Drivers of Digital Transformation, F. Abolhassan (ed.), Springer International Publishing, pp. 13–22.

Curran, S., and Mitchell, H. 1982. “New Technology: Understanding the Impact,” in Office Automation, London: Palgrave Macmillan UK, pp. 20–39. (https://doi.org/10.1007/978-1-349-05975-1_3).

Dichanz, H., and Ernst, A. 2001. “E-Learning.”

Djaouti, D., Alvarez, J., and Jessel, J.-P. 2011. “Classifying Serious Games: The G/P/S Model.”

Fitzgerald, M., Kruschwitz, N., Bonnet, D., and Welch, M. 2013. “Embracing Digital Technology: A New Strategic Imperative,” MITSloan Management Review, pp. 1–12. (https://doi.org/10.1057/palgrave.ejis.3000650).

Fleisch, E., Legner, C., and Thiesse, F. 2014. “Informationstechnologie-Basierte Geschäftsmodelle – Stand Und Ausblick,” in Wirtschaftsinformatik in Wissenschaft Und Praxis, pp. 103–114. (https://doi.org/10.1007/978-3-642-54411-8_7).

Gassmann, O., Frankenberger, K., and Csik, M. 2013. The St . Gallen Business Model Navigator, pp. 1–18.

Göbel, S., Mehm, F., Wendel, V., Konert, J., Hardy, S., Reuter, C., Gutjahr, M., and Dutz, T. 2014. Erstellung, Steuerung Und Evaluation von Serious Games. (https://doi.org/10.1007/s00287-014-0824-2).

Groh, F. 2012. "Gamification: State of the Art Definition and Utilization," Research Trends in Media Informatics, pp. 39–46.

Hamidian, K., and Kraijo, C. 2013. "DigITalisierung – Status Quo," in Digitalisierung Und Innovation, pp. 1–23. (https://doi.org/10.1007/978-3-658-00371-5_1).

Kempf, D. 2013. "Die Rolle von ITK Bei Industrie 4.0," in German Conference on Information Systems (WI2013), Leipzig, Germany.

Kerres, M., Bormann, M., and Vervenne, M. 2009. Didaktische Konzeption von Serious Games.

Lewis, A. 2015. "Blockchain Technology Explained," Blockchain Technologies, pp. 1–27. (https://doi.org/10.15358/0935-0381-2015-4-5-222).

Mouaheb, H., Fahli, A., Moussetad, M., and Eljamali, S. 2012. "The Serious Game: What Educational Benefits?," Procedia - Social and Behavioral Sciences (46), pp. 5502–5508. (https://doi.org/10.1016/j.sbspro.2012.06.465).

Naskali, J., Kaukola, J., and Matintupa, J. 2018. Well-Being in the Information Society. Fighting Inequalities, (Vol. 907), Springer International Publishing. (https://doi.org/10.1007/978-3-319-97931-1).

Nofer, M., Gomber, P., Hinz, O., and Schiereck, D. 2017. "Blockchain," Business & Information Systems Engineering (59:3), pp. 183–187. (https://doi.org/10.1007/s12599-017-0467-3).

Picot, A., Hopf, S., and Sedlmeir, J. 2017. "Digitalisierung Als Herausforderung Für Die Industrie – Das Beispiel Der Automotive Branche," in Technologie, Strategie Und Organisation, pp. 87–112. (https://doi.org/10.1007/978-3-658-16042-5_5).

Pivec, M., Dziabenko, O., and Schinnerl, I. 2003. "Aspects of Game-Based Learning."

Schallmo, D. R. A., and Williams, C. A. 2018. Digital Transformation Now! Guiding the Successful Digitalization of Your Business Model, Springer International Publishing. (https://doi.org/10.1007/978-3-319-72844-5).

Schallmo, D., Rusnjak, A., Abzebgruber, J., Werani, T., and Jünger, M. 2017. "Digitale Transformation von Geschäftsmodellen: Grundlagen, Instrumente Und Best Practices," Jetzt Digital Transformieren. (https://doi.org/10.1007/978-3-658-14569-9).

Webster, J., and Watson, R. 2002. "Analyzing the Past to Prepare for the Future: Writing a Review," Management Information Systems Quarterly (26:2), pp. 13–23.

Winkelhake, U. 2018. "Information Technology as an Enabler of Digitisation," The Digital Transformation of the Automotive Industry. (https://doi.org/10.1007/978-3-319-71610-7_8).

Wolan, M. 2013. Digitale Innovation: Schneller. Wirtschaftlicher. Nachhaltiger, BusinessVillage.

Zhang, Q., Cheng, L., and Boutaba, R. 2010. "Cloud Computing: State-of-the-Art and Research Challenges," Journal of Internet Services and Applications (1:1), pp. 7–18. (https://doi.org/10.1007/s13174-010-0007-6).

Zyda, M. 2005. "From Visual Simulation to Virtual Reality to Games," Computer. (https://doi.org/10.1109/MC.2005.297).

SERVIEW PLAYITIL (PLAYITIL) (2018). Retrieved from https://www.serview.de/training/different/gamification-spiele/

Towards an Application Programming Interface for Automated Testing of Artificial Intelligence Agents in Massively Multi-Player On-Line Role-Playing Games

Markus Schatten, Bogdan Okreša Đurić, Igor Tomičić
Artificial Intelligence Laboratory
Faculty áf árganization ánd ánformatics, University of Zagreb
Pavlinska á, á2000 áaraždin, ároatia
{markus.schatten, dokresa, igor.tomicic}@foi.hr

Abstract. *An initial implementation of an application programming interface (API) for automated testing of artificial intelligence agents in massively multi-player on-line role-playing games (MMORPGs) is presented and analyzed. The API, which is based on results from the Large-Scale Multi-Agent Modelling of Massively Multi-Player On-Line Role-Playing Games (ModelMMORPG) project, is implemented in Smart Python Agent Development Environment (SPADE) and allows for the implementation of artificial agents which are able to play The Mana World (TMW), an open source MMORPG. The API provides access to basic player behaviours including but not limited to movement, fight, pick-up of items and interaction with non-player characters (NPCs) as well as more advanced features like interaction with other players through chat, creation of player parties and trade. The basic proposition of the API is to allow concurrent testing of various types of artificial agents in playing the game which allows for agent behaviour analysis as well as gameplay analysis.*

Keywords. application programming interface, automated testing, artificial intelligence, massively multi-player on-line role-playing game

1 Introduction

The gaming industry is recognized as a rapidly growing domain spanning various intended uses, from entertainment and education, to programming and design industries, to serious games and research. The importance and success of the gaming world is emphasized by the Electronic Sports (eSports) movement as well. As one of the identified application domains of large-scale multiagent systems (LSMASs), along with the Internet of Things (IoT), smart cities, smart grid, complex systems and smart transportation (Schatten, Tomičić, and Đurić, 2017), to name a few, the gaming industry gained a lot of attention in recent research trends, especially when MMORPGs are considered. Other than being a good test-bed for technologies such as virtual reality and augmented reality, the gaming domain succeeded in becoming an effective environment for social studies as well.

Computer games have been a source of training data for various approaches to development of a form of an artificial intelligence (AI) using different heuristics as well as machine learning methods. These forms of specialized AI were trained and developed for use with computer games that are of a simple point-and-click, first person shooter, or simple arcade game genre. Only recently has the gaming industry turned towards developing a form of a general AI, since games can be modified quickly, with many re-runs of experiments. Such a short response and customization time support the processes of learning when artificial agents are considered. In this context, Peng et al. have presented multiagent bidirectionally-coordinated network "with a vectorised extension of actor-critic formulation", which could "handle different types of combats under diverse terrains with arbitrary numbers of AI agents for both sides" (Peng et al., 2017). Authors used StarCraft game as a test-bed scenario and demonstrated that their network facilitates learning of various types of agent coordination strategies. A literature overview on AI techniques used in real-time strategy games was presented in (Robertson and Watson, 2014), with a focus on a StarCraft game. Authors have identified the main areas of current research: tactical and strategic decision making, plan recognition, and learning. Another application on StarCraft game was presented in (Synnaeve and Bessiere, 2016), where authors showed the use of Bayesian models in three distinct core components: micro-management/units control, tactics, and strategy, arguing the possibilities of probabilistic models on sources of uncertainty and incompleteness, which are inherent to the AIs in real-time strategy games.

This paper therefore presents initial steps towards implementation of an API for automated testing of AI agents in MMORPGs. Since testing an MMORPG is a tedious task that requires countless hours of game-

play, having it performed by an artificial agent greatly increases relevance of the received testing results, and the sole amount of such testing results.

The rest of this paper is organized as follows: firstly in section 2 we provide an overview of related work. Then in section 3 we show the architecture of the implemented API and in section 4 we discuss what additional steps have to be taken for the API to be complete. In the end in section 5 we draw our conclusions and provide guidelines for future work.

2 Related Work

The majority of related research is performed on simple games that usually use only two dimensions, and are linear in gaming style - either it's simply moving, collecting items, and avoiding obstacles, or almost only avoiding obstacles and moving in a single direction, or something similar.

The major distinctive difference between the published research and the research proposed in this paper, is the game genre that is put in the focus. While related research is focused primarily on games that are linear in the nature of their objectives, the API of this research is being built for an MMORPG which is a part of the role-playing games (RPGs) genre that usually utilizes non-linear story that is a combination of various quests and side-quests thus introducing a level of creativity in gameplay.

The topic of APIs for RPGs has received an increase in popularity when DeepMind and Blizzard (Vinyals et al., 2017) opened Blizzard's StarCraft 2 game as an AI research environment. The published set of tools designed for AI research combines the API developed by Blizzard, and a modified open source version of DeepMind's toolset that can be used to integrate Python-based agents with Blizzard's API. The combination of the two, along with the additional elements of the set of tools, can freely be used by researchers and developers in their AI-related pursuits. StarCraft 2 is a sequel and a modernized version of an old but highly popular and valued real-time strategy (RTS) computer game that has secured its popularity as a long-time champion in the domain of eSports. The cited research is an example and proof of the growing importance of computer games in research and their importance in AI development.

Another research on the topic of how RPGs can be used to further and enhance AI research is presented in (Tian et al., 2017), where a research-oriented platform is presented that conforms to a defined set of features of an ideal game environment for fundamental reinforced learning – one that is diverse in the context of properties of offered games, efficient in running simulations, and provides highly customized environment settings. RTSs are used as a way of providing richer environments for reinforced learning, ensuring that training using e.g. raw pixel data is supported, since not many RPGs exist that can be used for research directly (Tian et al., 2017). The research presented by Tian et al. (2017) therefore provides a platform for enhanced use of reinforced learning, in the context of computer games of RTS genre, which is significantly different than RPG games that are the focus of this paper's research.

Another take on the concept of machine learning, using visual input from a computer game for training the AI, is associated with a game of first-person shooter (FPS) genre (Kempka et al., 2017). *ViZDoom* is purposed to be used for developing bots that play Doom, the classical FPS computer game released in 1997. The game is played based on visual reinforced learning principles, with the software emulating actions corresponding to key presses and mouse movements used to control the gameflow, and providing the developer with some directly available in-game data. Doom, the computer game used in this context, is far more simple than an average modern RPG since the provided and requested gameplay is much simpler and does not require devised planning.

Classic machine learning approaches have been used in development of artificial agents that can play so-called platformer games, such as Super Mario World, Super Mario Kart, or Pacman. While all of these examples are interesting to observe from the point of view of using computer games in training an AI agent, they do not represent APIs by themselves.

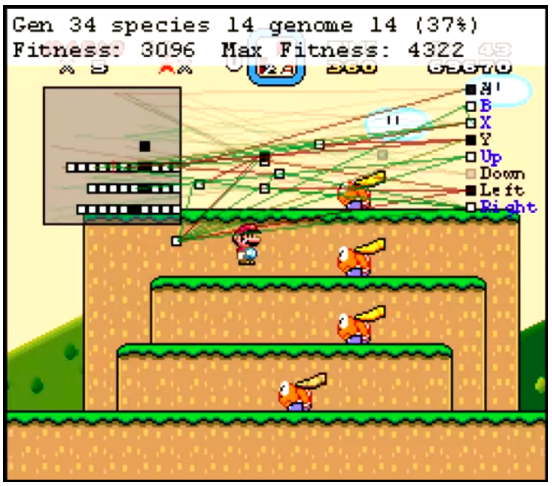


Figure 1: MarI/O agent training, a screenshot from YouTube

MarI/O[1], shown on figure 1, is a piece of software developed by an Internet user known as SethBling that uses neural networks and generic algorithms to build an agent that is capable of successfully playing a simple Super Mario World game. The software interacts with the game using emulated keystrokes, just like seen in ViZDoom above. The use of visual input coupled with

[1]For further information, visit: https://www.youtube.com/watch?v=qv6UVOQ0F44, and https://pastebin.com/ZZmSNaHX

the use of recurrent neural networks is shown by the same author on another example, working with a driving computer game named Super Mario Kart – MariFlow[2]. As opposed to the earlier described example, this time the AI does not learn by genetic modifications, but is fed input data in form of sample games. The last such example presented here, for the sake of emphasizing the use of computer games in information communication technology (ICT) research, is another showcase of neural network methods being used to train and play a game, this time of Pacman[3].

3 Implementation

The importance and added value of the research laid out in this paper comes from the chosen application domain – MMORPGs. The difference of such a genre, when compared to some of the examples presented in Section 2, is obvious for those accustomed to computer games.

FPS is the most simple genre, when strategizing is considered, as it is based on destroying as many opponents as possible or demanded, and may be the simplest to implement, since the agent only needs to identify its target, and shoot at its most vulnerable point.

The genre of RTS is complex in terms of strategic thinking, but is mostly based on a single goal, packaged into successfully outrunning the opponent in the context of, usually, economic or militaristic prowess and power. While the possible ways of achieving victory are diverse and mostly depend on the nature of the given player and the way they want to interact with the game (e.g. emphasizing technological advancement, or building their armies in numbers), the goals are seldom more complex than those derived from the aforementioned context.

Finally, an RPG boasts a much more complex world in general, since many more decisive elements are present, more so than in either FPS or RTS genre. Furthermore, role-playing games, being based on developing a player's character (often referred to as an avatar), are built around a story that helps develop this avatar, most often defined by their various traits, features, abilities, inventory, etc. The in-game storyline is not as linear as that of a typical FPS or a typical RTS either. Although these two genre may provide a storyline to be followed during the game in extent, the gaming experience is heavily divided by levels that act as a whole. An RPG is more often modeled as an open-world game, allowing the players to roam the in-game world and choose objectives for themselves, while the game in general is directed by a set of quests that demand interaction with concepts of the in-game world. Thus the possible set of actions and their utilization is much larger than that of either of the aforementioned computer game genre.

In addition to computations in terms of devising plans on how to solve a quest, the context of MMORPGs demands solving social-level problems and challenges as well, since socialization is an important and integral part of the *massively multi-player* part of this genre. More so, socialization is very important in utilizing all the features of a MMORPG. Therefore, building an API for an MMORPG is not only about enabling an agent to play a game and interact with the in-game world, but to interact with other players as well, which is an element that is not present in research presented in Section 2.

During the ModelMMORPG project, which aimed on developing artificial agent organizations that are able to play MMORPGs, a byproduct was the development of an interface to a very particular open source MMORPG called TMW. The interface was implemented in Python and comprises two parts: (1) a lower-level interface – dealing with particular low-level communication with the game server; as well as a (2) higher-level interface – dealing with the actual implementation of a playing agent (Schatten, Tomičić, Đurić, and Ivković, 2017; Schatten, Đurić, Tomičić, et al., 2017; Schatten, Đurić, Tomičič, et al., 2017).[4] The architecture of the ModelMMORPG agent implementation system is shown on figure 2, on which the MMORPG plug-in is the interface to TMW. Detailed model of both lower- and higher level interfaces is shown in figure 3.

The lower-level interface, shown on figure 4 implements the network protocol of the TMW servers and emulates an actual client connecting to the server and playing the game. It also implements a number of functions that allow actual gameplay including navigation (moving the avatar including automated navigation like moving to some character or dropped item, following a player or mob etc.), attack, sitting down, standing up, whispering communication to other players, picking and dropping up items, equipping items, creating parties of players (including inviting, responding to invitations, leaving parties and communicating on the group chat), environment sensing (what is my location, what is the location of nearby players/items/NPCs), and trading of items with other players, interacting with NPCs.

The higher-level interface builds upon the lower-level by introducing an agent template implemented in SPADE (Gregori et al., 2006) and adding additional parts like a STRIPS-based planning system implemented in Prolog, as well as an agent knowledge base implemented using the SPADE knowledge-base system for SWI Prolog. The higher-level interface acts as a layer around the low-level interface by providing higher-level agent behaviours based on low-level func-

[2]For further information, visit https://www.youtube.com/watch?v=Ipi40cb_RsI

[3]For further information, visit: https://www.youtube.com/watch?v=t5--kLRI4UE

[4]The current version of the interface is available at https://github.com/tomicic/ModelMMORPG

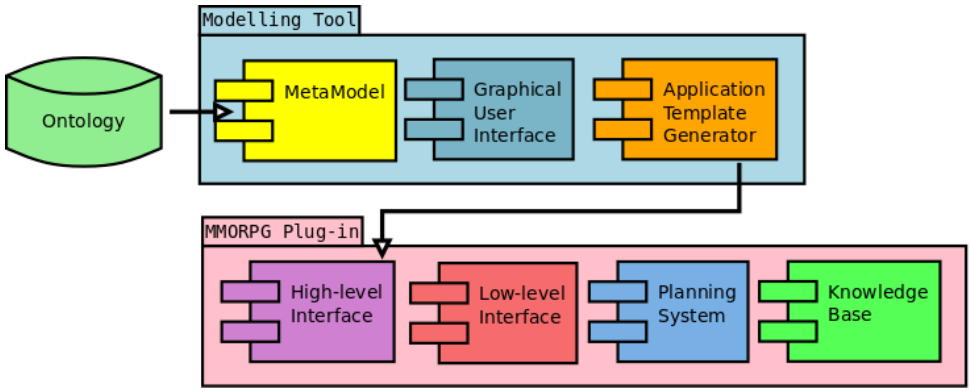


Figure 2: Architecture of the ModelMMORPG agent implementation system (Schatten, Đurić, Tomičić, et al., 2017)

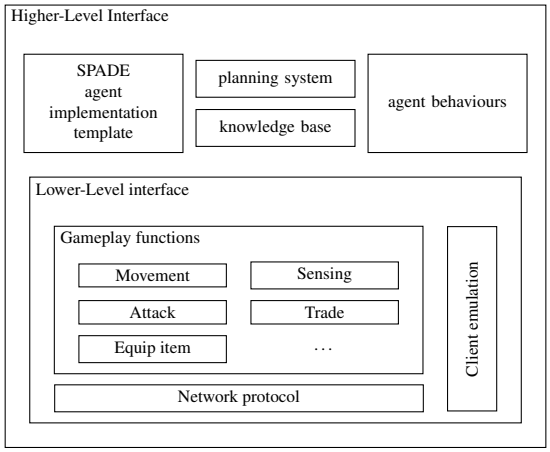


Figure 3: Detailed model of lower- and higher-level interfaces and their featuers

tion including but not limited to: navigation (including random walks), NPC conversation handling, fight handling, and party management. It is a typical belief-desire-intention (BDI) agent that firstly senses the environment, updates it's knowledge base, chooses an objective to accomplish next, generates a plan for this particular objective and then starts executing it, while the objective and the plan still make sense.

4 Discussion

In its current state, the interface cannot be really called an API, since a number of core functionality is missing, most prominently adequate documentation on how to implement an agent able to play TMW including a few simple examples as well as the abstraction of core functions from the actual implementation of an agent to be reusable. Currently, the implementation features a finished agent prototype that is able to play some of the initial quests of the game, as well as create, join or leave parties based on its assigned role. This prototype has to be dissected into parts that can be reused for new types of agents, possibly using different AI, reasoning and reaction technology like machine learning for example, since the prototype is based on BDI and STRIPS.

In order to do so, some basic rewriting of the code related to the high-level interface has to be done, for example: all prototype related code that was used for large-scale experiments has to be moved into an example usage, while core functionality like random walks, creation of parties, conversation with NPCs and similar have to be retained for future use. On the other hand, the low-level interface can stay more or less intact, since it is already used as a module for the high-level interface and acts as a low-level API.

Also, the planning system as well as the knowledge base of the agent can be easily reused for future agents' implementation, but the interface towards them has to be made more accessible for new types of agents.

Even though the API is to be used for testing MMORPGs in the context of their logical, quest, and story soundness, and balanced gameplay, the basic use case of the work presented and proposed in this paper is to develop testing scenarios whose details depend on the developer and testing needs. Therefore the success of testing largely depends on the game developers and testers, and the implementation details beyond the scope of this research which aims to provide functions and agent behaviours necessary for successful interaction between agents and an MMORPG.

5 Conclusion

Herein we have presented an initial step towards the implementation of an API for the automated testing of AI agents for MMORPGs based on some byproducts of the ModelMMORPG project. We have implemented a lower-level interface, a higher-level interface, a STRIPS based automated planning system as well as a knowledge base for the future API that will allow for the implementation of AI agents for the open source MMORPG TMW.

The most important contribution of the API is its fo-

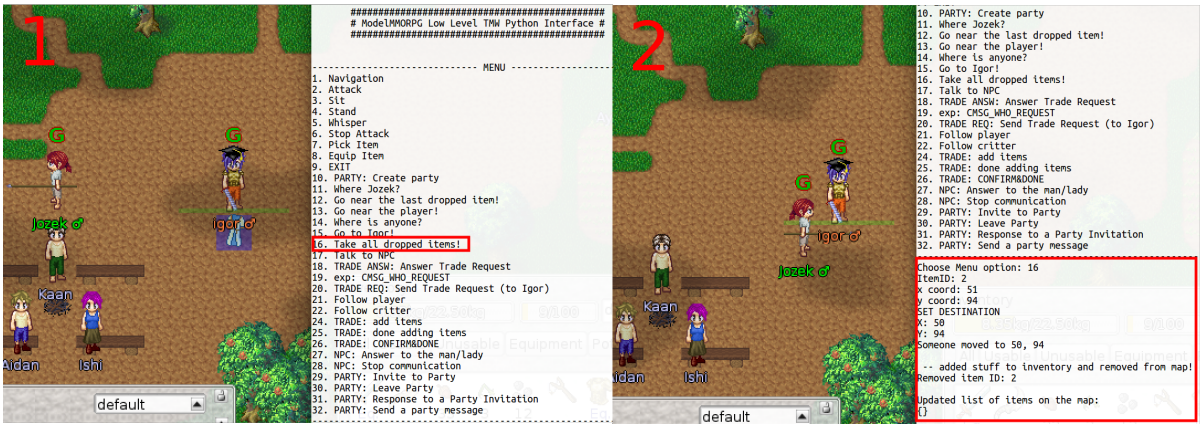


Figure 4: Lower-level interface to TMW (Schatten, Đurić, Tomičič, et al., 2017)

cus on a new genre, as opposed to current approaches. The RPG genre, and more recently the MMORPG genre hasn't been part of recent research for automated testing of AI agents, most probably due to the complexity of the domain, since such games include lots of complex behaviour of players to be exhibited in order to advance in the game.

Our future research will therefore be aimed towards completing the API by dissecting the prototype code into well established parts that can be reused in future AI agent implementation projects, by developing a number of simple AI agents for developers to start with as well as more profound documentation.

Acknowledgments

This work has been fully supported by the Croatian Science Foundation under the project number 8537.

References

Gregori, M. E., Cámara, J. P., & Bada, G. A. (2006). A jabber-based multi-agent system platform. In *Proceedings of the fifth international joint conference on autonomous agents and multiagent systems* (pp. 1282–1284). ACM.

Kempka, M., Wydmuch, M., Runc, G., Toczek, J., & Jaskowski, W. (2017). ViZDoom: A Doom-based AI research platform for visual reinforcement learning. *IEEE Conference on Computatonal Intelligence and Games, CIG*. doi:10.1109/CIG.2016.7860433. arXiv: 1605.02097

Peng, P., Yuan, Q., Wen, Y., Yang, Y., Tang, Z., Long, H., & Wang, J. (2017). Multiagent bidirectionally-coordinated nets for learning to play starcraft combat games. *arXiv preprint arXiv:1703.10069*.

Robertson, G. & Watson, I. (2014). A review of real-time strategy game ai. *AI Magazine*, *35*(4), 75–104.

Schatten, M., Đurić, B. O., Tomičić, I., & Ivković, N. (2017). Agents as bots–an initial attempt towards model-driven mmorpg gameplay. In *International conference on practical applications of agents and multi-agent systems* (pp. 246–258). Springer.

Schatten, M., Đurić, B. O., Tomičič, I., & Ivkovič, N. (2017). Automated mmorpg testing–an agent-based approach. In *International conference on practical applications of agents and multi-agent systems* (pp. 359–363). Springer.

Schatten, M., Tomičić, I., & Đurić, B. O. (2017). A review on application domains of large-scale multiagent systems. In T. Hunjak, V. Kirinić, & M. Konecki (Eds.), *Central european conference on information and intelligent systems* (pp. 201–206).

Schatten, M., Tomičić, I., Đurić, B. O., & Ivković, N. (2017). Towards an agent-based automated testing environment for massively multi-player role playing games. In *Information and communication technology, electronics and microelectronics (mipro), 2017 40th international convention on* (pp. 1149–1154). IEEE.

Synnaeve, G. & Bessiere, P. (2016, December). Multiscale Bayesian Modeling for RTS Games: An Application to StarCraft AI. *IEEE Transactions on Computational Intelligence and AI in Games*, *8*(4), 338–350. doi:10 . 1109 / TCIAIG . 2015 . 2487743

Tian, Y., Gong, Q., Shang, W., Wu, Y., & Zitnick, C. L. (2017). ELF: An Extensive, Lightweight and Flexible Research Platform for Real-time Strategy Games. (Nips), 1–14. arXiv: 1707.01067

Vinyals, O., Ewalds, T., Bartunov, S., Georgiev, P., Vezhnevets, A. S., Yeo, M., … Tsing, R. (2017). *StarCraft II: A New Challenge for Reinforcement Learning*.

Implementing Agent Roles in Massivley Multi-Player On-Line Role-Playing Games

Igor Tomičić, Bogdan Okreša Đurić, Markus Schatten
Artificial Intelligence Laboratory
Faculty of Organization and Informatics, University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
{dokresa, igor.tomicic, markus.schatten}@foi.hr

Abstract. *Organizational roles in multiagent systems (MASs) are an important concept in modelling and implementation of complex interactive systems like massively multi-player on-line role-playing games (MMORPGs). The paper presents a novel approach to implementing such roles as sets of agent behaviours. An initial implementation in Smart Python Agent Development Environment (SPADE) is presented and applied to the implementation of artificial players that are able to enact various roles in The Mana World (TMW) an open source MMORPG. The expressivity and applicability of the presented approach is discussed and examples of usage are provided.*

Keywords. organizational roles, multi-agent systems, massively multi-player on-line role-playing games, artificial player implementation, agent behaviour

1 Introduction

Agent roles are a novel implementation technique in the development of artificial agents as well as MASs and more recently large-scale multiagent systems (LSMASs). Herein we will introduce the usage of this implementation technique in MMORPGs since they provide an important application domain for LSMASs (Schatten, Tomičić, et al., 2017).

The most often used tool for modelling artificial game characters like non-player character (NPC) include finite state machines (FSMs), where each state relates to the state of the character and defines its choice from available actions. For a more complex behaviour, FSMs is inherently insufficient, and thus more more flexible planning capabilities are required. There are other techniques that are used for modelling NPC and player behaviour like behavioural trees. Behavioural trees, as opposed to FSMs that model states, model actual behaviours and include sequences, probability and priority selectors as well as decorators (Yannakakis and Togelius, 2017). Another ad-hoc method is using a utility function similarly to fuzzy logic in which the decision about current action depends on a vector of variables which is evaluated against a fuzzy set function (Yannakakis and Togelius, 2017).

Stanford Research Institute Problem Solver (STRIPS) and goal oriented action planning – based for example on belief-desire-intention (BDI) – are used for planning techniques in games such as F.E.A.R. (Orkin, 2006), and lately in research of MMORPGs such as (Schatten, Đurić, et al., 2017).

When modelling more complex behaviours, all these methods can become quite bulky and cumbersome, since a complex character might encompass many different behaviours and even more states to act in a game environments such as MMORPGs. Herein we would like to outline agent roles as a possible solution for the implementation of such complex sets of behaviours as well as provide examples by using SPADE (Gregori et al., 2006).

The rest of this paper is organized as follows: firstly in section 2 we provide an overview of related research. Then in section 3 we show how agent roles can be implemented as sets of behaviours in SPADE and give an example role implementation in 4. In the end in section 5 we draw our conclusions and give guidelines for future research.

2 Related Work

Agent roles are a concept introduced from organization theory (see (Schatten, Ševa, et al., 2016) for an in-depth review) that allow for the implementation of agents that are able to enact a given role based on the role specification. There have been a few propositions to use such agent roles in games in related literature.

For example, in (Westra, F. Dignum, et al., 2008) the authors propose to view games explicitly as organizations designed and developed for achieving certain goals and requirements. Within the organization, individual agents use appropriate behaviours in order to reach external goals. The agent organization defines constraints and capabilities of organizational concepts: roles, tasks, interaction protocols, and parties. Within their paper, authors also propose a system for agents that are adapting to the user during gameplay, but con-

sidering that the game does not reach unwanted states, e.g. breaking the game's story line, which might happen should the game contain randomly adapting agents. The proposed agents are adapting individually, but the adaptation is guided using an agent organization. The OperA (M. Dignum, 2004) framework which is used within the paper, distinguishes organizational aspects from the individual ones, enabling the specification of organizational requirements and objectives, but at the same time allowing individuals to act according to their own demands and capabilities.

In (Westra, Van Hasselt, et al., 2008) the authors contemplate about adaptability within on-line adapting games. They highlight the limitations of using centralized control in dynamic adjustability, and to meet the rising complexity and the number of adaptable elements, authors suggest the use of an multi-agent approach, specifically for *"adapting serious games to the skill level of the trainee"*. Considering the research on "reconciling" the two main aspects in games – the flexibility of adaptability, and the control of a game's story line, authors use the idea of agent organizations as a means for mediation. The authors also use the OperA model for agent organizations, which *"enables the specification of organizational requirements and objectives, and at the same time allows participants to have the freedom to act according to their own capabilities and demands."* Figure 1 shows a simple example of an OperA interaction structure through scenes (depicted with squares), which can progress in parallel.

Agent organizations are also considered in (Huber and Hadley, 1997), where authors describe the architecture and performance of autonomous agents that are able to play Netrek, a complex, multi-player, multi-team, real-time internet game. Authors recognize the challenge in creating agents which are able to play the game, not only autonomously as an individual, but also by cooperating and coordinating with other members of the same team, and coordinating against members of the opposing team. There are several agent roles that authors are suggesting, each tied to specific actions they are required to perform. The roles are named "engage", "assault", "escort", "ogg", "protect", "get armies", and vary from the relatively simple behaviours (like attacking the closest opponents) to the more complex ones (bombing a planet, dropping armies on an opponents planet, etc.). In the conclusion, author argue that their agents were able to pursue complex goals "within a very complex, dynamic environment", with roles as the key enablers for task solving processes.

Merging agent technology with game technology (game engines) is not a trivial task, as argued by (F. Dignum et al., 2009); for this consolidation to work, agents should run in their separate threads and only loosely be coupled with the game engine. The synchronization between the two proved to be an important aspect within this context; the communication between agents, and between agents and the game world should be enabled, but also a means of translation between the agent and the gaming world, as the agents operate on a more abstract level. Authors argue that these challenges can be faced by using *"agent technology to its full extent"*. Also, authors argue that *"using a conceptual stance allows for connecting the agent concepts to the game concepts such that agent actions can be connected to actions that can be executed through the game engine and that agents can reason intelligently on the information available from the game engine."* These connections could be implemented through agent roles.

The method based on MAS architecture which would be used for defining a game is described in (Aranda et al., 2012), and the first phase authors are remarking is the role definition phase. The roles are herein specified with names and sets of attributes, with attributes defining properties bound to game-playing agents. Also, agents have some predefined set of roles in order to provide basic features that any massively multi-player on-line game (MMOG) should have by default, and these basic features can be extended by the game designers in the form of new features within the game. Hierarchy is used to relate roles, and agent organizations are used to represent agent behaviour related to forms of collaboration, such as player clans.

3 Implementing Agent Roles

We will use agent roles as sets of agent behaviours. According to (Marian et al., 2004) agent behaviours can be:

- **role factory** (a role added/deleted at runtime to be enacted/stopped by the agent);
- **itinerary** (allows mobile agents to travel across various locations and perform tasks);
- **periodic** (looped behavior possibly with a given period of time intervals between iterations);
- **observer** (an agents awaits an event in order to perform its actions);
- **listener** (a special type of observer in which an agent awaits a special message of some other agent);
- **client/server** (resembles the client-server model);
- **one-shot behavior or task** (represents a simple task or activity);
- **finite state machine** (resembles a finite state machine in which every node is an activity to be performed);
- **sequential behavior** (a sequence of other behaviors);
- **parallel** (various behaviors are run in parallel).

In addition to these types of behaviour, we would like to add an additional one, which from our practice, has shown to be very useful for various practical implementations where an agent is using a resource concurrently: **exclusive behaviour** – allows an agent to run the behaviour exclusively, by stopping all other

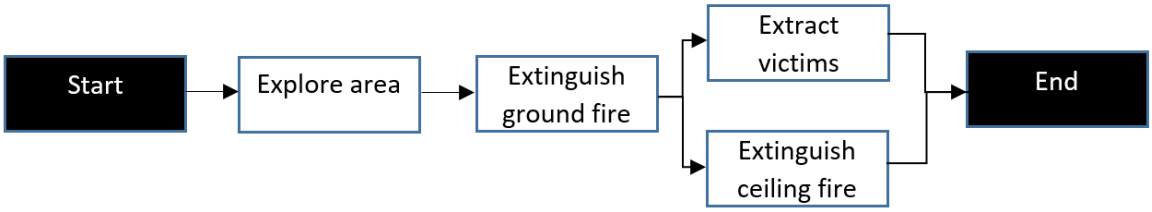


Figure 1: A simple example of an OperA interaction structure (Westra, Van Hasselt, et al., 2008)

exclusive behaviours. Such a behaviour, usually implemented with locking or semaphores, allows the agent to use some kind of resource (like a database, a file, a network socket etc.) that isn't thread-safe, by putting all other concurrent behaviours, that might use the same resource, at hold.

As per definition, we can define an agent's role as a set of behaviours $\mathbb{R} = \{b_1, b_2, \ldots, b_n\}$ which are added to the agents behaviours at run-time at the exact moment the agents acquires the role and starts enacting the defined behaviours. In the same manner, these behaviours are removed in the moment the agent stops enacting the role.

Translated into Python and more precisely SPADE the implementation of an agents role is straightforward as shown in listing 1.

Listing 1: An agent role class in Python/SPADE

```
class Role:
    ''' An (organizational) role is
        basically a set of behaviours.
    The behaviours should be a list of
        elements having the form: (
        behaviour instance, template
        instance ).
    The second item only applies to
        EventBehaviours (template),
        else it is None'''
    def __init__( self, behaviours=[] )
        :
        self.behaviours = behaviours
```

Thus, an instance of a role is independent of actual agents. It is just a container to hold various behaviour instances possibly with templates used for event behaviours (which are the implementation of observer behaviours for messages in SPADE).

In addition to a role class, we define two behaviours which implement the role factory behaviour which allows agents to acquire or stop enacting a role. The first behaviour allows an agent to acquire a role and is shown in listing 2.

Listing 2: An agent behavior class for adding a role implemented in Python/SPADE

```
class AddRole( spade.Behaviour.
    OneShotBehaviour ):
    """Behaviour to add a Role to the
        Agent. The Agent will acquire
        behaviours of the given Role."""
    def __init__( self, role, *args, **
        kwargs ):
        spade.Behaviour.OneShotBehaviour.
            __init__( self, *args, **kwargs
            )
        self.role = role

    def _process( self ):
        if not hasattr( self.myAgent, "
            roles" ):
            self.myAgent.roles = []
        self.myAgent.roles.append( self.
            role )
        for behaviour, template in self.
            role.behaviours:
            self.myAgent.addBehaviour(
                behaviour, template )
```

Basically, the behaviour appends the role to the list of roles the agent already has (or initializes the list of roles to an empty list first if the agent has no roles), and then adds all behaviours from the role to the agent to start using them.

The second behaviour removes a role from the agent, and is shown in listing 3.

Listing 3: An agent behavior class for deleting a role implemented in Python/SPADE

```
class DeleteRole( spade.Behaviour.
    OneShotBehaviour ):
    """Delete a role of the Agent. The
        Agent will lose all behaviours
        of the given Role."""
    def __init__( self, role, *args, **
        kwargs ):
        spade.Behaviour.OneShotBehaviour.
            __init__( self, *args, **kwargs
            )
        self.role = role

    def _process( self ):
        if not self.role in self.myAgent.
            roles:
            raise ValueError, "The agent isn'
                t playing the role to be
```

```
        deleted!"
    for behaviour, _t in self.role.
        behaviours:
      self.myAgent.removeBehaviour(
          behaviour)
    self.myAgent.roles.remove(self.
        role)
```

The behaviour basically checks if the agent has the role to be removed, and if yes, removes all behaviours of the role, and then removes the actual role.

4 Example

In order to provide additional clarification we provide an example of how such an implementation of agent roles could be used and has been used in the Large-Scale Multi-Agent Modelling of Massively Multi-Player On-Line Role-Playing Games (ModelMMORPG) project. In listing 4 a *Leader* role is implemented.[1]

Listing 4: An example Leader role implemented in Python/SPADE

```
class Leader(Role):
  class LeaderBehaviour(spade.
      Behaviour.OneShotBehaviour):
    def _process(self):
      (...)

  class PartyStats(spade.Behaviour.
      OneShotBehaviour):
    def _process(self):
      (...)

  class InvitePlayers(spade.
      Behaviour.PeriodicBehaviour):
    def _onTick(self):
      (...)

  def __init__(self):
    lb = self.LeaderBehaviour()
    ipb = self.InvitePlayers(30)
    psb = self.PartyStats()
    self.behaviours = [(psb, None),
        (lb, None), (ipb, None)]
```

The implemented role has been used for the implementation of artificial intelligence (AI) players inside the open source MMORPG TMW to create players which are able to create parties of players, invite new players and provide party statistics to party members.

The behaviours of the role have been implemented as nested classes inside the role class to be grouped together in one logical sequence of code, but could have been used separately, for example if the same behaviour is part of multiple roles. During the initialization of the role, instances of the behaviours are created and added to the list of behaviours.

To add the role to some agent (for example a TMW-player agent) one could use the code in listing 5.

Listing 5: Adding a role to an agent in Python/SPADE

```
role = Leader()
a = ManaWorldPlayer(SERVER, PORT,
    USERNAME, PASSWORD, CHARACTER, '%
    s@127.0.0.1' % USERNAME,
    SERVER_PASSWORD)
a.addBehaviour(AddRole(role))
a.start()
```

With these simple few lines of code, we were able to implement a very complex agent that acquires a role and starts enacting the defined behaviours.

5 Conclusion

Through the course of this paper we have shown a way of implementing agent roles in the context of the MMORPG domain, using the implementation possibilities provided by SPADE multi-agent development platform.

Roles that are defined as a part of this research represent normative concepts that denote sets of behaviors that can be played or performed by agents when enacting a specific role. Role enactment makes it possible for in-game characters to change their available actions dynamically, based on the role they enact. Having behaviour and actions defined in such a way makes the system's agents capable of adapting to various states of the system.

In this way very complex sets of behaviours can be implemented in a clear way with a higher level of abstraction then using common FSMs or behavioural tree methods. By using only a few lines of Python/SPADE code, we were able to define agent roles as sets of behaviours that can then be enacted by particular agents.

Our future research is aimed towards implementing more complex organizational features into MMORPGs by using agent technologies especially in regard of structural features.

Acknowledgments

This work has been fully supported by the Croatian Science Foundation under the project number 8537.

[1] For the sake of readability we have removed the details of each behaviour implementation, but the interested reader can refer to https://github.com/tomicic/ModelMMORPG/blob/master/TMWhlinterface.py for the whole implementation.

References

Aranda, G., Trescak, T., Esteva, M., Rodriguez, I., & Carrascosa, C. (2012). Massively multi-player online games developed with agents. In *Transactions on edutainment vii* (pp. 129–138). Springer.

Dignum, F., Westra, J., van Doesburg, W. A., & Harbers, M. (2009). Games and agents: Designing intelligent gameplay. *International Journal of Computer Games Technology*, *2009*.

Dignum, M. (2004). *A model for organizational interaction: Based on agents, founded in logic*. SIKS.

Gregori, M. E., Cámara, J. P., & Bada, G. A. (2006). A jabber-based multi-agent system platform. In *Proceedings of the fifth international joint conference on autonomous agents and multiagent systems* (pp. 1282–1284). ACM.

Huber, M. J. & Hadley, T. (1997). Multiple roles, multiple teams, dynamic environment: Autonomous netrek agents. In *Proceedings of the first international conference on autonomous agents* (pp. 332–339). ACM.

Marian, T., Dumitriu, B., Dinsoreanu, M., & Salomie, I. (2004). A framework of reusable structures for mobile agent development. In *Proceedings of ieee international conference on intelligent engineering systems (ines2004)* (pp. 279–284). IEEE.

Orkin, J. (2006). Three states and a plan: The ai of fear. In *Game developers conference* (Vol. 2006, p. 4).

Schatten, M., Đurić, B. O., Tomičić, I., & Ivković, N. (2017). Agents as bots–an initial attempt towards model-driven mmorpg gameplay. In *International conference on practical applications of agents and multi-agent systems* (pp. 246–258). Springer.

Schatten, M., Ševa, J., & Tomičić, I. (2016). A roadmap for scalable agent organizations in the internet of everything. *Journal of Systems and Software*, *115*, 31–41.

Schatten, M., Tomičić, I., & Đurić, B. O. (2017). A review on application domains of large-scale multiagent systems. In *Central european conference on information and intelligent systems*.

Westra, J., Dignum, F., & Dignum, V. (2008). Modeling agent adaptation in games. In *Bnaic 2008 belgian-dutch conference on artificial intelligence* (p. 381).

Westra, J., Van Hasselt, H., Dignum, V., & Dignum, F. (2008). On-line adapting games using agent organizations. In *Computational intelligence and games, 2008. cig'08. ieee symposium on* (pp. 243–250). IEEE.

Yannakakis, G. N. & Togelius, J. (2017). *Artificial intelligence and games*. Springer.

Computer Implementation of "Rational Pigs Game" Extended Model and Humans' Decision-Making in Related Scenarios

Robert Fabac, Danijel Radošević
Faculty of Organization and Informatics, University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
rfabac@foi.hr; darados@foi.hr

Abstract. *This paper examines the well-known Rational Pigs Game (RPG) scenario from game theory where two boxed gluttons try to consume as much food as possible from a common source. In our expanded game version, we introduced a third glutton and formulated computer game for playing in the laboratory, based on our previously featured software engine called Autogenerator. Thanks to this mechanism, the game theory model of the RPG acquires the attributes of a computer game and it can run by interaction of people, players on their computers. A laboratory experiment was performed where a total of 60 players using computers were engaged in decision-making in the role of gluttons. The players' behavior allowed us to gain useful insights into their rationalization, decision-making and learning processes.*

Keywords. computer implementation; rational pigs game; Nash equilibrium; decision making; learning styles; strategic choices; Autogenerator; clusters

1 Introduction

The problem of decision-making of rational gluttons has been described in various papers, e.g. in (Baldwin and Meese, 1979) the most relevant among which is (Hykšová, 2004). It has also been analyzed in several game theory books (McMillan, 1992), (Rasmusen, 2006). Recently, the rational gluttons scenario has been predominantly explored by Chinese scientists, e.g. in (Cheng et al., 2015), whose papers are mainly published by the CNKI scientific papers publisher.[1]

The scenario involves two pigs (gluttons) boxed in a cage and placed at its opposite ends. One glutton is large, the other one is small. By pressing the lever, one of the gluttons releases the food into the cage, but it falls far from him and close to the other (rival) glutton. According to various sources, for example in (Rasmusen, 2006), this game can be presented as a table which contains payoffs (earnings, obtained resources) for the players, as shown in Table 1. The gluttons' dilemma is whether or not press the lever and how often (in repeated games), because if they do not press the lever they will remain hungry and if they do they will cede the priority of feeding to their opponent (the other glutton).

Table 1: Rational gluttons game – payoffs

		Large glutton (LG)	
		pressure	*no pressure*
Subordinate glutton (SG)	*pressure*	15,45	1,50
	no pressure	30,30	0,0

McMillan (1992) provided an interpretation of this game, finding an analogy in the area of business, i.e. competition of oil cartels. We believe that an analog scenario can be recognized in the struggles of companies to conquer new markets, as well as in investment efforts in creating and applying innovations, and so on. RPG game elements can be identified at the opening of Chinese economy to foreign global companies. The state has allowed foreign companies to produce and sell on the Chinese market, providing them with possibility of a large income, but there is a condition - joint venture formation with some domestic company (model 50% - 50%).

Opening the pool with resources goes hand in hand with the formations of a partnerships between the big and the small "gluttons" and therefore some specific companies were formed such as: Changan Ford Automobile Co., Ltd. (joint venture between Changan Automobile and Ford Motor Company), BMW Brilliance Automotive Ltd. (BMW and Brilliance Auto), Beijing Benz Automotive Co., Ltd. (BAIC Motor and Daimler AG).

[1] http://en.cnki.com.cn

There are two Nash equilibria in the presented single-stage strategic game, and these are the solutions of this game in normal form in one stage that suggest optimal strategies. Besides, the game has a mixed-strategy Nash equilibrium, which also recommends appropriate mixed strategies, as described below.

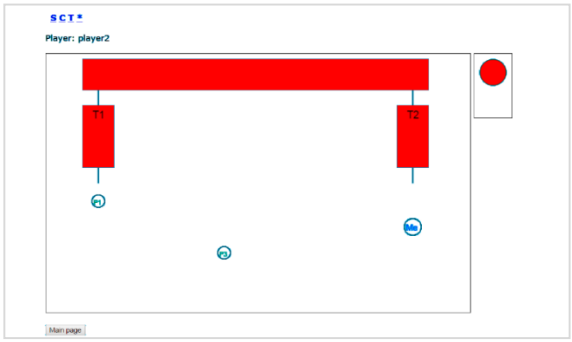


Fig. 1: Position of players – rational gluttons in extended scenario

The circumstances change in the case of a repeated game scenario. In addition, the innovative model formulated in (Fabac et al., 2014) involves a third glutton which acts asymmetrically to the benefit of the two original players. According to the formulated scenario of the game, the third player can not open the food container doors but he can and he wants to take a food. By introducing a third player, the game becomes more complex, the players' earnings are changed and the previous successful actions or the strategies needs to be re-examined. Such a rise in complexity is a common case for strategic interactions in business, military domain and other social contexts.

Table 2: Rational gluttons game (extended models) -- expected payoffs (Fabac et al., 2014)

TG (L, R)		**LG**	
		Pressure (P)	*No pressure (NP)*
SG	*P*	11,7; 35,0; 13,3 for (L) 12,0; 35,4; 12,7 for (R)	5,3; 35,4; 19,3 for (R)
	NP	20,6; 1,4; 38,0 for (L)	0; 0; 0

Furthermore, it is noteworthy that the gluttons in our model are given changed attribute values (size, speed of movement, speed of feeding and initial position in the cage). While the two gluttons who release food are positioned each at the opposite end of the cage, the third one (*tramp glutton (TG)*) is initially positioned in the middle. Although unable to release food, he moves around the open food tank to get hold of his portion. If both food tanks are open, the third glutton stops to figure out which end of the cage to go to (action L or R). The table presenting payoffs is now changed in relation to the initial one (Table 2).

Our research was aimed at establishing the behavior of humans in the roles of gluttons in a sequence of repeated games. This required setting of computer environment, i.e. an adequate game interface. The playing of a repeated game presented in tables 1 and 2 raises the question of the level of successfulness of players' strategies. As regards the finite repeated games, relevant is the term of the subgame perfect Nash equilibrium as described in, for example (Slantchev, 2004). On the other hand, unlike the well-known Prisoner's Dilemma, which was extensively studied, e.g. (Kendall et al., 2007), the studies of the iterated "rational gluttons game" almost don't exist, with the exception of the contribution of (Cheng et al., 2015).

This work attempts to contribute to the research of repeated RPG game. It is worth mentioning that according the designed rules the players/participants at first had no information on the game payoffs for various combinations of actions (strategies). Rather, they were gradually becoming aware of them through experience and learning. While the first task of this research was to enable the implementation of this game on computers, the second major task was to analyze the behavior of the game participants.

Related work was conducted in the domain of games according to the game theory models as well as in the domain of popular PC digital games. Among the numerous papers based on Iterated Prisoner's Dilemma (IPD), we highlight the research of (Burguillo, 2010) which offered a framework for using game theory and competition-based learning to increase the students' learning performance. Experimental "dictator game" was used and analyzed to determine the patterns of cooperative behavior of players, i.e. the presence of altruism, the reciprocity, the preferences for egalitarian choices, etc. (Diekmann, 2004), (Bardsley, 2008), (Engel, 2011).

In the domain online multiplayer computer games, Suznjević et al. (2011) analyzed the behavior of WoW (World of Warcraft) players based on categories of their actions. For the same game (WoW), based on certain behavioral variables, classes of player behavior were defined using clustering data methods in (Drachen et al., 2013). An empirical model of player motivations in online games was proposed in (Yee, 2006).

Concerning our first goal, computer implementation of the repeated RPG game was

achieved by the simulation mechanism based on Autogenerator, which had been presented in the aforementioned work (Fabac et al., 2014). Autogenerator embeds several features that were found to be useful in the implementation of the simulation mechanism. The most important one is that it allows specification of game parameters on a high level of abstraction and is capable of receiving the modified states of these parameters, including game results. Also, Autogenerator is aimed at building web applications, which allows networking among players.

2 Simulation mechanism based on Autogenerator

This mechanism is based on the specific implementation of our SCT generator model (Radošević et al., 2011), known as Autogenerator (Magdalenić et al., 2013), as can be seen in Fig. 2.

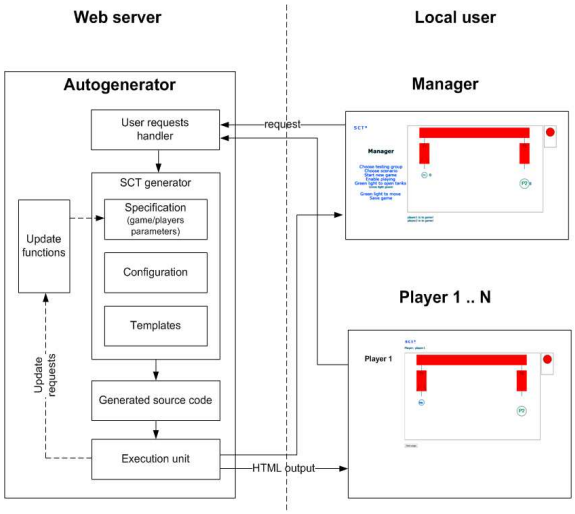


Fig. 2: Simulation mechanism used to implement strategic game

The main feature of Autogenerator is its capability to produce a demanding programming code and other application elements (such as HTML/CSS /Javascript code in the case of web application). Once created, the code is activated automatically. Autogenerator uses the option of script languages (Python in our example) to perform the programming code from variables and not just from the programming databases.

In the example of the strategic mechanism of games simulation, Specification consists of attributes that describe the generated code units (as virtual databases), the usual parameters of the game and the parameters pertaining to the tanks (of food) and players, as shown in the following example (Fig. 3):

```
OUTPUT:output
OUTPUT:out2
OUTPUT:out3
OUTPUT:out4
OUTPUT:out5
```

Output types (types of codes used for generating)

```
output:output/index.html
out2:output/player.cgi
out3:output/semaphore.cgi
out4:output/manager.cgi
out5:output/control_semaphore.cgi
```

Virtual databases used for generating

```
common:
+semaphore:red
+game_start:yes
+connected_tanks:yes
++fuel:60
```

Common parameters of the game

```
tank:T1
+position:1
+capacity:20
+fuel:20
tank:T2
+position:6
+capacity:25
+fuel:25
```

Properties of fuel tanks

```
player:player1
+earned:0
+player_position:1
+step:3
+size:1
player:player2
+earned:0
+player_position:6
+step:3
+size:2
```

Properties of players

Fig. 3: Simulation mechanism implementing strategic game

The types of outputs are associated with the highest-level templates in Configuration, as shown in Fig. 3, and are used to define the parts of the code which are generated and stored in variables (virtual databases). In the case of simulation mechanism, there are five code units to be generated; one that contains HTML and four scripts in Python. The common parameters of the game are defined under a common group, where "+" denotes subordination of attributes (for instance, parameter *fuel* is subordinated to attribute *connected_tanks*). These parameters change in the course of simulation performance in the way that attribute *semaphore* defines the current situation on the semaphore used in the simulation; *game_start* defines the current simulation stage (later it changes its value in the *game_open_tank* and

game_choose_direction stages of the game); *connected_tanks* is an option that defines the usage of the common tank (with the capacity of 60 units) or several unconnected tanks with their respective capacities. Each tank of the group is defined by its position, capacity and current level of fuel (food) in the tanks (it is used when the tanks are not connected). Finally, the group *player* defines properties of the players, such as the current amount of fuel, speed of movement (towards attribute *step*: higher value means slower movement) and size, which defines the speed of feeding (higher value means faster feeding).

Configuration consists of configuration rules and defines the connections between Specification and Templates (Radošević et al., 2011). Each configuration rule is defined by three elements:

• **Connection**. Each connection is physically positioned in Templates, usually marked by "#" signs; for example, *#title#*, defining the position where the real content must be placed. Connection is the key element which appears only once in Configuration, but it may appear one or more times in Templates.

• **Source**. Each connection has a respective source in Specification, for example, the source for connection, *#player#*, is the value of attribute *player*. The source can be defined as: particular, tank, or group.

• **Code template**. If the source is a tank or a group, the connection has its subordinated code template. If not, this element is left out.

Connections used in code templates define inclusion of content which may be from another code template or from the source, if template code is left out. Recursive connections (the ones leading to the same template) should be avoided. Similar to Specification, Configuration is organized in a hierarchical order so as to define the structure of the trunk. Configuration of the generated application is presented in Fig. 4.

Highest-level templates

#1#,,index.template
#2#,,player.template
#3#,,semaphore2.template
#4#,,manager.template
#5#,,control_semaphore.template

Connecting with attributes from specification and subordinated code templates

#links#,player,link.template
#tanks#,tank,tanks.template
#buttons#,tank,buttons.template
#player#,player
#players#,player,players.template
#players_display#,player,players_display.template
#earned#,earned
#fuel#,fuel
. . .

Fig. 4: Configuration of Autogenerator-based simulation mechanism

The initial rows of Configuration define Templates of the upper level that are connected with the initial rows of Specification, thus defining output types (e.g. *#1#* is connected with *output*, *#2#* with *out2*, etc.). Each row of Configuration includes Connection within '#' signs; attributes from Specification; and (possibly) the used code template. If the code template is not specified, the row defines substitution of Connection (each occurrence in the template code) by the value of the specified attribute from Specification. On the other side, the specified code template is used as many times as the specified attribute occur in Specification.

The application of Autogenerator enables implementation of the rational gluttons scenario in the way that humans can participate in the game in the role of gluttons and the game can be repeated through a sequence of iterations.

In the research, an iterative repeat of the same game has been successfully performed through the series of successive time phases.

3 Humans in the role of rational gluttons

The research was conducted on a sample of 60 respondents divided in 20 groups, each consisting of three members in the role of game players. The respondents were 2nd and 4th year students of the Faculty of Organization and Informatics, both male and females. Each group of players was led by a coordinator (arbiter, manager) who created a pre-planned scenario for each game. He signalized the beginning and the end of the game (Fig. 5). The scenarios were marked by numbers 1-7, at which the scenarios from the categories 1-4 referred to the standard RPG game with two players and those in the categories 5-7 included three players, i.e. they represented variants of the extended game (RPGE).

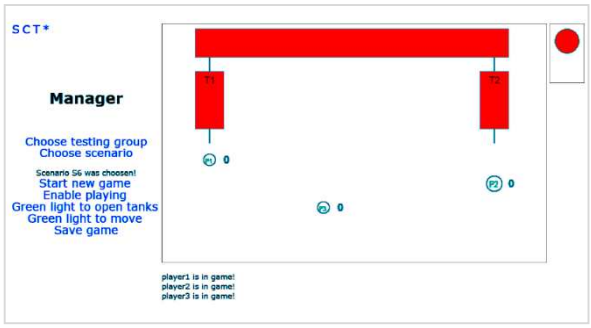


Fig. 5: The role of the game coordinator

The first 10-15 iterations of each game follow the two-player strategic model (Table 1). Further on, a third player is introduced into the same game, and the next 10-15 iterations are played following the same model, with payoffs as shown in Table 2. Because of the duration of the movement and feeding of the players (objects) in the simulation, a large number of iterations should be avoided, as it would create the saturation of people/participants.

Different game scenarios are implemented in the form of different automated generation specifications (Autogenerator). Once a new game is started, each player can see the current status of the game (Fig. 1).

In the course of the game, three persons are by their computers and the coordinator turns on the green light, thus signalizing the opening of the tank(s). Players 1 and 2 can open each tank on the opposite side (allowing the other player access to resources, i.e. food). That is action P (pressure); if there is no opening, action NP (no pressure) is chosen. It is also possible that players leave both tanks closed when given the opportunity to take action, but this option brings no gain to the players.

The players can see if one or more tanks are open, and choose to move towards one of the open food tanks. Those closer to the open tanks start feeding at once, whereas other players move over time towards the source of food and begin to feed themselves only when they reach them. One of the possible situations at the end of an iteration of the game is shown in Fig. 6.

The common food tank capacity (60 units) is distributed among the players who use them to the fullest, in keeping with this model. At the end of the game, after 25-30 iterations, the coordinator signalizes the end and stores the results. Then he can start a new game with a new group of players.

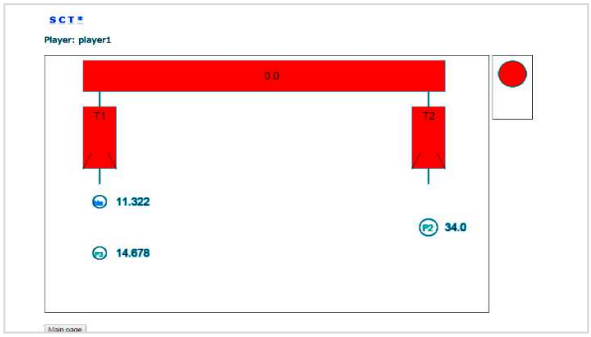


Fig. 6: A possible situation at the end of the game (payoffs).

Respondents participated in the model with two different game scenarios: the basic two-player RPG (RPG) with two players, and then the expanded game (RPGE), which includes the third player.

The main issue and aim of our analysis was to establish the behavioral characteristics of players in the computer design of the iterated "rational gluttons" model. In addition, we wanted to determine to what extent this real-life behaviour is in (in)congruity with the expected behavior. The behavior of individuals in the course of experiments as opposed to the behavior in real-life situations outside laboratories was discussed in various works, among which the most relevant are those of (Camerer, 2003), (Levitt & List, 2007) and (Benz & Meier, 2008). The topics and issues of motivation, engagement and learning in digital games have been dealt with in a number of contemporary research papers, such as (Iacovides et al., 2011).

4 Iterated RPG(E) game outcomes

Assuming the participation of rational players, the structure of payoffs (Table 1) indicates the presence of two Nash equilibria (NE) for one-stage game. According to definition, the output $x^* \in X_N$ is Nash's equilibrium from game G with n players, if the strategy x_i * of *player i* is the best response to x_{-i} * for all i = 1, ..., n, where x_{-i} * are the strategies of other players.

In this case strategy $x_i^* \in BR_i(x_{-i}^*)$ for all i = 1, ..., n. The set BR_i (BR as the "best response") form strategies that represent the best responses of (each) *player i* to the possible strategies of other participants.

NE exist for the following pairs of strategies of players SG and LG respectively:

(P, NP): $\boldsymbol{u}$(P, NP)= (10, 50), and

(NP, P): $\boldsymbol{u}$(NP, P)= (30, 30) (1)

at which $\boldsymbol{u}$ (utility) denotes players' payoffs. The third Nash equilibrium is in the domain of mixed strategies and can be recorded as follows:

$S^{NE}(SG)=(6/7; 1/7)$ and $S^{NE}(LG)=(2/5; 3/5)$, (2)
with associated payoffs:

$\boldsymbol{U}$ ($S^{NE}(SG)$; $S^{NE}(LG)$)= (12, 42.86) (3)

Thus, if the players choose their "pure" strategies "Pressure" and "No pressure" in the course of iterations of the base game with the frequency as defined in expression (2), they can expect payoffs (12, 42.9).

In the case of extended game (Table 2), payoffs change because of the third player who enters the game and wins for himself a part of the total sum (of resources, food).

This gives rise to the following question: how do the players behave when pursuing conflicting interests in the role of gluttons, which strategies do they choose, what characterizes their decision-making? All these issues are relevant in an iterated game with more than 20 repetitions. We therefore analyzed the

behavior of players in 20 groups, taking into consideration several indicative elements that impact the decision-making (see description of selected variables, below): obvious errors made during the game; reciprocal behavior; activity, or lack of it, regarding the opening of food tank (resources); non-cooperative actions; change of behavior at the start of extended game.

In the following, we applied the K-means clustering method using the JMP 13 software tool. Three different clusters were found in decision-making analysis at the position of Player 1 (small glutton), as shown in Fig. 7.

For example, the red cluster (cluster 2) describes the small player who is above average successful in the game of winning points (taking food). Furthermore, he is more active in that he presses the food lever more often and tends to move towards the opposite side although he can feed himself right away (a kind of small irrationality). His behavior is more cooperative in the way that he often opens the gate with food in some iteration ***k*** after his opponent released food lever in step **(k-1)**. Approximately 1/3 of all players in the role of small glutton belong to this cluster.

Table 3 shows results of the points won by Player 1 for the categories of clusters. We can see that the results of the players from cluster 2 exceeds the expected result of the mixed-strategy Nash equilibrium, whereas in the case of cluster 1 the achieved result is slightly lower than NE payoff (expression (3)).

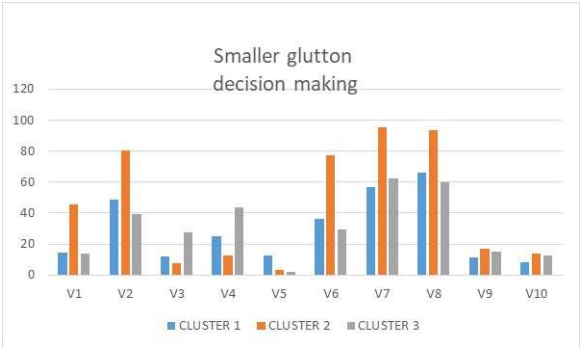


Fig. 7: Decision-making features of the player in the role of a small glutton

The characteristic *selected variables* that mark the actions i.e. the decision of the players are:

v1_ frequency of the choice to move towards the tank at the opposite end, when the one next to the player is open as well (category of mistake)

v2_ frequency of the choice to press the lever (opening) after the opponent has already pressed (opened) it in the previous iterative step (the concept of reciprocity)

v3_ frequency of the choice not to open after the opponent has already chosen to press (open) it in the previous step (no reciprocity)

v4_ frequency of the choice not to press the lever (not to open)

v5_ frequency of the event where one of the tanks is open for the player, but he chooses not to go for food (category of mistake)

v6_ frequency of situations where both players repeat the same strategies as in the previous step

v7_ frequency of pressings, i.e. openings in the second part of the game

v8_ frequency of openings throughout the game

v9_ number of the points won (payoff; resources used) throughout the game

v10_ number of the points won (payoff; resources used) in the second part of the game when the third glutton is present

Table 3: Results obtained by participants in the role of a small glutton

Cluster/ number of members	**RPG points/ standard deviations**	**Points for the 2nd part (RPGE)/ standard deviations**
1 / 5	11.07196/ 1.00011	7.990429/ 1.86532
2 / 7	17.06259/ 3.26629	13.83261/ 3.60227
3 / 8	15.33144/ 2.79018	12.32286/ 3.61995

The decision-making analysis described above was applied analogously to player 2. Fig. 8 shows three clusters and the typical decision-making of the participant in the role of large glutton (Player 2).

The blue cluster (cluster 1), performed by Player 2, turned out as the most successful one, although his result is lower than the expected result for the mixed-strategy Nash equilibrium (Table 4, expression (3)). Only about 1/7 of all players in the role of large glutton belong to this cluster.

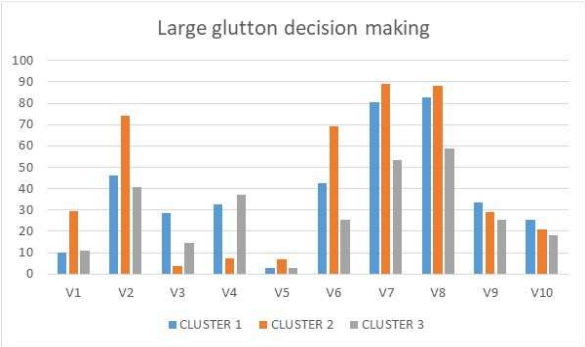


Fig. 8: Decision-making features of the player in the role of large glutton

While the aforementioned equilibrium stimulates the action of opening for Player 1 (see expression 2), a moderate activity of opening is recommended for Player 2 (in 40% of cases).

Obviously, the blue cluster (cluster 1) is not in the lead as regards the number of openings and is prone to playing with less cooperativeness and reciprocity (Fig. 8). Owing to its attributes and the constellation of the game, such attitude brings him more success.

Table 4: Results obtained by the participants in the role of large glutton

Cluster/ number of members	**RPG points/ standard deviations**	**Points for the 2nd part (RPGE)/ standard deviations**
1 / 3	33.45026/ 3.110606	25.16635/ 5.63685
2 / 9	28.79602/ 3.86564	20.95162/ 5.38918
3 / 8	25.54748/ 2.77229	18.16316/ 2.59669

The data on standard deviation indicate a statistically significant difference, especially between groups 1 and 3 in terms of the points won. The third glutton (TG) that is active in the second part of the game creates more damage to the result of the "large glutton" (Table 3 and Table 4).

5 Concluding remarks

Autogenerator makes it possible to implement scenarios of the well-known "Game of Rational Pigs", in a way that people can take the roles of gluttons and also that the game can be repeated in the same or similar form through a series of iterations. Since the previous research does not provide a systematic study of players' behavior in the model of the iterated "game of rational gluttons" this work is an original effort in that direction. Twenty groups of players participated in the computer-based gaming with two different scenarios: basic RPG in ten to fifteen repetitions and then expanded (RPGE), including third-party players, also in approximately 10-15 iterations. The analysis of players' decision-making in laboratory conditions has provided useful insights into rationality of players, the peculiarities of their decision-making and learning styles. At the beginning of each game, the participants did not have information about the payoffs (the abilities to conquer resources) for the game or various combinations of actions (strategies). Rather, they were acquiring knowledge through experience and learning through gaming. The applied behavioral analysis focuses on decision-making features leading to a greater or lesser success in the game.

It was found that among the most successful and least successful group of players, in the role of both the bigger and the smaller glutton, there was a statistically significant difference in overall success.

Preferred activities of the decision makers (players) and behavioral logic that lead toward success are different in the cases of larger and smaller players, as described earlier. Some individual players succeed in grasping and applying the right logic to success, while others fail to do so. Within the "large glutton" population, a smaller proportion of players succeed in identifying those activities that bring greater success (about one-seventh of them), unlike the situation found in the group of "small gluttons" (about one-third of them).

Why the players in the role of a "large glutton" are not so successful like players that play the role of a "small glutton"? A successful large glutton has such pattern of behavior that he is not in the lead as regards the number of openings and is prone to playing with less cooperativeness and reciprocity.

One of the key reasons for the large glutton's failure is their gaming philosophy and general approach to playing, which typically does not imply patience and persistence in postponing action. In a real-world situation, when an individual needs some resources (food and the like), he is ready to invest himself and consistently implement a strategy that may include waiting. However, in laboratory conditions, there is no struggle for survival. Instead, people are, consciously or subconsciously, aware that they are not participating in a scenario that can have actual implications.

Furthermore, through the gaming experience, the "large glutton" learned that with the same action of feeding he wins much more points than his opponent who acts as a "small glutton". He does not have a real need to take more resources than the other player, in each interaction. Therefore, his decision to open access to resources, at a particular iteration of the game is an unselfish behavior similar to the "actions of giving" recorded in the well known "dictator game". This seemingly irrational behavior of players is recorded in numerous laboratory experiments of other famous games. According to (Andreoni and Miller, 2002) "…subjects in economic laboratory experiments have clearly expressed an interest in behaving unselfishly".

The extended game RPG(E) scenario also contributes to a weaker "large glutton" result as the third player in the second stage of the game wins on average a lot more resources at the expense of a bigger than at the expense of a smaller glutton.

Finally, it should be emphasized that in this research, as well as in other works dealing with elements of game theory, social psychology, economic rationalization, the results must be considered taking into account the difference between the possible behavior in a possible real situation and

the determined behavior of the examinees in the laboratory. The behavior of the respondents who play in the lab on their computers is not characterized by the same motivation that drives the (hungry) gluttons who resort to confrontation and circumvention in the cage.

Since the shown experiment can be reproduced with different groups of respondents, hence the challenging issues about motivation and rationality could be in focus of our future researches of iterated RPG(E) game playing on computers.

References

Andreoni, J. & Miller, J. (2002) Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. Econometrica 70, 737–753 (2002).

Baldwin, B. A.; Meese, G. B. (1979) Social Behaviour in Pigs Studied by Means of Operant Conditioning. Animal Behaivour, 27(1979), 947–957.

Bardsley, N. (2008) Dictator game giving: altruism or artefact? Experimental Economics, 11 (2). pp. 122-133.

Benz, M., Meier, S. (2008) Do people behave in experiments as in the field?—Evidence from donations. Experimental Economics 11, 268–281

Burguillo, Juan. (2010) Using game theory and Competition-based Learning to stimulate student motivation and performance. Computers & Education. 55. 566-575.

Camerer, Colin (2003) Behavioral Game Theory. Princeton: Princeton University Press.

Cheng, Daizhan; Fenghua He, Hongsheng Qi, Tingting Xu (2015) Modeling, Analysis and Control of Networked Evolutionary Games. IEEE Trans. Automat. Contr. 60(9): 2402-2415 (2015)

Diekmann, A. (2004) The power of reciprocity: Fairness, reciprocity, and stakes in variants of the dictator game. The Journal of Conflict Resolution 48: 487–505.

Drachen, A.; Thurau, C.: Sifa, R. & Bauckhage, C. (2013) A Comparison of Methods for Player Clustering via Behavioral Telemetry. In Proceedings of Foundations of Digital Games 2013 (Chenia, Greece). Society for the Advancement of the Science of Digital Games (SASDG) Publishing.

Engel, C. (2011). Dictator games: A meta study. Experimental Economics, 14, 583–610.

Fabac, R.; D. Radošević, I. Magdalenić (2014) Autogenerator-Based Modelling Framework for Development of Strategic Games Simulations: Rational Pigs Game Extended, The Scientific World Journal, Hindawi, August 2014.

Hykšová, M. (2004) Several Milestones in the History of Game Theory. Jubiläen - Chance oder Plage? VII. Österreichisches Symposion zur Geschichte der Mathematik (pp. 49-56). Technische Universität Wien, Vienna, Austria, 2004.

Iacovides, Ioanna; James Aczel, Eileen Scanlon, Josie Taylor, Will Woods (2011) Motivation, Engagement and Learning through Digital Games. *IJVPLE* 2(2): 1-16 (2011)

Kendall, G.; X. Yao and S. Chong (Eds.) The Iterated Prisoner's Dilemma: 20 Years On. Singapore: World Scientific, 2007

Levitt, Steven D. and John A. List (2007) What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?," *Journal of Economic Perspectives*, 2007, 21 (2), 153–174.

Magdalenić, I.; D. Radošević, T. Orehovački (2013) Autogenerator: Generation and Execution of Programming Code on Demand. Expert Systems with Applications (ESWA), Volume 40, Issue 8, Pages 2845-2857, June 2013.

McMillan, J. (1992) Games, Strategies &Managers, HowManagers Can Use Game Theory to Make Better Business Decisions, Oxford University Press, 1992.

Radošević, D.; I. Magdalenić (2011) Source Code Generator Based on Dynamic Frames, Journal of Information and Organizational Sciences (JIOS), ISSN 1846-3312, 35 (2011), 1, 2011.

Rasmusen, E. (2006) Games and Information, Fourth Edition: An Introduction to Game Theory, *Wiley-Blackwell*

Slantchev, Branislav L (2004) Game Theory: Repeated Games, Department of Political Science, University of California – San Diego. Retrieved from http://users.auth.gr/~kehagiat/Research/GameTheory/02Courses/Course3/07-repeated-games.pdf

Suznjevic, M, I. Stupar, and M. Matijasevic (2011) MMORPG player behavior model based on player action categories, in NetGames, IEEE, 2011.

Yee, N. (2006). Motivations for play in online games. CyberPsychology & Behavior, 9(6), 772-775.

Data and Knowledge Bases

Challenges Due to Excessive Amount of Online Data and (Mis)Information

Katarina Rojko, Dejan Jelovac

Fakulteta za informacijske študije v Novem mestu

Ljubljanska cesta 31a, 8000 Novo mesto, Slovenia

{katarina.rojko, dejan.jelovac}@fis.unm.si

Abstract *The overload of data and information (including misinformation) that fill the World Wide Web and which are generally in circulation, became our contemporary problem, as information reliability is very often doubtful. Insofar as there is more information in circulation, their usage is reduced proportionally and, consequently, their meaning is reduced proportionally. Digital skills of EU citizens in general are underdeveloped, while amount of digital data grows exponentially. For this reason, constant development of digital competences and ICT literacy as a result of systematic education, social control, criteria for verifying reality, recognition of fake data, etc, is required.*

Keywords digital data, data quality, digital skills, ICT literacy, (mis)information

1 Introduction

Internet enables efficient, cost-effective data collection and facilitates access to large amount of data. However, since its use requires certain skills, among others, Hewson, Vogel and Laurent (2015) wrote a guide how to use internet as a tool for conducting research. Such guides are required since many researchers lack of skills for conducting research, using options offered by immense amount of data available online (Rojko, 2016). Moreover, the entire society has access to online data conditionally, and people are looking for information on World Wide Web (WWW) all the time.

The excessive amount of data is another problem, and every second it is more serious, as data is growing at a rapid pace in contemporary information society. Micro focus (2017) stated that 44 billion GB of data was created per day in 2016, while the growth to 462 billion GB of data created per day is predicted in 2025. Micro focus also revealed that 90% of data created is unstructured, so big data analytics and archiving tools are critical in being able to manage all this data.

Every day, we create 2.5 quintillion bytes of data. To put that into perspective, 90% of the data in the world today has been created in the last two years alone – and with new emerging technologies, devices, and sensors, the data growth rate will likely accelerate even more (IBM Marketing Cloud, 2016). For this reason, we are overwhelmed by the excessive amount of information. Namely, regardless of industry or profession, people need the right information at the right time to make truly confident and well-judged, productive decisions. But due to rapidly growing amount of information, it is harder and harder to separate the "signals from the noise", and to "discern the insights from the hindsight" (IBM Marketing Cloud, 2016).

As the internet has become flooded with untrustworthy information, some of which is intentionally misleading, it is necessary to know, how to recognize misinformation. However, the main question is, how to find and recognize "correct" information? For illustration, in 2016 there was 215 billion of emails sent every day, and 269 billion in 2017, while more than half were spam (Micro Focus, 2017). How can "typical user" deal with this problem? Good digital skills are important, but there is much more to know, and for this reason already since the turn of the century handbooks as "Web of Deception: Misinformation on the Internet" from 2002, by Anne P. Mintz, ed., are published.

There is another dilemma – a reflection on what is "correct" data. Seifert (2017) argues that what must be added to our understanding of misinformation in the "post-truth era" is our experience of misinformation, as the processing of information changed in the "post-truth" world.

In June 2016, the U.K. held a referendum on its membership in the European Union, and in November

2016, the U.S. held its national elections. In the run-up to both of these important decisional events, the internet with its burgeoning collection of information dissemination appli-cations, influenced the decisions of voters (Cerf, 2017). The disturbing aspect of these (and many other decisional events) is the quantity of poor-quality content, the production of deliberately false information, and the reinforcement of bad information through the social media. Besides, people began to read superficially, since because of the excessive amount of data, it is even more difficult to deepen into a single subject to read, and the quantity often wins.

In summer 2017, Pew Research Center and Elon University's Imagining the Internet Center conducted a large canvassing of technologists, scholars, practitioners, strategic thinkers and others, asking them to choose one of the two answer options:

a) The information environment will improve – in the next 10 years, on balance, the information environment will be improved by changes that reduce the spread of lies and other misinformation online.
b) The information environment will NOT improve – in the next 10 years, on balance, the information environment will NOT be improved by changes designed to reduce the spread of lies and other misinformation online.

Out of 1,116 respondents, 51% chose the option b), and 49% chose the answer a). Participants were next asked to explain their answers. In continuation we reveal some of these follow-up responses (Pew Research Center, 2017):

The quality of information will not improve in the coming years, because technology can't improve human nature all that much. Christian H. Huitema

In the arms race between those who want to falsify information and those who want to produce accurate information, the former will always have an advantage. David Conrad

We live in an era where most people get their 'news' via social media and it is very easy to spread fake news. … Given that there is freedom of speech, I wonder how the situation can ever improve. Anonymous project leader for a science institute

In order to reduce the spread of fake news, we must deincentivize it financially. Amber Case

When the television became popular, people also believed everything on TV was true. It's how people choose to react and access to information and news that's important, not the mechanisms that distribute them. Irene Wu

We can't machine-learn our way out of this disaster, which is actually a perfect storm of poor civics knowledge and poor information literacy. Mike DeVito

These responses clearly indicate the problem we are facing today. Some respondents are optimists, while others believe that information environment will not improve, as it is becoming harder and harder to find the right information, and predict that a larger digital divide[1] will form.

The internet's continuous growth and accelerating innovation also allows more people and artificial intelligence to create and instantly spread manipulative narratives. Furthermore, human tendencies and infoglut drive people apart and make it harder for them to agree on "common knowledge" (Pew Research Center, 2017).

Beside structured data, there is also much wider pool of unstructured data and different advanced tools are getting available for analysis of. But unstructured online data analysis and textual analytics require business intelligence skills and use of appropriate software tools. This means that not only digital skills are sufficient; there is also a condition of access to required sources, most often determined by the income level or inclusion in certain communities (Rojko, 2016).

Due to above mentioned concerns we decided to take a focus on needed skills for finding the information online and digital skills are one of the most important for the ability to find reliable information in required time.

2 Research methodology

Within the initial phase of the research, we did a systematic analysis of sources and a review of literature, in combination with experiences obtained during past two decades of active internet usage. This provided us with the theoretical and practical basis for formulation of our hypotheses.

For the purpose of empirical analysis, we used the data from Eurostat, which allows conclusions on reliable official dataset of data and thus increases certainty of our findings. The basis for the analysis

[1] Classical sociological theories of inequality, as well as empirical evidence (Ragnedda and Muschert, 2013), define digital divide as the unequal access and utility of internet communications technologies and explore, how it has the potential to replicate existing social inequalities, as well as create new forms of stratification. They examine how various demographic and socio-economic factors including income, education, age and gender, as well as infrastructure, products and services affect the internet use and access.

were digital[2] skills of individuals, variables that show level of digital competences.

To determine different impacts on digital skills, we also studied certain social exclusion indicators, and compared digital skills' levels to information[3] skills' and software[4] skills' levels.

We used only indicators' values for the last available year, to present the most current situation. We nonetheless encountered certain limitations, e.g. data for the age group under 16 was not available. Nevertheless, qualitative data analysis in combination with the literature review and obtained practical experiences, enabled us to conduct critical assessment of sources and theories, and credible verification of set hypotheses.

3 Research goal

Besides presented impact factors form Eurostat database, also others have significant impact on the ability for successful separation of the right from wrong, misleading information. Those are e.g. the kind of information we are looking for, level of ICT availability, usage and data quality, personal ability to develop a critical distance to obtained information, etc.

The goal of our research was to present the most up-to-date situation to convince readers that this topic requires much greater attention and coordinated actions of all actors involved to decrease levels of poor civics knowledge and poor information literacy. This can be improved only by better awareness, constant education, continuous research and data accessibility. Namely, other studies do not focus on current situation from the angle of challenges and conditions for society's prosperity based on information access and use, although it is generally accepted that "The oil of 21st century is data".

We decided to focus on measurable data, including data on digital skills in frame of European Union (EU) as a whole (28, 15), and we exposed situation in Slovenia and Croatia. We also researched other Eurostat's "Digital Society" data, to find explanations for observed situation, which enabled us to make conclusions that are more credible. Besides, we added some figures on internet size and number of internet users.

Based on initial data and literature review, we have set the following theses:

- Digital skills of EU citizens do not reach sufficient level to enable them to find reliable information on world wide web in required time.

- Development of digital skills is crucial in the contemporary period of excessive amount of online data and (mis)information.

4 Data Analysis

To support our theses, we firstly checked the level of digital skills in EU 28 and EU 15, while we also focused on situation in Slovenia and Croatia. For the purpose of our research, we decided also to consider certain social exclusion parameters that supposed to have impact on the level of digital skills. Moreover, we compared digital to information and software skills, to provide explanations for different rates.

Nonetheless, we are aware that the skills' rates do not answer the question, who is able to find reliable information faster; they only provide measurable indicators for exploring certain obstacles and conditions for it.

Data below thus show the most recent levels of digital skills in EU and other measurable indicators that can serve as the explanation of different skills' levels and provide understanding why certain groups of people have problems to find reliable information on the internet, and why there is a greater potential to mislead them by partial or wrong information.

[2] Persons that have been using internet during last 3 months are attributed a score on four digital competence domains: information, communication, content-creation and problem-solving, depending the activities they have been able to do. The scores are basic, above basic and below basic. Individuals not using internet are classified without digital skills. The four digital competence domains are aggregated in four logical groups (Eurostat, 2018).

[3] Information processing skills refers to the ability to identify, locate, retrieve, store, organize and analyze digital information, judging its relevance and purpose. The indicator is based on five activities internet users have been able to do online during previous 3 months. The scores are basic, above basic and none. Individuals not using internet are classified without digital skills (Eurostat, 2016).

[4] Software skills for content manipulation refer to the ability to create and edit new content (from word processing to images and video); integrate and re-elaborate previous knowledge and content; produce creative expressions, media outputs and programming; deal with and apply intellectual property rights and licenses. The indicator is based on six activities internet users have been able to do during previous 3 months. The scores are basic, above basic and none. Individuals not using internet are classified without digital skills (Eurostat, 2016).

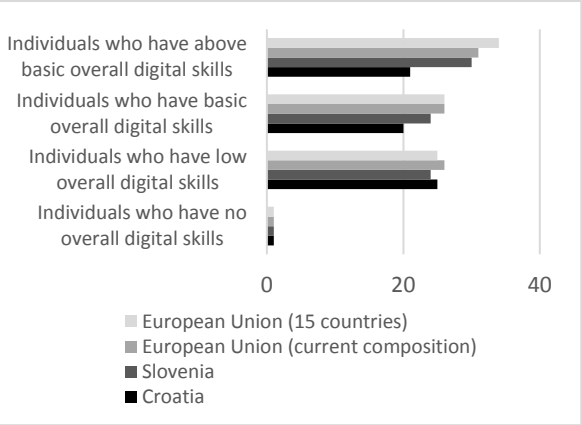


Figure 1. Digital skills of individuals (%) in 2017 (Eurostat, 2018)

Figure 1 shows the rates of digital skills in EU in 2017 and includes comparison with situation in Slovenia and Croatia. In EU 28, 52% of individuals have only low or basic overall digital skills, while 31% have above the basic overall digital skills, and 1% has no overall digital skills. In EU15 digital skills are, as expected, developed better, as 34% of individuals have above the basic overall skills.

Comparison of Slovenia and Croatia to the EU average reveals that both Slovenia and Croatia lag behind with 30% and 21% of individuals with above the basic overall digital skills respectively, while this percentage in EU28 varies from 10% in Romania to 58% in Iceland. Slovenia occupies nineteen spot in this term, while Croatia twenty-sixth, which indicates, that Slovenia and Croatia have to invest stronger in the development of digital skills, to avoid staying the laggards in this term.

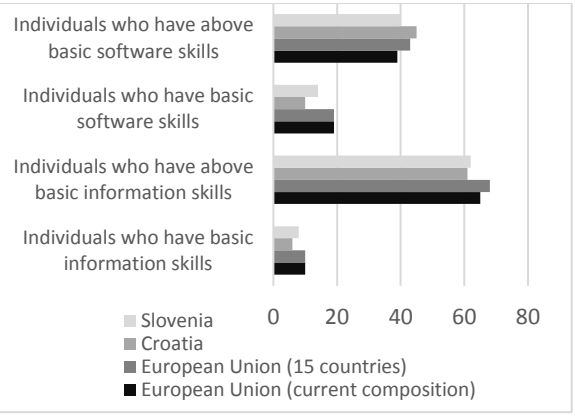


Figure 2. Information and software skills of individuals (%) in 2015 (Eurostat, 2016)

From the Eurostat's (2016) data on information and software skills' rates, we can find out that in 2015 as much as 65% of individuals had above the basic information skills in EU 28, which is more than double in comparison with above the average digital skills (28%) in the same year.

This clearly indicates that digital skills are significantly underdeveloped and that with such insufficient level, a considerable number of individuals cannot exploit their potential to search for digital information successfully.

Even worse situation is observed in terms of software skills' comparison, as Eurostat (2016) data revealed that 42% of individuals in EU28 had no software skills at all in 2015. Furthermore, the data on both information and software skills' rates in total, again reveal a smaller rate of skills in Slovenia and Croatia as in the EU 28.

In sought of the answer, why there is such a big gap between individuals in terms of digital, informational and software skills' rates, we decided to look at three social exclusion indicators (55 to 74 years old; low education; unemployed or inactive or retired) which showed us the correlation (Eurostat, 2016).

We could conclude, that in all three different kinds of skills, there is a big gap between all individuals and individuals marked with at least one or two social exclusion indicators, however still the biggest difference was in terms of information skills - the most "traditional" skill (Eurostat, 2016).

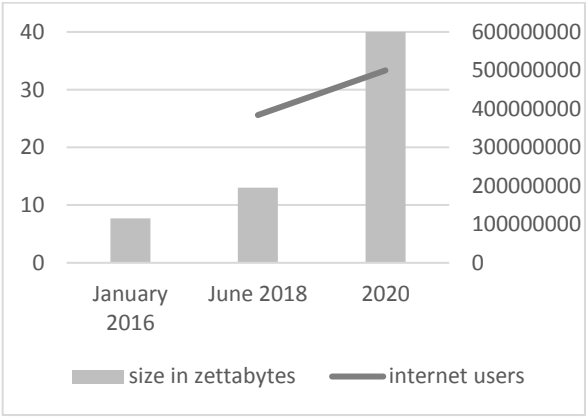


Figure 3. The size of the internet and number of internet users (live-counter.com, 2018)

The size of the internet doubles about every 2 years (live-counter.com, 2018). In July 2018, the counter revealed that the size of internet was 13,4 zettabytes, while, year and a half earlier (in January 2016) it was 7,7. In another year and a half (by 2020), the amount of data is expected to enlarge to 40 zettabytes. It is estimated that by then 50 billion devices will be connected to the internet. Those devices are used currently (July 2018) used by 3,8 billion of internet users, while by 2020, the total number of internet users is predicted to rise to over 5 billion (live-counter.com, 2018).

Data presented in Figure 3, clearly show huge data growth in terms of quantity in past two years, while the projections for the next years do not exhibit any

slowing. This clearly indicates that in the next "Big data years" digital skills will be even more required.

5 Discussion

Based on our research we were able to confirm our first thesis that digital skills of EU citizens do not reach sufficient level to enable them to find reliable information on world wide web in required time. Besides digital skills, also information and software skills are important basis.

Data we used show the most recent levels of digital skills in EU and other measurable indicators that can serve as the explanation of different skills' levels and provide understanding why certain groups of people have problems to find reliable information on the internet and why there is greater potential to mislead them by partial or wrong information.

There are certain conditions that must be met, to be able to look over the WWW for the right information. There is a requirement of having access to and be able to use internet connected device in order to access online sources. Certain groups of people are in disadvantage in this requirement, and for this reason we also considered social exclusion indicators.

Besides, our research revealed a big gap between digital, information and software skills. Comparing only above the basic skills of individuals with high, medium and low education, the smallest difference in skills' rates is among individuals with high formal education, while the biggest difference in skills' rates is among individuals with no or low education.

Moreover, we found out that digital skills are significantly underdeveloped, meaning that with the low rate of digital skills, notable amount of individuals cannot exploit the potential they have with own information skills. Thus much greater attention should be devoted to obtaining digital skills, especially among individuals with no or low formal education, while also among mid and high educated individuals.

Nonetheless, the target value in percentage with respect to the digital skills level in general, should be the same or close to informational skills' rates, we argue. Namely, if these levels would be close or the same, much greater exploitation of internet as an immense source of information would be possible.

Furthermore, there are also other requirements that must be met, as is the knowledge how to check data reliability, having accesses to different databases, and ability to develop critical distance to (mis)information. Here, dissemination of fresh knowledge and solutions by universities and ICT suppliers can help to solve the problem of society on how to find and recognize the "correct" information in Big- data years.

However, since the ability to find the reliable information on WWW is not only a technical issue but also a social issue, as internet changed the ways of information obtaining, constant education from early childhood, technology possession, information and digital skills' development, etc. are required. In addition, the knowledge where to find appropriate data with the ability to access this data is conditional. Here notable difference among different age groups and education level are observed, while also other factors certainly have impact, as income, location, origin, etc.

Based on presented data we also found out, that Slovenia and Croatia lag behind EU in terms of digital skills level, which leads to slower transition to knowledge society. Slower transition to knowledge society might further lead to consequent slower economic growth, as the quick access to information is becoming even more important in globalized e-society.

We could also confirm our second thesis that development of digital skills is crucial in the contemporary period of excessive amount of online data and (mis)information. Namely, the amount of data is growing at a very rapid pace, as presented by the live-counter.com (2018) data, while vast majority data of is available online and generated only in recent years. For this reason, people with underdeveloped digital skills are in big disadvantage, which is also recognized by governments, who regularly provide free options for the inclusion of certain groups of people (usually marked with social exclusion indicators) into the programs which help them to raise the level of ICT literacy and digital skills.

6 Conclusion

In the digital era the ability to find correct information is conditional, as individuals without good skills for online research experience big disadvantage. In addition, the future of internet as a source of reliable information might become questionable, as it has become flooded with untrustworthy information, some of which is intentionally misleading or corrupted.

Only a small share of society will find, use and perhaps pay premium for information from reliable sources. Those will separate from those who are not selective enough or who do not or are not able to invest either time or money in doing so. For this reason, concerns about how vast majority of society will be able to find and use accurate information, are justified.

Moreover people usually do not use or have access to big data software tools, thus the flood of digital data and sources available is making it even more difficult now for them to avoid (mis)information gathering, as was in the period when there was no or less digital data available.

However, the amount of data is growing exponentially, and finding the correct information becomes restricted to WWW users who are able to follow the changing situation in globalized "post-truth" society.

It is necessary to know, how to recognize misinformation. Good digital skills are important, but we also have to be able to recognize poor-quality content, the deliberately false information, and the bad information, many of these spread through the social media.

The rise of "fake news" and the proliferation of narratives that spread online are challenging also publishers and platforms. Some are trying to stop the spread of false information, but their task is Sisyphus' work.

7 References

Cerf, V. G. (2017). Information and Misinformation on the Internet. *Communications of the ACM*, 60(1), 9-9. ACM, NY, USA. Retrieved from https://dl.acm.org/citation.cfm?id=3018809

Eurostat. (2016). *Digital society*. Retrieved from: http://ec.europa.eu/eurostat/data/database

Eurostat. (2018). *Digital economy and society*. Retrieved from: http://ec.europa.eu/eurostat/data/database

Finnegan, R. (2005). *Participating in the Knowledge Society: Researchers Beyond the University Walls*. Palgrave Macmillan UK.

Halfpenny, P., & Procter, R. (Eds.). (2015). *Innovations in digital research methods.* Los Angeles [etc.]: Sage.

Hewson, C., Vogel, & C., Laurent, D. 2016. *Internet research methods.* Second Edi. Los Angeles [etc.]: Sage.

IBM Marketing Cloud. (2016). *10 Key Marketing Trends for 2017 and Ideas for Exceeding Customer Expectations.* Retrieved from https://public.dhe.ibm.com/common/ssi/ecm/wr/en/wrl12345usen/watson-customer-engagement-watson-marketing-wr-other-papers-and-reports-wrl12345usen-20170719.pdf

Live-counter.com. (2018). *How big is the internet.* Retrieved from http://www.live-counter.com/how-big-is-the-internet/

Micro focus. (2017). *Growth of internet data in 2017*. Retrieved from https://www.slideshare.net/Micro-Focus/growth-of-internet-data-2017.

Ó Dochartaigh, N. (2012). *Internet Research Skills*. Sage.

Pew Research Center. (2017). *The Future of Truth and Misinformation Online.* Report, October 19, 2017. Retrieved from http://www.pewinternet.org/2017/10/19/the-future-of-truth-and-misinformation-online/

Ragnedda, M., & Muschert, W. G. (2013). *The Digital Divide: The Internet and Social Inequality in International Perspective*. Routledge.

Rojko, K. (2016). Requirements and obstacles of e-Research, *Research in Social Change*, 8(2), 53-74.

Salmons, J. (2016). *Doing qualitative research online*. Los Angeles [etc.]: Sage.

Seifert, M. C. (2017). The Distributed Influence of Misinformation. *Journal of Applied Research in Memory and Cognition*, 6(4), 397-400. Elsevier Inc.

Van Deursen, A., & van Dijk, J. (2014). *Digital Skills: Unlocking the Information Society*. Springer.

Wishart, J., & Thomas, M. (2016). *E-Research in Educational Contexts: The Roles of Technologies, Ethics and Social Media.* Taylor & Francis.

Preliminary Study for a Survey-Based Fuzzy Membership Function Definition for Imprecise Quantification in Croatian*

Leo Mršić, Ph.D.
Algebra University College
Zagreb, Croatia
leo.mrsic@algebra.hr

Sandro Skansi, Ph.D.
University of Zagreb
Zagreb, Croatia
skansi.sandro@gmail.com

Robert Kopal, Ph.D.
Algebra University College
Zagreb, Croatia
robert.kopal@algebra.hr

Abstract. *In this preliminary report we propose a survey-based method for defining imprecise quantification for Croatian. By using the results of a survey conducted among students, a fuzzy membership function for each of the precise and imprecise quantification terms can be defined, with possible extensions to type-2 fuzzy memberships. An earlier version of this paper was submitted and subsequently withdrawn from the ACE-X 2017 conference.*

Keywords. Fuzzy Set Membership, Fuzzy Quantifiers, Linguistic Variables, Croatian Quantification

1 Introduction

Quantifiers in language and linguistic variables in fuzzy logic share a common theme. The study of quantification dates back to at least Aristotle's Organon, and in the modern period was revived by G. Frege (Frege 1879) and perfected by C. S. Peirce (Peirce 1885). Quantifiers today are an essential feature of all major applied logic systems with a few noteworthy exceptions (SAT-related logics (Marek 2009), propositional modal logic (Blackburn, de Rijke, Venema 2002)). Quantifiers however, are an essentially linguistic constructs allowing the reference to a number of terms (Peters and Westerstahl 2006). Unprecise quantifiers were a motivational factor behind the development of mathematical fuzzy logic (Hajek 1998), and a natural approach to their meaning is via fuzzy membership functions. Linguistic variables were at first considered in the context of process theory (D'Ambrosio 1989), but as fuzzy logic became more know today they are discussed in the context of fuzzy logic (Ross 2010) (Mrsic 2017).

A second interesting phenomenon is that nonconventional quantifiers (Torza 2015) (Peters and Westerstahl 2006) are relative to a given language, and different languages possess natural quantifier terms both for precise and imprecise numberings. Precise quantifiers can be thought of as number terms, uniquely denoting a precise quantity. An example of such term could be "ten", but their apparent precision fails to take in account their ability to reference an entity (Donnellan 1966) (Donnellan 1972). A statement "Fetch me that jar with ten bolts" may be successful at referring to the correct jar (containing e. g. 12 bolts). The example might be made even more elaborate by stipulating the presence of another jar with e.g. 143 bolts next to it. This may seem like a minor point, but it points out to the inherent imprecision even with precise numbering terms when considering the everyday communicational aspects of the language quantifiers.

(A first version of this paper under the title "Learning Fuzzy Membership functions for Slavic Quantifier Terms" was intended to be published with the ACE-X 2017 conference, but we felt that the paper at that time needed substantial revision and expansion, as well as refocusing, and we have subsequently withdrawn it from publication prior to the conference. The present paper is a revised and refocused version of the previous unpublished and unpresented paper.)

2 Basic quantifier and number term usage

The basic quantifiers in first order logic are "all" and "exists", and they are defined to be true when combined with a property that holds. An example of this could be "For all x, P(x)". Naïve set theory, prior to Russell's paradox (van Heijenoort 1967) claimed that every property P(*) defines a set of objects x satisfying P(*). Classical logic together with naïve set theory was proven inconsistent by Russell's paradox. It is an open question whether naïve set theory and fuzzy logic is inconsistent (Behounek and Hanikova 2014), but we will assume it is for the scope of this paper, since it simplifies the exposition. The quantifier "Exists" holds true if the property holds for at least one object. One would think quantifiers like "Ten" could be easily defined by an extension of this principle, but it then we run into the problem of reference. If we define "Ten" to mean exactly 10, the reference should have failed for the jar with ten bolts.

* This paper is published and available in Croatian language at: http://ceciis.foi.hr

It is a point of fact that reference in such cases succeeds and the solution is to model the desired quantifier terms with fuzzy memberships. Let us give an example. As we stated earlier, "There exists an x such that P(x)" is true provided there is an object with the property P(*). We could make the same analogy for "Ten" then it would be that we have to find a truth value for "There are ten x such that P(*)". We could say it is true when there are ten items, and false otherwise, but we could also relax this condition and accept some border cases with some truth, e. g. with a "truthness" of say 0.8 for 9 and 11 items.

The terms analyzed were (i) the precise terms and (ii) the non-precise terms. The precise terms analyzed were "Jedan", "Dva", "Tri", "Četiri", "Pet", "Šest", "Sedam", "Osam", "Devet", "Deset", "Jedanaest","Dvanaest", "Trinaest" (numbers one to thirteen). The non-precise terms were "Jedva išta" (barely anything), "Par" (a couple), "Nekolicina" (few), "Nekoliko" (few), "Brojni" (numerous), "Dosta" (plenty), "Mnogo" (a lot), "Puno" (a lot), "Malo" (few), "Nešto" (some), "Osjetno" (quite a few), "Više" (a number of). As it can be seen we have included several different words but with very similar meaning, so the distinction between their membership functions can be seen as a contribution to the understanding of their semantics. Their translation in English is also tentative at best, as many of them are considered synonyms.

For finding a good represenation of the linguistic quantifiers, we have conducted a survey among 43 students of the University College Algebra asking them to fill in a chart with scores 1-5 representeting how good a term describes a quantity. For example, a score of 4 for "Mnogo" under the column "Quantity: 20" meant that "Mnogo" was 80% appropriate term for describing a quantity of 20 units. The 1-5 scores were subsequently normalized to a 0-1 scale.

A more comprehensive survey, as well as the interpretation of results as type-2 fuzzy sets is planned for further research. Another topis for further research is the use of our approach to facilitate anaphora resolution in South Slavic languages.

2.1 Precise numbering terms

It turns out that our surveys reported such fuzziness even for precise numbers, and we interpolated values to the results to find a function to describe the graph. The most common number terms in Croatian were considered and most of them showed a relatively high precision, along with some fuzziness on the border cases, which increased as the numbers increased. The notable exception was "Deset", or (ten) showing considerably more fuzziness than "Jedanaest" (eleven).

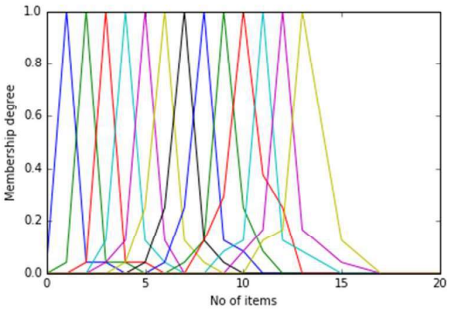


Figure 1. Precise numbering terms

All of these functions can be modelled by a minimal adjustment to the Gaussian function, or, for our needs, with a simple piecewise spike function. Eventual gaps in values are inconsequential, as long as the function is defined for all arguments. Also, the function should be used to assess only integer values.

2.2 Non-Precise numbering terms

The first non-precise quantifier term analyzed was "Nekolicina" (few), which displayed a bell-like shape and could easily be approximated by a Gaussian function. The visible hops on the right-hand side do not matter much for modelling, but the bounds do, so the function returns 0 on 0, 0.7 on 5, and again goes to 0 at 12. The graph is shown on the picture below.

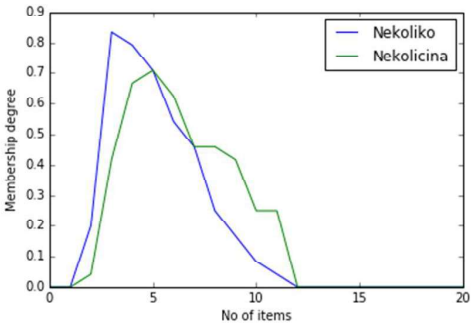


Figure 2. Non-Precise numbering terms

The second nonprecise term we modelled was "Nekoliko" (few). The first part of the function acts similarly to the previous one, but topping off at 0.8 when 3 is reached, and going down more steeply at 12. Notice how the values at 10 differ considerably. The most important difference is the argument for which the functions achieves it maximal value (in the previous case 5, and here 3), which points to the semantic differences in these two terms (both adequately translated into English with the word "few").

This points to the fact (which we shall see soon to hold) that for most quantifier term membership

functions, the argmax, argmin, maximal and minimal values give a very precise definition for a fuzzy membership function. We could approach this issue by using a Gaussian function to model them, but we shall use a piece-wise linear function instead (it is computationally more feasible).

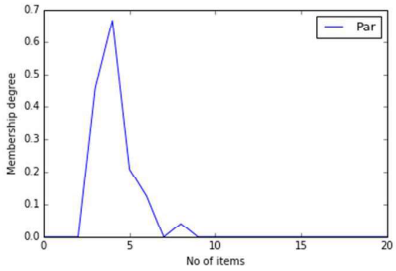


Figure 3. Word "Par" graph

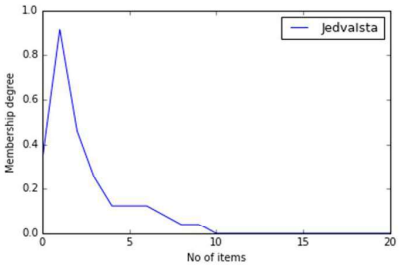


Figure 4. Word "Jedva išta" graph

The word "Par" (a couple), achieves its maximum at 4, with a very steep slope (ending at 6), but in practice it is the same as "Nekoliko", only with the maximum translated at 4, and a steeper curve. "Jedva išta" (barely anything) displays a similar pattern, with a high maximum at 0.95 and argmax at 1.

The next type of membership functions are ReLU-type functions, which are similar to appearance to the ReLU function (f(x)=max(0,x)). They are large quantifier terms, and they are dependent on scale: if the scale goes up to 1000, then at 1000 they will reach the value closest to 1. If on the other hand they are scaled up to 10000, then there the membership functions will be close to 1.

The term "Dosta" (plenty) is the most peculiar, starting the rise at 5, spiking at 10 and plateauing at 10-17, and taking off afterwards. This indicates the problem with the scale limit, and it can be restated to catch the fact that "Dosta" has a spike in mid-range, a max at the end of the range and a plateau between.

"Puno" and "Mnogo" (a lot) are sometimes considered synonyms in south Slavic languages, but they seem to different in semantics, and shown on the images below, in since "Mnogo" was assessed truer of a smaller number of items than "Puno".

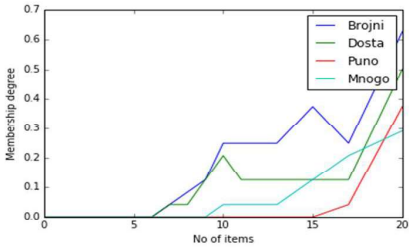


Figure 5. Comparison graph

The step in "Mnogo" is an interesting occurrence but it is inconsequential. The semantic difference is really a matter of nuances in this case, and there is no practical applicability of this, but it is an interesting curiosity nevertheless.

The term "Brojni" (numerous), has a more erratic membership function shown on the image below, but can still be approximated by a ReLU-like function. A detailed exposition of how the relevant ReLU's are defined is given in the next section.

The terms "Malo" (few) and "Nešto" (some) have a spike at 2 and 3, to decline later on. They are similar in behaviour to number terms, only with a prolonged tail as the values grow.

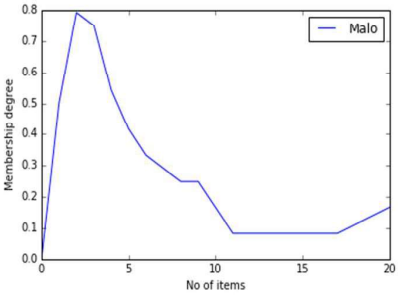


Figure 6. Word "Malo" graph

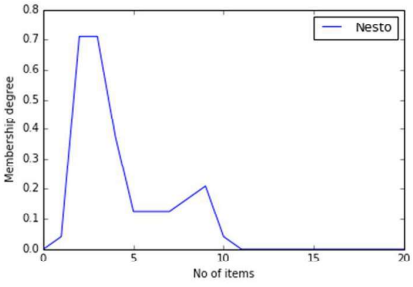


Figure 7. Word "Nešto" graph

The last two term are "Osjetno" (considerably) and "Više" (more), which share a middle spike at 10, and a subsequent decline, followed by a rise to 1 as the

number of counted terms rise. This is in a way surprising, since their semantics is traditionally not considered close, whereas "Više" is considered almost synonymous with "Puno" and "Mnogo". For the purpose of extracting membership functions however, we shall regard "Više" as being a ReLU-like function.

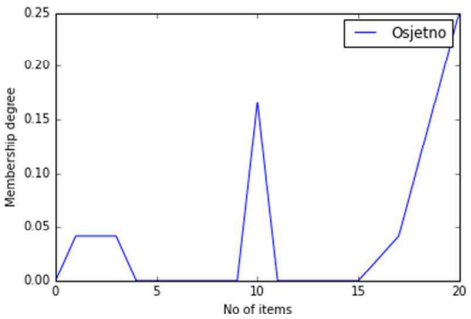


Figure 8. Word "Osjetno" graph

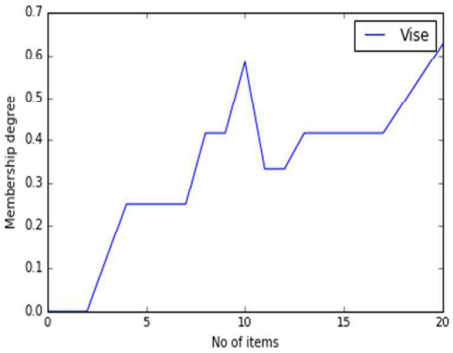


Figure 9. Word "Više" graph

3 Extrapolated functions

We give a table of the extrapolated functions:

Table 1. Extrapolated functions: translations, Non-Zero range

Term	Translation	Non-Zero range
Nekoliko	Few	1-12
Nekolicina	Few	1-12
Par	A couple	2-6
Jedva išta	Barely Anything	0-10
Brojni	Numerous	6-Inf(20)
Dosta	Plenty	6-Inf(20)
Puno	A lot	15-Inf(20)
Mnogo	A lot	9-Inf(20)
Više	A number of	2-Inf(20)
Malo	Few	0-20
Nešto	Few	1-11
Osjetno	Considerably	0-4, 9-11, 15-Inf(20)
[Number N]	Precise number N	(N-1)-(N+2)

Table 2. Extrapolated functions: Maximum Value At, MF type, ReLU kickoff

Term	Maximum Value At (Argument, Value)	MF type	ReLU kickoff
Nekoliko	(3, 0.85)	Spike	--
Nekolicina	(5, 0.7)	Spike	--
Par	(4, 0.65)	Spike	--
Jedva išta	(1, 0.9)	Spike	--
Brojni	(20, 0.65)	ReLU	0.33
Dosta	(20, 0.5)	ReLU	0.33
Puno	(20, 0.38)	ReLU	0.75
Mnogo	(20, 0.3)	ReLU	0.45
Više	(20, 0.6)	ReLU	0.1
Malo	(2, 0.8)	Spike	--
Nešto	(3, 0.7)	Spike	--
Osjetno	(2, 0.05), (10, 0.16), (20, 0.25)	Other	--
[Number N]	(N, 1.0)	Spike	--

To extract the exact version of the membership function needed from the above table, the membership function type column must be consulted first. In the case of ReLU functions, we use the following general form:

$$ReLU(x) = max(0, f(x)) \quad (1)$$

Where f(x) is a linear function calculated through two points after the beginning of the non-zero range. This is a trivial task, but we repeat the process for the reader's convenience (for details see Bronshtein et al. 2007). First the slope is calculated through two points of the non-zero range with slope(f) = (f(x2)-f(x1))/(x2-x1), where the usual denominator provisions are enforced (but even if not, this is not consequential since a different pair of points is chosen and the slope is calculated). After the slope is calculated, by using y-y1 = slope(f)(x2-x1), an explicit representation is obtained. A word of advice: when choosing points for a calculation, due to approximation errors is best to choose the points at which the piece-wise function connects with its other parts. This means that when calculating the slope, the last point for which f(x) = 0 should be one of the points used, and the second should be the end of the range point (which has also the

maximal value in ReLU-like functions). This is even more important in the spike functions, which have a general form:

$$Spike(x) = max(0, (up(x), down(x))) \quad (2)$$

Where the (up, inflection, down) is a shorthand for a two-piece function with the (global) maximum in the middle (this is the inflection point). Two slopes are needed. The first slope is the "up" portion. Its calculation is similar to ReLU's. Two points are used, the left one being the last point for which f(x)=0, which is the first point in the non-zero range from the table above, and for the second point the pair in the column Maximal Value At from the table above should be consulted. For the "down" part, the first point for the slope should be the pair in the column Maximal Value At from the table above, and the second point the first point for which f(x)=0 after the inflection point.

4 Conclusion

In our research we have focused more on smaller number terms, due to the fact that they are effectively bounded by 0. These terms are the spike-type functions. For the sake of completeness, we have also included the large quantifier terms represented by the ReLU-like membership functions, but there is an inherent problem with these terms, namely that they are interpreted relative to the scale offered: In some cases, 20 may be "a lot", and in some cases 100 may not fit "a lot" well. We used a range from 0 to 20, but to provide for this problem we have added a "ReLU kickoff" parameter in the table above, which defines after which percentage of the range the function stops being zero and takes off. This is better than the approach with log scales, as the log scale would require nonlinear function in the ReLU, and yet they would only reduce the problem and not eliminate it completely, since a right-hand bound would still have to be provided.

The membership function for "Osjetno" was left out, as it did not give clear readings (it has a max at only 25%), and it is rather difficult to describe. We feel that to make a useable representation for "Osjetno", a larger dataset and a deeper polynomial fitting algorithm should be used. In this way, more regularities might arise, so we leave this as an open problem for further research. There are of course a number of other terms just as complex as "Osjetno" which still have to be addressed.

We believe our research could be of great use for computational semantics and anaphora resolution in south Slavic languages. First, due to similarities, we believe that for workable engineering applications, the current representations of quantifier terms in Croatian could be used for all the south Slavic languages (Bosnian, Bulgarian, Croatian, Macedonian, Montenegrin, Serbian, Slovene). The application to computational semantics is quite straightforward: the two main segments of computational semantics are the relationships (extracted usually by machine learning) and the quantifiers, which we have partly addressed.

An additional application would be to be able to assess the anaphoric behaviour of quantifier terms, i.e. does a current unspecified quantity refer to a previous unspecified quantity and, if yes, to which one if more than one is mentioned. This can be done by learning how similar are two membership functions given a hypothesized quantity. This can no longer be done with a simple linear fitting, and we leave this for further research.

References

Behounek, L. and Hanikova, Z. (2014). Set Theory and Arithmetic in Fuzzy Logic. In Petr Hajek on Mathematical Fuzzy Logic, ed. F. Montagna, pp. 63-89.

Blackburn, P., de Rijke, M., Venema, Y. (2002). Modal Logic (Cambridge Tracts in Theoretical Computer Science). Cambridge: Cambridge University Press.

Bronshtein, I. N., Semendyayev, K. A., Musiol, G. and Muehlig, H. (2007). Handbook of Mathematics (Fifth Edition). Berlin: Springer.

D'Ambrosio, B. (1989). Qualitative Process Theory Using Linguistic Variables. Berlin: Springer.

Donnellan, K. S. (1966). Reference and Definite Descriptions. The Philosophical Review, vol. 75 No. 3, pp. 281–304.

Donnellan, K. S. (1972). Proper Names and Identifying Descriptions. U D. Davidson i G. Harman (ur.). Semantics of Natural Language.

Frege, G. (1879). Begriffsschrift: eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Halle.

Hajek, P. (1998). Metamathematics of Fuzzy Logic. Amsterdam: Kluwer Academic Press.

van Heijenoort, J. (1967). From Frege to Gödel: A Source Book in Mathematical Logic, 1879-1931, pp. 124-125. Cambridge: Harvard University Press.

Marek, V. W. (2009). Introduction to Mathematics of Satisfiability. Boca Raton, FL: Chapman & Hall/CRC Studies in Informatics Series.

Mrsic, L. and Klepac, G and Kopal R (2017). A New Paradigm in Fraud Detection Modeling Using Predictive Models, Fuzzy Expert Systems, Social Network Analysis, and Unstructured Data, Computational Intelligence Applications in Business Intelligence and Big Data Analytics, Auerbach Publications, pp. 157-194

Peirce, C. S. (1885). "On the Algebra of Logic: A Contribution to the Philosophy of Notation, American Journal of Mathematics, vol. 7, pp. 180–202.

Peters, S. and Westerstahl, D. (2006). Quantifiers in Language and Logic. Oxford: Oxford University Press

Ross, T. (2010). Fuzzy Logic with Engineering Applications. New York: Wiley Press.

Torza. A. (2015). Quantifiers, Quantifiers, Quantifiers: Themes in Logic, Mataphysics and Language. New York: Springer.

Education and Learning Analytics

Acceptance of E-Learning System at Faculty of Technical Sciences

Teodora Lolić, Sonja Ristić, Darko Stefanović, Uglješa Marjanović
University of Novi Sad, Faculty of Technical Sciences
Department of Industrial Engineering and Management
Trg Dositeja Obradovića 6, Novi Sad
{teodora.lolic, sdristic, darkoste, umarjano@uns.ac.rs}

Abstract. *This paper shows a research carried out at the Faculty of Technical Sciences, University of Novi Sad regarding the acceptance of e-learning system by students. The questionnaire used as a data collection tool was created in accordance with the UTAUT[1] model with the aim of determining the acceptance of the implemented system. Within the paper, the obtained research results are analyzed using different statistical methods, after which a final model for measuring the acceptance of the e-learning system was obtained. The results contribute to a better understanding of how to measure the acceptance of e-learning systems, and how to improve its usage.*

Keywords. E-learning system, IS acceptance, UTAUT model, SEM, Moodle

1 Introduction

Nowadays, technology usage as a part of teaching process is necessary. The concept of knowledge is changing. Bearing in mind that each organization is striving to become a learning organization, the use of technology in the process of learning is becoming inevitable (Marjanovic, Delic & Lalic, 2016). Modern education requires transformation of "traditional model of knowledge reproduction" into a model of active knowledge construction, where teachers and students work together to create a knowledge base which needs to be adopted.

Universities and faculties, as institutions of higher education, are also focusing on information and communication technologies and systems, with the aim of competitiveness, efficiency and effectiveness improvement. Technology usage in teaching process varies at different levels – from simple tools like MS PowerPoint and Prezi, to usage of more complex systems for e-learning such as Moodle. Usually, focus is on creating learning content which is necessary for students learning process.

E-learning is defined as a learning which is facilitated and supported through the usage of information and communication technologies (Jenkins & Nunamarker, 2003). Accordingly, e-learning concept includes utilization of information and communication technologies (Internet, computer, mobile phone and multimedia) as a way to support teaching and learning.

Continuous search to provide the acceptance of technology by clients is current challenge for management (Schwartz & Chin, 2007). The research of technology acceptance in the field of information systems (ISs) and information technologies is nowadays considered as one of the basic research topics (Venkatesh et al., 2003). Various techniques are used for acceptance testing of different systems and technologies in many different contexts. Therefore, diverse perspectives of stakeholders, technologies and contexts, theories and research methods are discussed (Williams et al., 2009). This situation has led to confusion between researchers, since they are commonly forced to choose characteristics and components from the wide spectrum of, usually competitive, models and theories. As a reply to this challenge Venkatesh et al., (2003). have developed unique model which sums up alternative views of user and acceptance of innovations – Unified Theory of Acceptance and Use of Technology (UTAUT).

In this paper are analyzed different aspects and possibilities of measuring acceptance and usage of systems for e-learning with a focus on post implementation phase. After that, it summarizes the results about how much these approaches are relevant for measuring acceptance and usage of systems for e-learning.

In this research the acceptance of e-learning system by the students at the Faculty of Technical Sciences, University of Novi Sad has been examined. Different stakeholders are interested in e-learning system's acceptance, like: creators and managers of information system, e-learning process managers, professors who use e-learning system for more efficient and easier

[1]Unified Theory of Acceptance and Use of Technology

communication with their students at one place. Surely, the main goal for all stakeholders is to determine whether the implemented system is positively accepted by its primary users – students. For the research purpose UTAUT model has been used.

2 Background and related work

Acceptance of an IS one of the arguable questions that draws attention of researchers in the field of IS. Problem is even more complex because acceptance is multidimensional concept which can be graded from different levels, such as technical, individual, group or organizational.

E-learning systems are gaining importance and are increasingly being used in everyday teaching activities. In order for teachers to have an insight into the state of contribution of this system, there are various models that allow the measurement of the acceptance of an e-learning system. Many researchers are focusing on this issue in past decades (Weerasinghe, 2017), as well as today (Ayele & Birhanie, 2018; Yakubu & Dasuki, 2018).

2.1 E-learning system

An e-learning system is a type of IS based on Internet technology that provides training of the learner in an independent and flexible way (Wang et al., 2007; Lee & Lee, 2008), supporting teaching and learning processes (Hassanzadeh et al., 2012). Users or e-learners access the system through Internet or intranet portals in order to acquire information, knowledge and skills (Chen, 2012). During the learning process, users can interact with other participants, such as the instructor or other users. A lot of web based learning and training programs are developed that make learning self-paced. The e-learning process is usually designed in the way to enable users to control learning elements. These two features, self-pacing and control over learning, are providing new opportunities for the individuals and for the companies. In the process of design and delivery of e-learning and training programs, enterprises have to consider both effectiveness and acceptance of e-learning systems.

2.2 Unified Theory of Acceptance and Use of Technology

In the field of ISs, many researchers have concluded that information technologies are insufficiently used in organizations, which leads to huge economic loss in their business. As a result, a lot of theories and models about acceptance of technology are developed or used for studying the acceptation of information systems. These models include: Theory of rational action (Fishbein et al., 1977), Model of technology acceptance (Davis, 1989) and extended Technology Acceptance Model (TAM) (Venkatesh & Davis, 2000), UTAUT (Venkatesh et al., 2003) and many others.

UTAUT suggests that there are three constructs of the main determinant of behavioral intention towards using information technologies. Those three constructs are: Performance Expectancy (PE), Effort Expectancy (EE) and Social Influence (SI). They have direct impact on construct Behavioral Intention (BI) which, in the end together with Facilitating Conditions (FC), influences on construct System Usage (SU).

2.3 Hypotheses development

In this section, the rationale for each of the proposed hypotheses, stating connection between constructs from the proposed e-learning acceptance model is explained. The hypotheses about the relationships in the UTAUT model, with corresponding discussions, are presented below.

H1: Performance Expectancy has positive effect on the Behavioral Intention.

Maldonado et al., (2011) have found strong and positive effect while observing educational portal for e-learning in South America. Abdekhoda, Dehnad & Gavani (2016) confirmed that statistically significant relationship between PE and BI exists. They examined this relationship on an example of e-learning system at the University of Tabriz, Faculty of Medicine. Chaka & Govender (2017) have also confirmed significance of this connection. Fourth group of authors, in their research of student behavior while using e-learning systems, showed that the relationship between the two, previously mentioned, factors is strong and significant. Wang, Wu & Wang (2009) were testing the acceptance of Mobile Learning (m-learning) Systems, and as the result they have represented that the connection between PE and BI is statistically significant and strong. In his research about effects of Technological expectations on BI, Chen (2012) has shown the statistical significance of this relationship.

H2: Effort Expectancy has positive effect on the Behavioral Intention.

Previous researches whose topic was significance of relationship between these two factors have got the results which indicate that significance of the relationship exists. Maldonado et al., (2011) have shown, as a part of the Motivation for e-learning factor, that relationship between these two factors is strong and statistically significant. Some other authors (Abdekhoda, Dehnad & Gavani, 2016; Chaka & Govender, 2017) have also confirmed that the relationship between EE and BI is strong and significant. On the other side, Masa'deh et al., (2016) have found that previously mentioned relationship is weak and that it does not have statistical significance. However, the relationship between EE and BI has also been confirmed as strong and significant in the research which was conducted by Wang, Wu & Wang (2009). While searching for the answer about the

impact factors of Technological expectations on BI, Chen (2012) confirms significance of this relationship, too.

H3: Social Influence has positive effect on the Behavioral Intention.

Almost all previous researches have shown that connection between these two factors exists (Maldonado et al., 2011; Abdekhoda, Dehnad & Gavani, 2016; Ra'ed Masa'deh et al., 2016). Maldonado et al., (2011) have shown that the connection between SI and Behavioral Intention is very strong. Abdekhoda, Dehnad & Gavani (2016) in their research claim that if SI increases by one, BI will increase by 24%. Significance of the relationship between SI and BI is shown as strong and important (Chaka & Govender, 2017). While researching usage and acceptance of technology with m-learning concept, authors (Wang, Wu & Wang, 2009) have concluded that relationship between these two factors exists and it is statistically significant and strong. The fact that SI has positive and statistically significant impact on BI was also shown by Chen (2012).

H4: Behavioral Intention has positive effect on the System Usage.

All relevant previous researches have shown statistical significance (Maldonado et al., 2011; Chaka & Govender, 2017) and strong relationship between BI and SU (Abdekhoda, Dehnad & Gavani, 2016), which is natural – if someone has intention to use the system, he will.

H5: Facilitating Conditions have positive effect on the System Usage.

On one hand, some of the previous researchers have found that the connection between FC and SU is small, but statistically significant (Masa'deh et al., 2016). On the other hand, a few researchers have found that this relationship is at the borderline of statistical significance (Maldonado et al., 2011; Abdekhoda, Dehnad & Gavani, 2016).

3 Materials and Methods

Within this section, materials and methods that have been used for the proposed research are shown.

3. 1 Measures

The indicators and constructs[2] of the conceptual model have been determined based on previous research on IS success. The measures of constructs that have been used to analyze the success of various types of IS, in previous studies, which have been adopted in this study are listed in Table 1.

[2]In the methodology of Structural modelling notions construct and factor are used interchangeable as dimensions to be measured (Pallant, 2010).

Table 1. Construct measures

Construct	Indicator	Ref.*
Performance Expectancy	PE1. Usage in learning	1, 2
	PE2. Faster obligations fulfillment	1, 2
	PE3. Increase in work productivity	1, 2
	PE4. Easier learning	2
	PE5. Better learning performance	2
Effort Expectancy	EE1. System usage – clear and understandable	1, 2
	EE2. Fast system understanding	1, 2
	EE3. Simplicity of using	1, 2
	EE4. Learning to handle the system	1, 2
	EE5. System responsiveness	2
Social Influence	SI1. Effect of people that have an influence on student	1, 2
	SI2. People that student care about	1, 2
	SI3. Older Faculty colleagues	1, 2
	SI4. Influence of the Faculty	1, 2
	SI5. Other colleagues	2
Facilitating Conditions	FC1. Owning a resources	2
	FC2. Owning a knowledge	2
	FC3. Compatibleness with other systems	2
	FC4. Fitting into way of working	2
	FC5. Instructions for using	2
Behavioral Intention	BI1. Intention to use system in the future	1, 2
	BI2. Prediction of future usage	1, 2
	BI3. Planning to use the system in the future	1, 2
System Usage	SU1-11: Forum, Chat, Learning material, Lesson Video resources, Quiz, Assignments, Marking of completed activity, Messages, Gradebook, Participants directory	3, 4

*References: 1 – Wang, Wu & Wang (2009); 2 – Venkatesh et al., (2003); 3 – Persico, Manca & Pozzi (2014); Lolić, (2018) p.80.

3.2 Sample and data collection procedure

Data which has been used for this research was collected from the students at the Faculty of Technical

Sciences (Department of Industrial Engineering and Management), University of Novi Sad. These students use e-learning system Moodle eLLab.

The researched system is open source code software – Moodle which has been used for development of e-learning system (Romero, Ventura & Garcia, 2008). Moodle owns flexible set of modules that can be organized by using activities and resources. By developing different activities and resources, it also supports creation of different types of statistical and interactive materials. This software has tools for collaboration, which makes it a collaboration system, too. Moodle keeps and records detailed notes about all activities that participants take (Rice, 2006). Recording these notes makes possible keeping a track of materials and tools used by participants (Romero, Ventura & Garcia, 2008). This system reports every activity that participants do and uses it for navigational purposes. Besides that, system has a built-in record log, which is also accessible, as well as information about activities of certain participant and his performance (Martín-Blas & Serrano-Fernández, 2009).

Data collecting process lasted one month. Students' opinions about using the e-learning system Moodle eLLab were collected online with help of the questionnaire (Lolić, 2018. p. 78-80) which has been made based on the theory (Table 1) that was distributed through Internet tool as instrument for data collection. For research implementation SurveyMonkey[3] has been used.

For the needs of data collection process, researchers used e-mail addresses which were available in the database of Faculty of Technical Sciences e-learning system Moodle eLLab. Participants were contacted and asked to give their opinion about usage of the e-learning system. Participation in this research was voluntarily and none of the participants, in any way, was forced to respond.

Participation invitation in this research was sent to 2017 e-mail addresses of students, while valid replies were collected from 796 of them. Response rate is 37,78%. Initial data screening showed that 50 cases had very low standard deviations (below 0.2). Thus, to minimize the non-engage bias, they were removed from further analysis (Nunnally & Bernstein, 1994). The final number of validly filled questionnaires that was used in analysis was 746.

In this research, 387 of the respondents were female (51.9%), while 359 of them were male (48.1%). Respondents are categorized in four categories according their age: respondents younger than 21 year, respondents at age between 21 and 24, respondents at age between 25 and 30 and respondents older than 30 years. Most of the respondents are found to be in the second category – between 21 and 24 years (46.8%). Degree of computer literacy was divided in the three categories according to years of experience in using e-learning system: professional user (more than 3 years), middle user (1-3 years) and beginner (less than 1 year). Respondents mostly declared themselves as a middle user, which means that most of them belong to the second category – 489 (65.5 %).

4 Results

After the data collection process was done, next step was to analyze the results. That was done with the IBM SPSS Amos (version 20) tool, which is explained in the following part of the paper.

4.1 Statistical methods used for analysis

For the need of showing the results, researchers used the analysis of the basic items characteristics (descriptive statistical analysis) which represents a group of methods that describe results and have a goal to group, arrange and show the statistical data, as well as to determinate basic indicators of statistical series (Marjanovic, 2014).

Exploratory Factor Analysis (EFA) was used to identify the structure of factors by examination of correlation matrices.

After that, Confirmatory Factor Analysis (CFA) is used to statistically confirm a definition of dimensions by manifest variables. This analysis is used as a tool to test the conceptual model and hypothesis. Based on the CFA, a Structural Equation Modeling – SEM (Tabachnick & Fidell, 2007) has been conducted.

4.2 Identifying the factor structure

On courses supported by the e-learning system, students have the need to use the system in order to get all the materials necessary for passing the exam. For this reason, the dimension Voluntariness of Use from the UTAUT model is not being observed because its role is insignificant for this research. Following that, we have researched the relationships between other constructs of UTAUT model.

A thirty indicator instrument is tested on the bases of collected data. With the aim to improve validity of model by using EFA, next steps were conducted:

1. Applying the Keizer-Guttmans rules or "validity of variance higher than 1";
2. Screen Plot – visual representation of variance value;
3. Eliminating the variables that made other factors, not important for this research;
4. Suitability of Chi-square index and Usage of suitability index (Root-mean-square error of approximation – RMSEA and Comparative fit index – CFI) (Tabachnick & Fidell, 2007).

[3] www.surveymonkey.net

Factor analysis was done in an iterative procedure until an adequate model and factor structure that satisfies all the criteria shown in the previous five steps was made. This model explained 72.74% of variance, and the value of variance and its percent explanations for factors were in a range of 6.75% and 16.82%.

At the end of the factor analysis process, nine indicators were eliminated EE5, SI4, FC5, SU1-SU5 and SU9 because they have built other constructs which are not important for this research.

Final model consisted of 6 factors and 24 indicators/variables.

4.3 Reliability and validity assessment

Reliability of measurement instruments is determinated by calculating a coefficient of Cronbach's alpha for each dimension. Calculated values for each of the dimensions are: PE = 0.879; EE = 0.866; SI = 0.760; FC = 0.771; BI = 0.957 and SU = 0.876. All these values satisfy a minimal criteria (0.60 or higher) accordingly to (Hair et al. 2009).

Besides from that, reliability and convergent validity of factors is estimated with a usage of Composite Reliability (CR) and Average Variance Extracted (AVE). Results are demonstrated in Table 2. All coefficients of Cronbach's alpha and CR values satisfy minimal criteria value – 0.70, which was suggested in (Hair et al., 2009).

Average variations are above recommended 0.50 level (Straub, 2012), which means that more than half of variations observed as indicators were calculated with their factors hypothesis. CR was higher than AVE for each factor. Because of that, we can conclude that all factors in the model of measurement have adequate convergent validity.

Discriminant value can be estimated by testing AVE, MSV and ASV. By Hair et al., (2009), if MSV is higher than AVE and ASV, it leads to discriminant importance. Regarding convergence (Table 2), all factors were satisfying. Summarized, model of measurement had adequate reliability, convergent validity and discriminant validity.

4.4 Structural model

This research has had a goal to examine relationships between constructs suggested by model and to examine suggested model with collected data.

Model consisting of six constructs had adequate suitability indexes. Values of suitability indexes are shown in the Table 3 which represents that all the values are in the acceptable range, which further indicates a good fitting of models.

Table 3. Suitability indexes for CFA and SEM

Model	$\chi 2$/df	NFI	CFI	RMSEA
Measurement model	4.617	0.909	0.927	0.070
Structural model	4.752	0.905	0.923	0.071
Recommended value	less, the better	>0.90	>0.95	<0.08

Path coefficients, p-values, z-results and explanation of variance are shown in the Fig. 1. All obtained values for path coefficients are above recommended values of 0.20 as suggested in (Chin, 1998) with the exception of the paths between assumed dimensions of BI and SU, and FC and SU.

Relationship between PE and BI is statistically significant and positive (path coefficient = 0.306; t = 6.834).

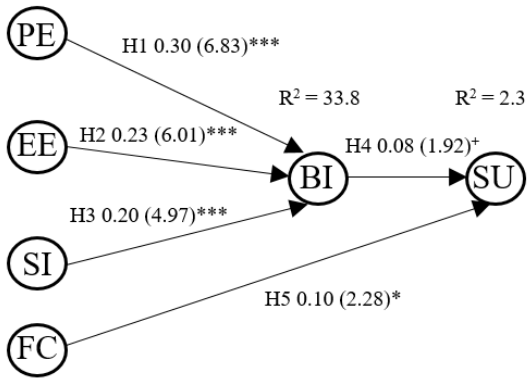


Figure 1. Structural model for examination of e-learning system *Moodle* eLLab acceptance

Table 2. Reliability, convergent validity and construct correlation

Construct	Middle value	SD	α	CR	AVE	MSV	ASV	SI	SU	PE	EE	BI	FC
SI	3.25	.84	.760	.844	.660	.234	.119	.812[a]					
SU	3.18	.90	.876	.840	.514	.078	.037	.247	.717[a]				
PE	4.13	.75	.879	.882	.603	.266	.194	.484	.280	.777[a]			
EE	4.55	.56	.866	.908	.714	.323	.161	.213	.116	.495	.845[a]		
BI	4.28	.90	.957	.957	.880	.266	.161	.395	.118	.516	.421	.938[a]	
FC	4.24	.66	.771	.819	.535	.323	.154	.310	.127	.382	.568	.437	.731[a]

Note: SD – Standard Deviation, MSV - Maximum Shared Variance, ASV - Average Shared Variance.
[a] Indicates the square root of AVE construct

This connection is the strongest and statistically most significant. EE has statistically significant effect on BI and relationship between these two dimensions is distinctively strong and positive (path coefficient = 0.233; t = 6.009). Statistically significant relationship between SI and BI is also proven (path coefficient = 0.20; t = 4.97). Dimension FC has statistically significant effect on factor SU, but with low statistical significance. As for relationship between BI and SU, we can conclude that this is the weakest relationship and it is at the borderline of statistical significance (path coefficient = 0.08; t = 1.92).

Together, both assumed dimensions explain 36.1% of variance in structural model. Assumed dimension BI is explained in 33.8% of dimensions PE, EE and SI. On the other side, dimensions FC and BI shows significantly lower effect (influence) on assumed dimension SU, with only 2.3%.

Results of SEM show that all five hypothesis are accepted, but also that hypothesis H4 is barely accepted since its value is a little below the borderline of statistical significance ($p < 0.1$).

5 Discussion

Measuring instrument that is used in this paper is tested with EFA. Results of this analysis indicate the existence of five strong dimensions of acceptance and usage of information system. All of these dimensions were taken from UTAUT model of acceptance and usage of technology.

Results of the CFA have shown a high degree of reliability and validity of relationships between dimensions in measurement model. A strong connection within all elements in the model is established, with one exception – relationship between two dimensions, BI and SU.

All five dimensions represent stable elements in model of acceptance of e-learning system Moodle eLLab, while the relationship between these dimensions is interdependent one. These results are in accordance with the original creators of technology acceptance and usage model, who also claimed that the relationship among these dimensions is interdependent (Venkatesh et al., 2009; Wang, Wu & Wang, 2009).

The results of this research purpose the acceptance of the e-learning system Moodle eLLab at the Faculty of Technical Sciences by the students. Results showed the existence of high expectations of system usage by the students, as well as their will and desire to be a user of e-learning system Moodle eLLab.

During the analysis, mutual influence of four independent and two dependent variables from conceptual model was observed. Connections of independent factors PE, EE and SI towards dependent factor BI were analyzed, as well as relationships between BI and SU and between FC and SU. Obtained results indicate that observed dimensions are reliable and valid acceptability measures in context of e-learning system. Analysis of data obtained in this research has strongly supported four out of five defined hypotheses.

Higher intention to use a system leads to higher usage of it, and consequently to higher total usage of the system by students.

Designers of e-learning systems (in this case professors) should pay a special attention to expectations of the system, when it comes to system usage. Students should be conscious that the usage of e-learning system will be useful in their learning process, and that it will upgrade their performance as well as make their learning easier.

Furthermore, it will make learning process much faster. Cognition like this will result in an increase of their intention to use the system.

Summarized results are shown in Table 4.

Table 4. Results of testing on the defined hypothesis

Hyp.	Relationship			Path coefficient	t-value	Result
H1	PE	→	BI	0.306***	6.834	Accepted
H2	EE	→	BI	0.233***	6.009	Accepted
H3	SI	→	BI	0.200***	4.970	Accepted
H4	BI	→	SU	0.082^{+}	1.918	Marginally accepted
H5	FC	→	SU	0.105*	2.280	Accepted

Note: *** $p < 0.001$, *$p < 0.05$, $^{+}p < 0.1$

Hypothesis H1: PE has positive effect on BI is confirmed. Results of structured model indicate that PE is the most important component from the aspect of measurement of technology acceptance and usage. Through the direct influence on BI, PE also has an influence on SU.

EE, as second important component that has an effect on BI, proved itself as important in a way that if students expect that their usage of e-learning system will be clear and understandable, that they will not have to put a lot of effort and time to learn how to use the system, their intention to use e-learning system will be higher. Therefore, hypothesis H2: EE has positive effect on BI, is confirmed.

In the most cases, students are imposed to use the e-learning system because it offers them the materials which are necessary for them to pass the exam. In general, this shows that Faculty encourages, and of course, supports usage of e-learning system. Besides the Faculty initiative, influence of the people that students care about and influence of the Faculty colleagues are segments of the SI factor which, if is strong, leads to higher Behavioral Intention for e-learning. This conclusion resulted in the acceptation of the hypothesis H3: SI has positive effect on BI.

If we look at the conditions that have to exist so that students can use the system, we can claim that their existence will result with usage in higher percent. By conditions, we mean having necessary knowledge and resources and that the e-learning system is compatible

with other systems that students already use. Based on that, hypothesis H5: FC have positive effect on SU, is confirmed, but with the statistical significance at the borderline. Very important data, collected with the usage of descriptive statistics over the factor FC is that students rated every item of this factor with the mark 5. In other words, this means that all necessary conditions for system usage exist, but it does not have high influence on the final system usage.

The most important hypothesis H4: BI has positive effect on SU that indicate the relationship between BI and SU, has shown itself as a borderline statistically significant ($p < 0.1$), which is the result that is not in compliance with the results of other researches that examined this connection. This way, the weakness of this model is presented.

In order to find the answer to this question, descriptive statistics was done over the SU factor, with the aim to find out which items are mostly used. In this respect, items Quiz, Lessons, Marking of completed activities, Gradebook and Participants directory were proved as important for research. Firstly, conclusion is made that Lessons and Gradebook are items which creators of e-learning system expect from students to use, and from that aspect students' intention to use this system does not have important influence od final system usage since they have to use the system anyway.

On the other side, from the descriptive statistics we can see that the items Quiz, Marking of completed activities and Participants directory are mostly answered with the mark 1, which means that students rarely use these items, almost never. If we look at the real situation more closely, reason why the hypothesis H4 is at the edge of acceptance is obvious. Hypothesis H5 shows that even though resources for usage, as well as all the other conditions exist, that does not necessarily mean that students will use the e-learning system. Likewise, even though the results have shown that students mostly have a high positive intention to use system – over 70% of them, that fact does not insure us that they will actually use it. These conclusions lead to answer on the question why the relationship between BI and SU is weak. In most of the teaching courses we cannot find Quiz nor Marking of completed activities, or Participants directory. If these items do not exist within the courses which are part of e-learning system, students surely cannot use them, even if they have an intention to.

6 Conclusion

In this paper, research about the acceptance and usage of the e-learning system Moodle eLLab which is implemented at the Faculty of Technical Sciences, University of Novi Sad is done. For this purpose, UTAUT model was used.

Model that has been tested in this paper can be used to estimate the acceptance and usage of e-learning system and its influence on learning performance from the students' perspective. For example, developmental teams of e-learning systems, mostly professors, should use the simplicity of usage and easiness of understanding this system with the aim to increase the satisfaction of students and system usage. This estimation will enable Faculty to collect the feedbacks on efficiency of implemented information system. IT managers, whose job is to develop and use the e-learning system at the Faculty, can use the model to successfully undertake corrective measures for its advancement.

There are many different ways to make learning process better through using ICT in it (Stevanov et al. 2017; Lolić, 2018. p.58). On the basis of established relationship in the model, Faculty can estimate on which dimensions to pay attention with the aim to advance system acceptance and usage by the students. Also, if professors want to improve students' behavioral intention, they have to pay attention to what they offer to students when they offer them this system as an option for use.

Presented research can help in determination of improvement or degradation in the process of e-learning system implementation. Changes could be made in a few teaching courses, and the next step would be to analyze previous and new state, on both courses that had some changes and the ones that have not had it. This would be done with the aim to determinate the result of changes made over system.

Acknowledgments

This article has been produced as part of a research project: no. 47028 "Advancing Serbia's competitiveness in the EU accession process" supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia for the period 2011th-2018th year.

References

Abdekhoda, M., Dehnad, A., & Gavgani, S. J. (2016). Factors influencing the adoption of E-learning in Tabriz University of Medical Sciences. *Medical Journal of the Islamic Republic of Iran*, 30 (1).

Ayele, A., Birhanie, W. (2018). Acceptance and use of e-learning systems: the case of teachers in technology institutes of Ethiopian Universities, *Applied Informatics*, 5 (1), p. 1.

Chaka, J. G., & Govender, I. (2017). Students' perceptions and readiness towards mobile learning in colleges of education: a Nigerian perspective. *South African Journal of Education*, 37 (1).

Chen, H. J. (2012). Clarifying the empirical connection of new entrants' e-learning systems

use to their job adaptation and their use patterns under the collective–individual training environment. *Computer Education*, 58(1).

Chin, W. W. (1998). Issues and Opinion on Structural Equation Modeling. MIS Quarterly. *MIS Quarterly & The Society for Information Management*, 22 (1), 1-4.

Fishbein, M., & Ajzen, I. (1977). *Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research*, Addison-Wesley, Reading, MA.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate Data Analysis* (7th edition). Prentice Hall, p. 816.

Hassanzadeh, A., Kanaani, F., & Elahi, S. (2012). A model for measuring e-learning systems success in universities. *Expert Systems with Applications*, 39 (12), 10959–10966.

Jenkins, M., & Hanson, J. (2003) *E-learning Series: A Guide for Senior Managers, Learning and Teaching Support Network (LSTN).* Generic Centre, United Kingdom.

Lee, J.-K., & Lee, W.-K. (2009). The relationship of e-Learner's self-regulatory efficacy and perception of e-Learning environmental quality, *Comput. Human Behavior*, 24 (1), 32–47.

Lee, S.H. et al. (2005). Technical Trend of Location-Based Service. *Electronics and Telecommunications Trends*, 20 (3), 33-41.

Lolić, T. (2018). *Electronic learning – acceptance of the system and its contribution to the learning process* (In Serbian). Zadužbina Andrejević

Maldonado, U. P., Khan, G. F., Moon, J., & Rho, J. J. (2011). E-learning motivation and educational portal acceptance in developing countries. *Information & Knowledge Management.* 35 (1).

Marjanović, U. (2014). *The model of the collaboration system development and its impact on the organizational performance of the company* (In Serbian). PhD thesis, University of Novi Sad.

Marjanovic, U., Delic, M. & Lalic, B., (2016). Developing a model to assess the success of e-learning systems: evidence from a manufacturing company in transitional economy. *Inf. Syst. E-Bus. Manage.* 14 (2), 253-272.

Martín-Blas, T., & Serrano-Fernández, A. (2009). The role of new technologies in the learning process: Moodle as a teaching tool in Physics. *Computer Education*, 52 (1), 35-44.

Masa'deh, R., Tarhini, A., Mohammed, A. B., &Maqableh, M. (2016). Modeling Factors Affecting Student's Usage Behavior of E-Learning Systems in Lebanon. *International Journal of Business and Management*. 11(2).

Nunnally, J., & Bernstein, H. (1994). Psychometric theory. New York: McGraw-Hill.

Pallant, J. (2010). *SPSS Survival Manual 4th edition*.

Persico, D., Manca, S., & Pozzi, F. (2014). Adapting the Technology Acceptance Model to Evaluate the Innovative Potential of E-Learning Systems. *Computers in Human Behavior 30*. Elsevier Ltd: 614–22. doi:10.1016/j.chb.2013.07.045.

Romero, C., Ventura, S., & García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computer Education*. 51 (1), 368–384.

Schwartz, A., & Chin, W. (2007). Looking forward: toward an understanding of the nature and definition of IT acceptance. *Journal of the Association for Information Systems*. 8 (4).

Stevanov, B., Stefanovic, D., Anderla, A., Sladojevic, S., & Tasic, N. (2017). New Approach to Information System Engineering Study Program to Meet Industry Expectations. *International Journal of Engineering Education*. 33 (4).

Straub, D. (2012). Does MIS Have Native Theories? MIS Quarterly, *MIS Quarterly & The Society for Information Management*, 36 (2), 3–7.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics.* Boston: Pearson Education, Inc., p. 980.

Venkatesh, V., & Davis, F. D. (2000). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Manage. Sci*. 46 (2), 186.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: toward a unified view. *MIS Quarterly*, 27 (3), 425-478.

Wang, Y. S., Wu, M. C., & Wang, H. Y. (2009). Investigating the Determinants and Age and Gender Differences in the Acceptance of Mobile Learning. *British Journal of Educational Technology*, 40(1), 92-118.

Weerasinghe, S. (2017). Technology Acceptance Model in the Domains of LIS and Education: A Review of Selected Literature, *Library Philosophy & Practice*, pp. 1-26.

Williams, M. D., Dwivedi, Y.K., Lal, B., & Schwarz, A. (2009). Contemporary trends and issues in IT adoption and diffusion research. *Journal of Information Technology*. 24 (1), 1-10.

Yakubu, M., Dasuki, S. (2018). Assessing elearning systems success in Nigeria: An application of the DeLone and McLean Information Systems Success Model, *Journal Of Information Technology Education*, 17(1), p. 182.

Employment of Business Informatics Graduates: Preliminary Results

Mirjana Pejić Bach
University of Zagreb
Faculty of Economics and Business Zagreb
Trg J. F. Kennedyja, 10000 Zagreb
mpejic@efzg.hr

Dario Šebalj
Josip Juraj Strossmayer University of Osijek
Faculty of Economics in Osijek
Trg Ljudevita Gaja, 31000 Osijek
dsebalj@efos.hr

Daniela Garbin Praničević
University of Split
Faculty of Economics, Business and Tourism
Cvite Fiskovića 5, 21000 Split
daniela@efst.hr

Igor Pihir
University of Zagreb
Faculty of Organization and Informatics
Pavlinska 2, 42000 Varaždin
ipihir@foi.hr

Abstract*. Business informatics graduate studies aims at developing both information and business related skills, such as business process management and business intelligence, which are nowadays crucial for the competitiveness of modern companies. Goal of the paper is to discuss the employment trends and opportunities of students that graduate at one of the Business informatics graduate studies taught at the Croatian economic and organizational faculties: Faculty of Economics in Osijek, Split, Zagreb, as well as Faculty of Organization and Informatics in Varaždin. In addition, the abovementioned graduate study programmes are discusses in relation to European e-Competence Framework.*

Keywords. Business Informatics, students, employment, ICT role profiles, European e-Competence Framework

1 Introduction

The rising digitalization and complexity of modern business has led to a demand for highly skilled graduates with different ICT skills and knowledge, aimed to support the economic activities. The named imply the imperative of connecting the ICT and business in tactical, operational and strategic decisions..

Numerous universities all over the world offer study programs specialized for the field of ICT in (and for) business, with excellent training and education respectively. In Croatia, there are many departments and study programs focused on both business and informatics field. In our research, we focus to the study programmes at the largest Croatian faculties: Faculties of Economics in Osijek (EFOS), Split (EFST) and Zagreb (EFZG), as well as Faculty of Organization and Informatics in Varaždin (FOI). Particularly, we focus to the following graduate programmes: *Business Informatics* at EFOS, *Information management* at EFST, *Managerial informatics* at EFZG, and at FOI. However, at FOI there are four graduate studies in Informatics: *Information and Software Engineering* (IPI), *Business Systems Organization* (OPS), *Databases and Knowledge Bases* (BPBZ) and *Informatics in Education* (IO). For the sake of simplicity, we shall refer to above-mentioned graduate studies under one generic name as *Business Informatics graduate studies*.

Numerous research indicate that public and professional ICT community still lacks knowledge regarding Business Informatics professionals' profiles, and their employment opportunities (e.g. Jelić, 2012). In this research, the authors intend to reduce the named gap and consequently present the ICT professional role profiles as defined within European e-Competence Framework (e-CF, 2018). Moreover, we aim to explore the characteristics of employment of graduated students of business informatics study programs offered on above-mentioned four Faculties in Croatia.

The structure of the paper is as follows. The first part presents in detail the ICT roles profiles, and argues the importance of the European ICT skills transparency. The second part covers the research methodology while the forth part reveals the research findings. Discussion and conclusions are enclosed on the paper end.

2 ICT professional role profiles

Being ICT professional indicates the possession of a respectable core body of knowledge, including in-depth ICT knowledge and practical ICT skills, but also business related skills, such as project management and human resources management (Marius, 2014). New

competencies are also constantly emerging as relevant, that are needed in order to address the challenges of ICT driven innovation. E-learning is more and more used to support the development of these competencies (Gyamfi et al., 2017; Leväsluoto et al., 2016). Noted competencies are considered as the ability to perform a work role to a defined standard with reference to real working environments (IFAC Education Committee, 2003). In professional literature, the set of work-related personal attributes, knowledge, and skills are also considered as competencies (Wang & Tsai, 2014) or job competencies (Lowry & Flohr, 2005; Sisson & Adams, 2013). Many researches focus on the skills and abilities that ICT professionals need to know, and will even need more in the future (e.g. Wessels, 2005).

On the contextual level, the ICT profiles origin from the need to achieve the respective graduate employability, which can be further perceived as a set of achievements, moreover skills, understood as personal attributes – "that makes graduates more likely to gain employment and be successful in their chosen occupations, which benefits themselves, the workforce, the community and the economy" (Yorke, 2004, p. 410).

However, competencies taught by universities may differ substantially from the actual competences that businesses need (Chapman & Lovell, 2006: Morgado et al., 2014). Accordingly, the universities and other related institutions need to verify the effectiveness of current strategies used for enhancing employability skills (Andrews & Russell, 2012) and to face the fact that both environment (Guthrie et al., 2008) and students (Suša, 2014) are rapidly changing.

To better explore the labour market needs, it seems as fruitfully to extend the research content toward skills and competences that the European Commission particularly points out as important and crucial for creating future jobs such as "EntreComp: The Entrepreneurship Competence Framework" (CWA, 2018). The stated framework significantly support the development of the European ICT Professional Role Profiles that contribute to overall transparency and convergence of the European ICT skills landscape. The competences of the European e-Competence Framework have been considered as "the main component of profile descriptions" (CWA, 2018, p. 2) and is accordingly incorporated in the ICT Professional Role Profiles formation.

Consequently, the group of 30 ICT professional role profiles has been developed, that include: (i) a generic set of typical roles performed by ICT professionals, covering the full ICT business process in organizations, and (ii) a common European reference language for developing and managing ICT professionals' needs in order to support mutual understanding particularly between HR and ICT departments. Onward, these ICT roles provide the usefuls structure for identifying the main and supporting activities in context of organisational digital strategy provision. In summary, the named ICT professional roles are useful as the: (i) starting point for the flexible creation of more specific ICT profiles, based upon organizational roles of individual job descriptions from wider variety of contexts, and (ii) tool for individuals and organizations who aims to develop and teach skills in line with the European e-Competence Framework standards.

Furthermore, the ICT professional role profiles „are not intended to represent a rigid standard, but are built as a foundation for the flexible creation of more context-specific profiles in a broad variety of areas" (CWA, 2018, p.2).

The roles refer to seven business groups supported by ICT such as process improvement, business general, technical, design, developments, service and operation, and support, and are categorised as following:

1. *standard role* (CWA, 2018, p.35) including: Account Manager Role, Business Analyst Role, Business Information Manager Role, Chief Information Officer Role, Data Administrator Role, Developer Role, Digital Media Specialist Role, Enterprise Architect Role, Digital Consultant Role, ICT Operations Manager Role, Cyber Security Manager Role, Cyber Security Specialist Role, Digital Educator Role, Network Specialist Role, Project Manager Role, Quality Assurance Manager Role, Service Support Role, Service Manager Role, Systems Administrator Role, Systems Analyst Role, Systems Architect Role, Technical Specialist Role, and Test Specialist Role.
2. *recently proposed roles* (CWA, 2018, p.38) including: Solution Designer Role, Digital Transformation Leader Role, Devops Expert Role, Data Scientist Role, Data Specialist Role, Scrum Master Role, and Product Owner Role.

Croatian graduate study programmes in business informatics will be briefly elaborated in relation to the presented UCT professional roles.

Faculty of Economics in Osijek offers graduate programme in *Business Informatics*, which is organized over the course of four semesters on graduate level. The goal of the programme is to enable students to get skills and knowledge of business information systems as well as ICTs, and are acquainted with the latest tools and methods for developing business software applications.

Regarding the Faculty of Economics in Split, the graduate program students have opportunity to attend the 2-year course: *Information management* which provides them with various ICT skills and knowledge oriented to ICT support to business processes, mainly in e-competency area such as ICT implementation, ICT planning, and ICT manage, less in ICT develop and ICT build.

Faculty of Economics and Business Zagreb delivers the graduate study in *Managerial informatics*, which is taught in both Croatian and English. Graduate study lasts for one year, and is focused to skills related to information management, electronic business, business process management, knowledge discovery in

databases, data warehousing and databases, decision support systems, information technology risk, and others.

Faculty of Organization and Informatics carries out undergraduate, graduate and postgraduate studies in Informatics. Graduate study has four study programme orientations available within the graduate studies in *Informatics*: *Information and Software Engineering* (IPI), *Business Systems Organization* (OPS), *Databases and Knowledge Bases* (BPBZ) and *Informatics in Education* (IO). Studies are held in Varaždin and enable the students to gain competences to solve the most complex problems in the areas of ICT development and application, for the purpose of increasing the efficiency of modern organizations. (FOI, 2018).

Although the described graduate studies differ in their names (Business Informatics, Information management, Managerial informatics, Informatics), they all develop similar competencies.

Diploma supplement for all study programmes contains the list of courses that are taught at the course. Based on the courses taught at analysed studies (EFOS, EFST, EFZG and FOI), students of these study programmes are currently taught for the following European ICT profiles mentioned above: Account Manager, Business Information Manager, CIO, ICT Operations Manager, Data Scientist, Project Manager, Business Analyst, Enterprise Architect, Data Specialist, and Developer.

3 Methodology

For the purposes of this research, survey was conducted among students who finished graduate study of Business Informatics, from 2000 to April, 2017. The survey was distributed over email to the list of graduates of each Faculty, and answers were collected using Google Forms, from February to April 2017. In addition, the survey was placed on social networks (e.g. Facebook groups of graduate business informatics students).

The research instrument was organized into three sections: demographics, employment, and satisfaction with the graduate study. First section consisted of 7 demographic questions: year of birth, gender, county of residence, year of graduate study enrolment, and year of completion of the study, average grade of the study and student status. Second set of questions referred to employment status. Respondents were asked when did they get their first job, are they currently employed, current job industry (according to National Classification of Activities - NKD 2007), county of employment, legal form of employment entity, current net salary etc. The third part of the questionnaire consisted of 14 statements for which the respondents were supposed to express their agreement on a five-point Likert scale (1-"entirely disagree"; 5-"entirely agree"). The statements referred to student's satisfaction with the study, satisfaction with professors and assistants, and study quality.

In Table 1, we provide the information about the total number of graduate students who finished Business Informatics study from 2006 to April 30, 2017 at all observed faculties, sample of students that has been contacted, as well as distribution of the respondents who took part in the survey. Data were processed and analysed using the statistical software package IBM SPSS Statistics v 24. It can be noted that the response rates are high, ranging from 43.5% (FOI) to 64.8% (EFST).

Table 1. Distribution of the respondents

Faculty	# of graduated students	# of students in sample	# of respondents	Response rate (%)
EFOS	262	123 (46.9 %)	60	48.7
EFST	108	71 (65.7%)	46	64.8
EFZG	633	200 (31.5%)	91	45.5
FOI	1472	200 (13.5%)	87	43.5

Source: Authors' work, based on survey 2017

4 Results

4.1 Sample characteristics

Table 2 presents the distribution of students by birth. Most of the students were born between year 1986 and 1993, which refer to the first generation of "Bologna students". The students are almost equally distributed by the gender (male – 56.3%, female – 43.7%).

Table 2. Year of birth of students

Characteristics	# of students	%
Year of birth		
1970-1980	3	1.1
1981-1985	32	11.3
1986-1989	140	49.3
1990-1993	109	38.3
Total	284	100.0

Source: Authors' work, based on survey 2017

Table 3 shows distribution of the respondents regarding graduation faculty. Most of the respondents finished EFZG, followed by FOI, but those two faculties produced the largest number of graduate students who enrolled business informatics graduate study. Most of the students were Bologna students (92%) enrolled in academic year 2005/2006 or later.

Table 3. Graduation faculty

	# of students	%
Graduation faculty		
EFOS	60	21.1
EFST	46	16.2
EFZG	91	32.0
FOI	87	30.6
Total	284	100.0
Bologna study programme		
No	22	7.7
Yes	262	92.3
Total	284	100.0

Source: Authors' work, based on survey 2017

Almost 58% of the respondents are very good students with final grade between 3.5 and 4.4. There were only 11 students with the grade below 2.9. The 2/3 of the respondents graduated between year 2011 and 2015 (see Table 4).

Table 4. Year of graduation and average grades

	# of students	%
Average grades		
2.5-2.9	11	3.9
3.0-3.4	73	25.7
3.5-3.9	81	28.5
4.0-4.4	83	29.2
4.5-5.0	36	12.7
Total	284	100.0
Year of graduation		
2006-2010	29	10.2
2011-2015	190	66.9
2016-2017	65	22.9
Total	284	100.0

Source: Authors' work, based on survey 2017

4.2. Employment of Business Informatics Graduates

Table 5 shows average time (in months) to first employment.

More than half of the students from all observed faculties found a job within one month from graduation. 30% of the students were employed between 2 to 6 months, and only 12 of them have been looking for a job for more than a year from graduation.

From the same table it can be seen that students mostly stayed at the first job they found after graduation. At the time of conducting this survey, almost 95% of the respondents were employed.

Table 5. Time to first employment; # of jobs from graduation; current employment status

	# of students	%
Time to first employment		
Up to 1 month	147	51.8
2-6 months	85	29.9
6-12 months	35	12.3
13-36 months	12	4.3
Total	284	100.0
# of jobs from graduation		
Only 1	129	45.4
2 jobs	86	30.3
3 jobs	46	16.2
4 to 10	18	6.3
Total	284	100.0
Current employment status		
Employed	269	94.7
Not employed	15	5.3
Total	284	100.0

Source: Authors' work, based on survey 2017

From the Table 6 it can be concluded that more than one third of all students work in ICT sector, followed by Professional, scientific and technical activities, indicating that students can find a job in different industries, both related to economics or ICT.

Table 6. Current job industry*

Job industry	# of students	%
Information and communication	103	36.3
Professional, scientific and technical activities	27	9.5
Finance and insurance	26	9.2
Education	16	5.6
Wholesale and retail trade	14	4.9
Other service activities	14	4.9
Administrative and support service activities	10	3.5
Construction	10	3.5
Public administration; compulsory social security	10	3.5
Manufacturing	10	3.5
Unemployed	15	5.3
Other (4 and less students)	21	7.4
No Answer	8	2.8
Total	284	100.0

**Job Industry classified according to NKD 2007. – Translation according to Statistical Yearbook of the Republic of Croatia (2017).*

Source: Authors' work, based on survey 2017

Table 7 indicate that almost 57% of the students have a work agreement for unspecified period. One quarter of them, have an agreement for fixed period. Surprisingly, students are not keen on establishing their own company or craft (only two students were self-employed).

Table 7. Form of employment*

Form of employment	# of students	%
Craftsmen	1	0.4
Through the agency	1	0.4
Self-employed, owner of own company / craft	2	0.7
Professional training without employment	26	9.2
Student agreement	2	0.7
Work agreement for unspecified period of time	164	58.0
Work agreement for fixed period of time	71	25.1
Work agreement	1	0.4
Unemployed	15	5.3
No answer	1	0.4
Total	284	100.0

**Form of employment translation according to Statistical Yearbook of the Republic of Croatia (2017).*
Source: Authors' work, based on survey 2017

Table 8 reveals that the students who finished Business Informatics study mostly work in a private sector (70% of them).

Table 8. Legal form of employment entity*

Legal form of employment entity	# of students	%
Trading companies in which the government hold majority in ownership	1	0.4
In government sector	35	12.3
In public sector (kindergartens. schools, faculties, …)	24	8.5
In private sector	198	69.7
In own company	11	3.9
Unemployed	15	5.3
Total	284	100.0

**Form of employment entity translation according to Statistical Yearbook of the Republic of Croatia (2017).*
Source: Authors' work, based on survey 2017

Net salary of the respondents differs a lot. Most of the students have a salary greater than average net salary in Croatia (which is approximately 6,000 HRK). From the Table 9 it can be seen that the largest number of the respondents have a net salary between 4.000 and 6.000 HRK. Less than 3.000 HRK gets only 10% of them. Those respondents mostly have professional training without employment. Only 16.2% of respondents have net salary greater than 10,000 HRK.

Table 9. Net salary

Net salary	# of students	%
3,000-3,999 HRK	18	6.3
4,000-5,999 HRK	80	28.2
6,000-7,999 HRK	61	21.5
8,000-9,999 HRK	36	12.7
less than 2,999 HRK	28	9.9
more than 10,000 HRK	46	16.2
Unemployed	15	5.3
Total	284	100.0

Source: Authors' work, based on survey 2017

Table 10 reveals that 2/3 of the respondents work on jobs related to ICT. According to European ICT professional profiles, 15% of them work as developers, followed by digital transformation leaders, business analysts and CIOs. In addition, teachers of ICT in secondary school as well as professors/assistants at the faculties can be found. However, about one fourth (27.8%) of the respondents have a job that is not directly related to ICT.

Table 10. Type of job

Type of job	# of students	%
European ICT professional profiles		
Developer	43	15.1
Digital transformation leader	38	13.4
Business Analyst	29	10.2
CIO	21	7.4
Data administrator	10	3.5
Service support	5	1.8
DevOps	1	0.4
Account manager	1	0.4
Other types of jobs		
Teacher of ICT in secondary school	5	1.8
Professor / Assistant at the faculty	2	0.7
Other job related to ICT	35	12.3
Other		
Not related to ICT	79	27.8
Unemployed	15	5.3
Total	284	100.0

Source: Authors' work, based on survey 2017

4 Discussion and conclusion

The main purpose of this paper is to presents the research results about survey conducted on sample of business informatics graduates of four Croatian largest faculties. In this preliminary study, characteristics of

their employment were explored. From the results, it can be seen that the time from graduation to their first job is relatively short. More than 50% of the respondents found their first job within one month from graduation. It can also be noted that some students found a job during study, so they were already employed after graduation. Most of the students (95%) were employed at the time of the survey, mostly in a private sector. Most of the respondents, about two-thirds, work on jobs related to ICT, such as developers, business analysts, CIOs etc. These preliminary results indicate that the knowledge and skills acquired at the Croatian faculties teaching business informatics graduate studies are sufficient to find a job in the ICT sector. However, the results of the study has several limitations that also indicate the directions for future research. First, we provide only descriptive statistics on the total sample, while larger sample would allow the comparison among faculties. Second, in our research we focus only to Croatia, while comparison with similar countries is lacking, indicating the need for the future research in neighbouring countries. Therefore, we plan to extend the sample in the future, both in Croatia, and to the other countries, which will provide the more objective basis for the conclusion on the employability of business informatics graduates.

References

Andrews, G. & Russell, M. (2012). Employability Skills Development: Strategy, Evaluation and Impact'. *Higher Education, Skills and Work Based Learning*, 2(1), 33-44.

CEN Workshop Agreement / CWA (2018). European ICT Professional Role Profiles, Retrieved from http://www.ecompetences.eu/wp-content/uploads/2018/01/CWA_Part_3_EU_ICT_PROFILES_METHODOLOGY_DRAFT.pdf

Chapman, J. A. & Lovell, G. (2006). The competency model of hospitality service: why it doesn't deliver. *International Journal of Contemporary Hospitality Management,* 18(1), 78-88.

Guthrie, K., Griffiths, R., & Maron, N. (2008). Sustainability and revenue models for online academic resources. Retrieved from http://www.jisc.ac.uk/publications/programmerelated/2008/scaithakasustainability.aspx

Gyamfi, S. A., & Gyaase, P. O. (2017). Virtualization of University Education: The Impact of ICT-Mediated Learning Environment on Students' Performance. *International Journal of E-Services and Mobile Applications (IJESMA)*, *9*(4), 24-40.

International Federation of Accountants, Education Committee. (2003). International Education Paper IEP 2: towards competent Professional Accountants. New York: IFAC.

Jelić, M. (2012). Perception of the higher education in informatics on the labour market. Unpublished graduate thesis, Faculty of Economics and Business, Zagreb.

Leväsluoto, J., Heikkilä, J., Tuovinen, J., & Viitanen, K. (2016). Gamification as an enabler of mutual learning in complex health care systems. *International Journal of E-Services and Mobile Applications (IJESMA)*, *8*(4), 35-47.

Lowry, L. L., & Flohr, J. K. (2005). No Student Left Behind: A Longitudinal Assessment of the Competency-Based Framework Used to Facilitate Learning in a Capstone Tourism Course. *Journal of Hospitality & Tourism Education*, 17(4), 28-35.

Marius, M. (2014). What does it mean to be an "ICT professional"?. Retrieved from http://www.ict-pulse.com/2014/11/ict-professional/

Morgado, L., Varajão, J., Dominguez, C., Oliveira, I., & Sousa, F. (2014). Balancing European SME managers' training contents: perceived importance & training needs. *Business systems research*, *5*(2), 4-22.

Sisson, L. G., & Adams, A. R. (2013). Essential Hospitality Management Competencies: The Importance of Soft Skills. *Journal of Hospitality & Tourism Education*, 25(3), 131-145.

Statistical Yearbook of the Republic of Croatia. (2017). Retrieved from https://www.dzs.hr/Hrv_Eng/ljetopis/2017/sljh2017.pdf

Suša, D. (2014). Digital immigrants and digital natives: Learning business informatics at higher educational level. *Business systems research*, *5*(2), 84-96.

Wang Y., & Tsai, C-T. (2014). Employability of Hospitality Graduates: Student and Industry Perspectives, *Journal of Hospitality & Tourism Education*, 26(3), 125-135.

Wessels, P. L. (2005). Critical information and communication technology (ICT) skills for professional accountants. *Meditari Accountancy Research*, 13(1), 87-103.

Yorke, M. (2004). Employability in the Undergraduate Curriculum: some student perspectives. *European Journal of Education*, 39(4), 409-427.

FOI. (2018). Graduate studies in informatics. Retrieved from https://www.foi.unizg.hr/en/students/gsp/inf

Analysis of Students' Experiences with Microsoft Dynamics NAV Solution Using Technological Acceptance Model

Simona Sternad Zabukovšek
Faculty of Economics and Business
University of Maribor
Razlagova ul. 14, 2000 Maribor, Slovenia
simona.sternad@um.si

Ruben Picek
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
ruben.picek@foi.hr

Polona Tominc
Faculty of Economics and Business
University of Maribor
Razlagova ul. 14, 2000 Maribor, Slovenia
polona.tominc@um.si

Samo Bobek
Faculty of Economics and Business
University of Maribor
Razlagova ul. 14, 2000 Maribor, Slovenia
samo.bobek@um.si

Abstract. *Enterprise Resource Planning (ERP) solutions are the most frequently used software tool in companies in all industries. The growing body of scientific literature about the acceptance of ERP solutions by users in companies reflects the growing percieved importance of ERP solutions for business management as well. The labour market requires the knowledge and skills for usage of ERP solutions from graduates – future employees. The main objective of our paper is therefore the identification of important factors that contribute to the acceptance of ERP solutions by students in economics and business and that shape their intentions to use this knowledge in the future. The conceptual model of our research is based on the Technology Acceptance Model (TAM), extended by identified important multidimensional external factors that refer to (1) students' personal characteristics and information literacy, (2) perceived system and technological characteristics of ERP solutions and (3) perceived support within the study process. The conceptual model formed was tested using the structural equation modeling. Research results revailed that several dimensions of the three external factors play an important role in shaping the attitudes towards acceptance of ERP solutions by students. Results of the study have important implications for higher education institutions, reforming and updating their study programs, as well as for educators in the field of economics and business sciences.*

Keywords: ERP solutions, TAM, economics and business' gruaduates, acceptance model

1 Introduction

The most widely used integrated solutions for business in companies from almost all industries worldwide are Enterprise Resource Planning (ERP) solutions. Number of ERP users within companies is growing very fast as well; employees are using ERP solutions daily at their work.

Therefore it is not surprising that the research studies regarding adoptions and acceptance of ERP solutions by users at different levels within companies are emerging (for example Costa et al., 2016). The most frequently used research approaches in these studies are (Awa et al., 2016): technology acceptance model (TAM) (Davis, 1989), theory of reasoned action (TRA) (Fishbein & Ajzen, 1975), theory of planned behaviour (TPB) (Ajzen, 1991), innovation diffusion theory (IDT) (Rogers, 2003), stage model (SM) (Poon & Swatman, 1999) and technology-environment-organization (T-O-E) (Tornatzky & Fleisher, 1990). In this area, TAM proved to be the most efficient model to study adoption in information systems (IS) (Shih and Huang, 2009; Sternad et al., 2011; Costa et al., 2016) and therefore numerous IS researchers apply this method to study ERP acceptance as well.

Because of that, there is also no doubt that the knowledge and skills of ERP solutions usage are among important competences of graduates in the field of economics and business, for achieving a competitive position in the labor market. In past few years, selected

topics of ERP solutions have become an integrative part in curricula in the management and business studies, within courses, such as: Accounting Information Systems, Enterprise Resource Planning, Information Systems etc. On the other hand all leading ERP vendors such as SAP, Microsoft, Oracle etc. have university academic alliances such as SAP University Alliances (SAP, 2018), Microsoft Dynamics Academic Alliance (Microsoft, 2018), Oracle University (Oracle, 2018) etc. which help higer education institutions to use their ERP solutions in their curriculum and thus preparing students with hands-on experience in using modern business applications. Despite the recognized importance of the ERP solutions as a business management tool within companies and the importance of this knowledge for graduates, researches aimed at identification of factors shaping students' attitudes towards the acceptance of ERP solutions, are rather scarce (Davis & Comeau, 2004; Shivers-Blackwell & Charles, 2006; Scott & Walczak, 2009; Iriberri, 2015).

The main objective of our paper is therefore the identification of important factors that contribute to the acceptance of ERP solutions by students in economics and business and that shape their intentions to use this knowledge in the future. The conceptual model of our research is based on TAM. The key purpose of TAM within our study is to provide a basis for testing the impact of additional external factors on students' internal beliefs (perceived usefulness - PU and perceived ease of use – PEOU), attitudes (AT), intentions (behavioural intention - BI) and actual use (Davis et al., 1989) of the ERP solutions. Identified important multidimensional external factors refer to (1) students' personal characteristics and information literacy, (2) perceived system and technological characteristics of ERP solutions and (3) perceived support within the study process.

2 Literature review

The organization Gartner Group first defined ERP as a concept more than 25 years ago (Montgomery et al. 2018). ERP systems initially focused on automating back office functions (functions which did not directly affect customers), while front office functions (functions which directly dealt with customers, e-business or supplier relationship management (SRM) became integrated later, when the Internet enabled the simplified communication with external parties. The organization Gartner Group (Ganly et al., 2013) in year 2013 introduced the term "postmodern ERP" (some call it also eXtended ERP – xERP). According to Gartner's definition of the postmodern ERP strategy, legacy systems of monolithic and highly customized ERP suites, in which all parts are heavily inter-dependent, should be replaced by a mixture of both cloud-based and on premise applications, which are more loosely coupled and can be easily exchanged if needed. The organization Gartner Group has evolved its definition over time and now defines ERP as an application strategy focused on several distinct enterprise application suite markets. They segment ERP into four major business process support areas: financial management systems, human capital management (HCM), enterprise assets management (EAM), and manufacturing and operations (Montgomery et al. 2018). Early ERP providers focused on large enterprises, but smaller enterprises are increasingly using ERP systems as well. The worldwide ERP market grew from 3.8% and 24.4B USD in 2012 to 25.4B USD in 2013. The global ERP software market is projected to reach $47.71 billion by 2022 growing at a CAGR of 7.0% during the forecast period (2016 to 2022). Company SAP is in market leadership position, follow by Oracle, Sage, Infor and Microsoft (SMRC, 2017). It is expected that ERP will remain the basic important software in the organisations.

Several theoretical approaches have been used to investigate the determinants of acceptance and the use of new information technology (IT), such as the theory of reasoned action (TRA; Fishbein & Ajzen, 1975), the theory of planned behaviour (TPB; Ajzen, 1991), the theory of the technology acceptance model (TAM; Davis et al., 1989), innovation diffusion theory (IDT; Rogers, 2003), stage model (SM; Poon & Swatman, 1999), technology-environment-organization (T-O-E; Tornatzky & Fleisher, 1990); and resource-based view (Caldeira & Ward, 2003). Compared to competing models, TAM is believed to be more parsimonious, predicative, and robust (Venkatesh & Davis, 2000; Lu et al., 2003; Liu & Ma, 2006), and therefore it is most frequently used by IS/IT researchers (Davis, 1989; Davis et al., 1989; Amoako-Gyampah & Salam, 2004; Lee et al., 2010; Costa et al., 2016). TAM posits that two beliefs − perceived usefulness (PU) and perceived ease of use (PEOU) − are of primary relevance for computer acceptance behaviour (Davis et al., 1989). PU is defined as "the degree to which a person believes that using a particular system would enhance his or her job performance" (Davis 1989, p. 320). PEOU in contrast, refers to "the degree to which a person believes that using a particular system would be free of effort" (Davis 1989, p. 320). The two central hypotheses in TAM state that PU and PEOU positively influence an individual's attitude towards using a new technology (AT), which in turn influences his or her behavioural intention (BI) to use it. Finally, intention is positively related to the actual use (U). TAM also predicts that PEOU influences PU, as Davis et al., (1989, p. 987) put it, "effort saved due to the improved perceived ease of use may be redeployed, enabling a person to accomplish more work for the same effort". The key purpose of TAM is to provide a basis for identifying the impact of external factors on internal beliefs, attitudes, and intentions (Davis et al., 1989). Original TAM is presented in Figure 1 by the grey rectangle. The original TAM is well established and tested and furthermore, a variety of extensions regarding external factors for

examining the antecedents of PU and PEOU have been developed such as TAM 2 (Venkatesh & Davis, 2000), UTAUT (Venkatesh et al., 2003) and TAM 3 (Venkatesh & Bala, 2008).

Even though TAM can be applied to a variety of technologies, the extensions and modifications of TAM are needed when analyzing specific information systems (Calisir et al., 2009). Although the number of studies analyzing the acceptance of ERP solutions by users in companies are emerging, they are still scarce and most of them investigate a very limited number of specific external factors (Calisir et al., 2009; Shih & Huang, 2009; Sun et al., 2009; Youngberg et al., 2009; Lee et al., 2010; Sternad et al. 2011; Sternad & Bobek, 2013, 2014; Mayeh et al., 2016; Costa et al., 2016). The researches aimed at analyzing factors influencing the ERP solution acceptance by students are even more scarce (see Shivers-Blackwell & Charles, 2006; Scott & Walczak, 2009; Iriberri, 2015). Shivers-Blackwell and Charles (2006) researched student readiness to use ERP technology using TAM, but they studied students' ERP acceptance in specific circumstances, namely, students read an online newsletter provided by the ERP communication, education, and training team entitled "What is ERP", first. Participants were then solicited by their professors to complete the survey, without any practical experience of ERP solution usage. Scott and Walczak (2009) examined cognitive engagement, prior experience, computer anxiety, and organizational support as determinants of computer self-efficacy in the use of a multimedia ERP system's training tool. They also examined the impact of computer self-efficacy on its acceptance. Iriberri (2015) researched the external factors' impact - training and teaching - on actual use.

3 Conceptual model and hypotheses

The main objective of our research is to identify the factors, included into the extended TAM as external factors, that are significantly shaping the antecedents of students' attitudes and future intentions of students to use the ERP solutions.

As already mentioned, the TAM intruduced by Davis (1989) and Davis et al. (1989), suggests the following relationships (this original TAM is presented by grey rectangle in Figure 1) among the multidimensional constructs, that are perceived ease of use – PEOU, perceived usefulness – PU, attitude toward using ERP system – AT, behaviour intention – BI, actual use – Use and in the case of our research refer to the ERP solutions:

H1: Perceived ERP ease of use (PEOU) has positive and direct effect on perceived ERP usefulness (PU).

H2: Perceived ERP ease of use (PEOU) has positive and direct effect on attitude toward ERP system (AT).

H3: Perceived ERP usefulness (PU) has positive and direct effect on attitude toward ERP system (AT).

H4: Attitude toward ERP system (AT) has positive and direct effect on behaviour intention (BI).

H5: Behaviour intention (BI) has positive and direct effect on actual use (Use).

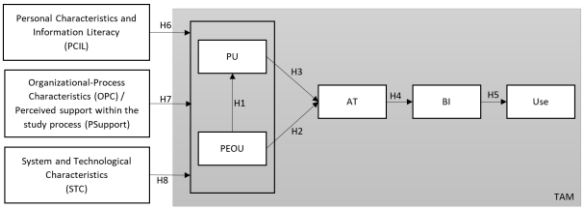


Figure 1. Conceptual Model

Even though TAM can be applied to a variety of technologies, it must be extended and modified for analysis of specific information systems (Calisir et al., 2009), as we already pointed out. The literature review reveiled that the external factors in general can be divided into three groups of factors: personal characteristics and information literacy (PCIL), system and technological characteristics (STC), and organizational-process characteristics (OPC) (see Sternad et al., 2011, Sternad & Bobek, 2013, 2014).

Personal characteristics and information literacy (PCIL), including personal characteristics that can influence individuals' perceptions of ERP system acceptance and usage, were analyzed in the past: personal innovativeness from the IT view-point (Yi et al., 2006; Thompson et al., 2006), computer anxiety (Venkatesh et al., 2003), computer self-efficiency (Venkatesh & Davis, 2000; Venkatesh et al., 2003; Shih & Huang, 2009) and perceived individual benefits (Hsu et al., 2015).

In contrast to the majority of researches regarding IT implementation which are very wide, the fact that ERP implementation research is focused on single technology-software solution, implies that the specific perceived technological characteristics should be examined. The literature review suggests that the following external factors are important within STC: system performance (Venkatesh et al., 2003; Kositanurit et al., 2006), user manuals (help) (Kelley, 2001; Kositanurit et al., 2006), quality of ERP system (Costa et al., 2016) and quality of information in ERP system (Hsu et al., 2015).

In the conceptual model of our research the modifications were implemented within the OPC construct, since the environment within the higher education institutions differs from the business environment in companies. Organizational-process characteristics (OPC) capture various social processes, mechanisms, and support organizations that guide individuals to facilitate the use of an ERP system. Since the students' acceptance of ERP solutions is in the focus of our research, the factors associated with their perceived support within the study process (during

course lectures and exercises regarding ERP solution) were taken into account; therefore the OPC construct was reshaped with the purpose to cover the educational organization view point. PSupport – Perceived support within study process includes perceived social influence (of teachers, other students and professionals participating in the educational process) (Venkatesh et al., 2003) and perceived characteristics of training and education on ERP system (Amonko-Gyampah & Salam, 2004; Bueno & Salmeron, 2008; Bobek & Sternad, 2011).

Therefore, the following hypotheses were formed:

H6: Personal characteristics and information literacy (PCIL) has a positive impact on the perceived ERP usefulness (PU).

H7: Perceived support within the study process (Psupport) has a positive impact on the perceived ERP usefulness (PU).

H8: Perceived system and technological characteristics (STC), has a positive impact on perceived ERP ease of use (PEOU).

4 Research design and methodology

The questionnaire was developed in three phases. In the first phase, we clarified the relationships between the constructs and the measurement scales for individual constructs, we reviewed the literature and resources. A questionnaire was employed. All items in the questionnaire were scored on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). The research design consisted of five constructs arising from the TAM model (PEOU, PU, AT, BI and Use) and three external constructs (PCIL, STC and Psupport), that we formed and included into the expanded TAM model. The external factors are therefore included by the three second-order constructs, based on all manifest variables of the underlying lower-order factors. PCIL includes: personal innovativeness toward IT, computer anxiety, perceived computer self-efficiency and perceived individual benefits. STC is composed of: system performance, user manuals (help), quality of ERP system and quality of information in ERP system. Psupport includes: perceived social influence (of teachers, other students and professionals participating in the educational process) and perceived characteristics of training and education on ERP system. Our conceptual model includes 15 first-order factors and 3 second-order factors.

In the second phase the instrument was pilot tested with a group of 30 ERP users in an organization. Based on the results of the pilot testing, revisions and additions were made to the instrument.

In the third phase the survey was conducted. Our sample included a total of 87 Croatian students in the second (4th semester) year of undergraduate study programme "Economics of entrepreneurship". The survey was carried out at the end of semester after students' full interaction with Microsoft Dynamics NAV ERP solution (after 14 lecture hours), within the course that includes all together 30 teaching hours of lectures of ERP topics with focus on selecting and implementing IS in methodological way and 30 hours in computer lab where students adopt the knowledge of the business processes functions in Microsoft Dynamics NAV (introduction, basic in finance and accounting process, purchasing process, sales process and some advance functionality simulating every day activities). The Microsoft Dynamics NAV 2016 (NAV) was used. On the last lecture in the semester (June 2017) 87 questionnaires were properly filled out by respondents and used for the purpose of analysis. Respondents were 14.9% (13) male and 85.1% (74) female. The average age of students was 20.70 years.

Demographic data was analysed by SPSS. All other empirical data was analysed in two steps analysis using partial least squares (PLS) technique, with Smart PLS 3.2.1. PLS path modelling is a variance-based structural equation modelling (SEM) technique which is widely used in education, business and social sciences in past two decades (Henseler et al., 2016; Garson 2016). We utilized this approach because of the relatively small sample size combined with the second-order factors analysis. In the first step, measurement model was assessed, and in the second step, structural model. Path significance has been estimated using bootstrapping resampling technique with 500 sub-samples as suggested by Ringle et al. (2015). While analysing data, we followed the guidelines specified by Henseler et al. (2016) and Garson (2016).

5 Analysis and results

All measurement scales were examined for their psychometric properties (reliability, convergent validity, and discriminant validity) prior to testing hypotheses. For external factors second-order procedure has been used. Because not all external factors met assessment requirements of the measurement model, they were excluded from further analysis. These factors are: computer anxiety from PCIL group and user manuals from STC group. Cronbach's alpha (α) and composite reliability (CR) as measures of reliability was examined, where $\alpha > 0.7$ and CR > 0.7 assures adequate reliability for confirmatory purposes. For convergent validity Fornell and Larcker's assessment criteria has been adopted: all item factor loadings should be significant and exceed 0.70, and the average variance extracted (AVE) for each construct should exceed 0.50. Discriminant validity between constructs was assessed following Fornell and Larcker's recommendation that the square root of AVE for each construct should be higher than its correlation with any other latent variable. Results of measurement model were satisfactory (results can be obtained by authors). The hypotheses testing rutilize bootstrapping (with 500 subsamples) to test the

statistical significance of each path coefficient, using *t*-tests, as recommended by Chin (1998). Results of this analysis are shown in Figure 2.

Our research partly confirms results of original TAM. PEOU has no significant effect on PU (H1; b = 0.070, p>0.05) and has a moderate significant effect on AT (H2; b = 0.403; p<0.01). PU has also moderate significant effect on AT (H3; b = 0.506; p<0.01). AT strongly influences BI (H4; b = 0.764; p<0.01) and BI moderately influences the Use construct (H5; b = 0.523; p<0.01). Hypothesis H1 was therefore rejected, but hypotheses H2 – H5 were not.

The hypotheses H6, H7 and H8 that refer to the extension of the TAM, were also partly confirmed. Second-order factor Psupport has strong significant positive effect on PU (H7; b = 0.576, p<0.01) and on PEOU (b = 0.624, p<0.01). PCIL has no significant effect on PU (H6; b = 0.172, p>0.05). STC has no significant effect on PU (H8; b = 0.138, p>0.05). Hypotheses H7 was not rejected, while hypotheses H6 and H8 were rejected.

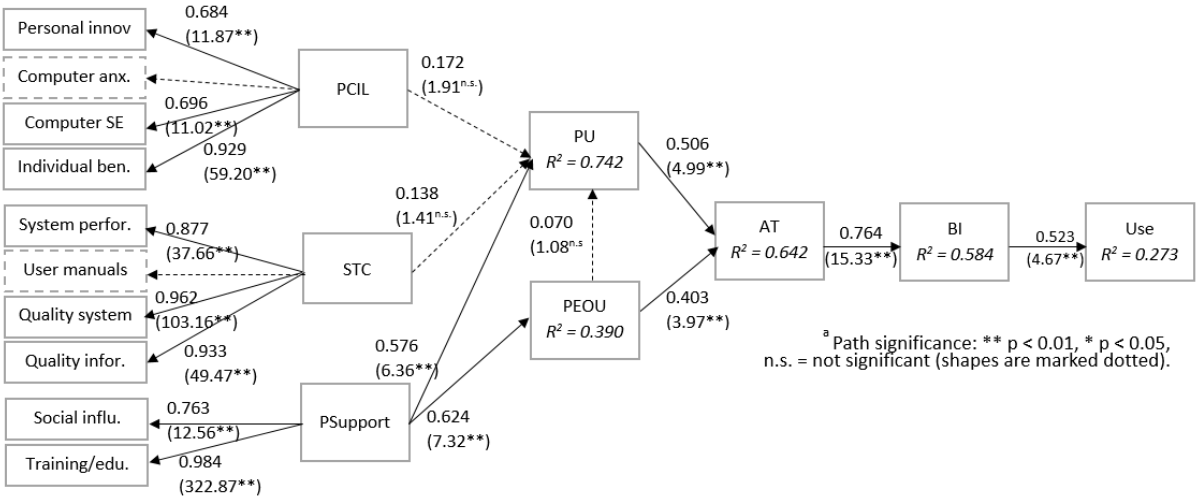


Figure 2: Results of the structural model analysis

6 Disscussion

Results of the present study regarding the hypotheses of original TAM model are consistent with several other research results regarding the IT/IS acceptance (Davis, 1989; Davis et al,. 1989; etc.). Both, PEOU and PU have strong positive effect on ERP usage, with the relationship of PU being a bit stronger. Therefore hypothesis H2 and H3 were confirmed. Also, PEOU has no statistical effect on PU. Hypothesis H1 was rejected. The findings about the importance of PEOU and PU in the literature are vague; Davis (1989), Davis et al. (1989) and Simon and Paper (2007) exposed that PU has stronger positive effect on IT/IS usage as PEOU, while PEOU has weaker or even no statistical effect on IT/IS usage after some time of usage. Since students were surveyed at the end of semester, where the ERP solution learning process took place, this could be the reason for the results obtained.

Hypotheses H4 and H5 were confirmed. Factor AT is vital in the TAM model and has very strong positive effect on BI and through it also indirect strong positive effect on Use, which is consistent with other researches (Pijpers & Montfort, 2006; Simon & Paper, 2007).

The main result of this research is the identification of external factors which influence students' ERP acceptance and have an impact on the antecdents of PU and PEOU.

The fact that ERP implementation research is focused on a single solution (technology) has enabled the possibility to study specific perceived system and technological characteristics. In the past, this external second-order factor (STC) was included into the research models of very few previous researchers (Sternad et al., 2011; Sternad & Bobek, 2013, 2014). Without the second-order factor Psupport being included into the model, factor STC was showing a statisticaly significant impact on PU, through the following first-order external factors: system performance, quality of NAV system and quality of information in NAV system. At the same time factor STC had no statistical impact on PEOU, (because of that, this relationship is not included in the Figure 2). When the second-order factor Psupport was added in the model, factor STC was not statistically significant any more. Therefore hypothesis H8 was rejected. The only remaining first-order factor within this second-order factor STC, namely user manuals, is not statistically significant – this is very likely the consequence of the fact, that users' manuals themselves are not included into the pedgogical process.

Similarly, the second order factor PCIL had significant and positive impact on the PU as long as the second-order factor Psupport was not included into the model (see Figure 2). Therefore hypothesis H6 was not

confirmed. The first-order factors within PCIL – namely personal innovativeness toward IT (software tools and applications), computer self efficiency and individual benefits - had significant impact on PU, but not on PEOU. First-order factor Computer anxiety, is not statistically significant – this can be explaned by the fact that the computer anxiety is probably a state of fear that is not known any more to the young population who grew up with the computers included in all (or at least many) aspects of the everyday's life.

Second order factor Psupport has strong significant and positive impact on PU and PEOU. This two relationships support hypothesis H7. As it was menitoned before, this factor was showing statistically significant impact on PEOU and PU through this two first-order factors: social influence and training and education. Factor social influence concerns opinions of teachers, other students and professionals participating in the educational process regarding students knowledge of ERP systems. It seems that students take into account the opinions of other (important) people for them.

The most important external factor is factor training and education. Therefore we suggest teachers to put an important effort into the preparation of excellent teaching materials and that try to explain ERP topics related content to students using simple routines, with the real business environment characteristics. To understand the ERP solutions is challenging for students, because they do not have practical experience of how ERP solutions are used in enterprises.

Conclusion

The aim of this research was to identify which external factors have impact on students' acceptance of ERP within study programme, while they are exposed to ERP solution (in our case Microsoft Dynamics NAV). We want to know how to motivate students to take course dealing with the ERP solution Microsoft Dynamics NAV, with all due seriousness and importance. That is why we studied 10 external factors which might have an impact on students' ERP acceptance. Studying the influence of the system of external factors on constructs not only contributes to the theory development, but also helps in designing teachers' curriculum.

Our research shows that most important external factors are especially two: training and education about ERP and social influence (of people - teachers, students, professionals - who have influence on students' perception regarding NAV/ERP). Factor training and education about ERP is more important than factor social influence. Therefore teachers have to put an important effort into the preparation of excellent teaching materials and that try to explain ERP topics related content to students using simple routines.

External first order factors within PCIL, namely personal innovativeness toward IT (software tools and applications), computer self-efficiency and individual benefits (regarding future job), were important personal factors, while computer anxiety was not important. Among important first order factors of STC were three: system performance, quality of ERP system and quality of information in ERP system, while factor user manuals was not important. As can be seen from Figure 2, all first order factors included into the second order PCIL and OPC became insiginificant, if the first order factors (training and education, social influence) forming the second order factor Psupport were included into the model.

Several implications for researchers and practitioners arise from the results of the extended version of TAM. Findings indicate that students have positive perception on the PU, PEOU, AT, BI and Use and that they understand the usefulness of ERP systems and their relevance as the support, important for their current or future jobs. These findings can help business schools assess students' engagement as they develop ERP software skills desired by employers. By many organizations, a big concern is whether students understand business processes (also process flows, subprocesses, etc.) behind ERP system. ERP system is very complex system and no single factor alone influences student's use of ERP. Our research showed that most important external factors are teaching and education of ERP system and social influence for students to understand the functionality of the system, its usefulness, and ease of use.

This study has certain limitations which are at the same time the opportunities for further research within this important and comperhensive topic. Since the respondents were limited to one group of students in Croatia, the study could be extended to other countries. Further research is needed to explore the importance of external factors included in different time frames (after introduction of course, at the end of course) as well as inclusion of additional external factors. Another limitation is also that research was conducted for one ERP solution only –namely for Microsoft Dynamics NAV; the importance of external factors may be different, when other ERP solutions are taking place (SAP, Infor ERP etc.).

References

Ajzen, I. (1991). The theory of planned behaviour. *Organizational Behaviour and Human Decision Processes*, 50, 179–211. DOI: 10.1016/0749-5978(91)90020-T.

Amoako-Gyampah, K., & Salam, A. F. (2004). An extension of the technology acceptance model in an ERP implementation environment. *Information & Management*, 41, 731–745. DOI:10.1016/j.im.2003.08.010.

Awa, H. O., Ukoha, O., & Emecheta, B. C. (2016). Using T-O-E theoretical framework to study the adoption of ERP solution. *Cogent Business &*

Management, 3, 1–23. DOI: 10.1080/23311975.2016.1196571.

Bueno, S. & Salmeron, J. L. (2008). TAM-based success modelling in ERP. *Interacting with computers*, 20 (6), 515–523. DOI: 10.1016/j.intcom.2008.08.003.

Calisir, F., Gumussoy, C. A., & Bayram, A. (2009). Predicting the behavioural intention to use enterprise resource planning systems—An exploratory extension of the technology acceptance model. *Management Research News*, 32(7), 597–613. DOI: 10.1108/01409170910965215.

Chin, W. W. (1998). Issues and opinion on structural equation modelling. *MIS Quarterly*, 22(1), 7–16.

Costa, C., Ferreira, E., Bento, F. & Aparicio, A. (2016). Enterprise resource planning adoption and satisfaction determinants. *Computers in Human Behaviour*, 63, 659–671. DOI: 10.1016/j.chb.2016.05.090.

Davis, C. H. & Comeau, J. (2004). Enterprise Integration in Business Education: Design and Outcomes of a Capstone ERP – based Undergraduate e-Business Management Course. *Journal of Inforamation Systems Education*, 15 (3), 287-300.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. DOI: 10.2307/249008.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982–1003. DOI: 10.1287/mnsc.35.8.982.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behaviour: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Ganly, D., Kyte, A., Rayner, N. & Hardcastle, C. (2013). Predicts 2014: The Rise of the Postmodern ERP and Enterprise Applications World. Gartner Group. Retrieved March 10, 2017, from https://www.gartner.com/doc/2633315/predicts--rise-postmodern-erp.

Garson, G. D. (2016). *Partial Least Squares: Regression and Structural Equation Models*. Asheboro, NC: Statistical Associates Publishers.

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS Path Modeling in New Technology Research: Updated Guidelines. *Industrial Management & Data Systems*, 116(1), 2–20. DOI: 10.1108/IMDS-09-2015-0382.

Hsu, P.-F., Yen, H. R., & Chung, J.-C. (2015). Assessing ERP post-implementation success at the individual level: Revisiting the role of service quality. *Information & Management*, 52(8), 925–942. DOI:10.1016/j.im.2015.06.009.

Iriberri, A. (2015). Integrating an ERP into the curriculum at a busienss school: the students' perceptions of SAP. *Academy of Ecucation Leadership Journal*, 19(2), 99- 108.

Kelley, H. (2001). *Attributional analysis of computer self-efficacy: dissertation*. (Unpublished doctoral dissertation). London, UK: Richard Ivey School of Business.

Kositanuri, B., Nqwenyama, O. & Osei-Bryson, K. M. (2006). An exploration of factors that impact individual performance in an ERP environment: an analysis using multiple analytical techniques. *European journal of information systems*, 15, 556–568. DOI: 10.1057/palgrave.ejis.3000654.

Lee, D. H., Lee, S. M., Olson, d. L., & Chung, S. H. (2010). The effect of organizational support on ERP implementation. *Industrial Management & Data Systems*, 110(2), 269–283. DOI: 10.1108/02635571011020340.

Liu, L., & Ma, Q. (2006). Perceived system performance: A test of an extended technology acceptance model. *Journal of Organizational and End User Computing*, 18(3), 1–24. DOI: 10.1145/1161345.1161354.

Lu, J., Chun-Sheng, Y., Liu, C., & Yao, J. E. (2003). Technology acceptance model for wireless internet. *Internet Research: Electronic Networking Applications and Policy*, 13(3), 206–222. DOI: 10.1108/10662240310478222.

Mayeh, M., Ramayah, T., & Mishra, A. (2016). The role of absorptive capacity, communication and trust in ERP adoption. *The Journal of Systems and Software*, 119, 58–69. DOI: 10.1016/j.jss.2016.05.025.

Microsoft. 2018. Microsoft Dynamics Academic Alliance. Retrieved May 20, 2018, from https://dynamics.microsoft.com/en-us/academic/

Montgomery N., Anderson R. P., Kostoulas, J. & Woodyer, A. 2018. High-Tech Tuesday Webinar: Best Opportunities and Bets for Growth in Enterprise Resource Planning. Retrieved May 20, 2018, from https://www.gartner.com/doc/3640429?ref=SiteSearch&sthkw=postmodern%20erp&fnl=search&srcId=1-3478922254.

Oracle. 2018. Oracle Univrstiy. Retrived May 20, 2018, from http://education.oracle.com/pls/web_prod-plq-dad/db_pages.getpage?page_id=3.

Pijpers, G. G. M. & Montfort, K. (2006). An investigation of factors that influence senior executives to accept innovations in information

technology. *International journal of management*, 23(1), 11–23.

Poon, S., & Swatman, P. (1999). An exploratory study of small business internet commerce issues. *Information & Management*, 35(1), 9–18. DOI:10.1016/S0378-7206(98)00079-2.

Ringle, C. M., Wende, S., & Becker, J.-M. (2015). SmartPLS 3. Boenningstedt: SmartPLS GmbH. Retrieved February 8, 2016, from http://www.smartpls.com.

Rogers, E. (2003). *Diffusion of Innovations* (4th ed.). New York: The Free Press.

SAP. 2018. SAP Universty Alliances. Retrieved May 20, 2018, from https://www.sap.com/training-certification/university-alliances.html.

Scott, J. E., & Walczak, S. (2009). Cognitive engagement with a multimedia ERP training tool: Assessing computer self-efficacy and technology acceptance. *Information & Management*, 46, 221–232. DOI: 10.1016/j.im.2008.10.003.

Shih, Y. Y., & Huang, S. S. (2009). The actual usage of ERP systems: An extended technology acceptance perspective. *Journal of Research and Practice in Information Technology*, 41(3), 263–276.

Shivers-Blackwell, S. L. & Charles, A. C. (2006). Ready, set, go: Examining student readiness to use ERP technology. *Journal of Management Development*, 25(8), 795–805. DOI: 10.1108/02621710610684268.

Simon, S. J. & Paper, D. (2007). User acceptance of voice recognition technology: an empirical extension of the technology acceptance model. *Journal of organizational and end user computing*, 19(1), 24–50.

SMRC. (2017). ERP Software Market Report 2017, Trends, Analysis, Share, Estimates and Forecasts to 2022. Statistics Market Research Consulting. Retrieved June 5, 2018, from http://www.mynewsdesk.com/us/pressreleases/erp-software-market-report-2017-trends-analysis-share-estimates-and-forecasts-to-2022-2287038.

Sternad Zabukovšek, S., Gradišar, M., & Bobek, S. (2011). The influence of external factors on routine ERP usage. *Industrial management + data systems*, 111(9), 1511–1530. DOI: 10.1108/02635571111182818.

Sternad, S., & Bobek, S. (2013). TAM-based external factors related to ERP solutions acceptance in organizations. *International Journal of Information Systems and Project Management*, 1(4), 25–38. DOI: 10.1016/j.protcy.2013.12.004.

Sternad, S., & Bobek, S. (2014). Comparative analysis of acceptance factors for SAP and Microsoft Dynamics NAV ERP solutions in their maturity use phase: enterprise 2.0 issues. In M.M. Cruz-Cunha, F. Moreira, & J. Varajao (Eds.), *Handbook of research on enterprise 2.0: technological, social, and organizational dimensions* (pp. 389–415). Hershey, New York: Business Science Reference, IGI Global.

Sun, Y., Bhattacherjee, A., & Ma, Q. (2009). Extending technology usage to work settings: The role of perceived work compatibility in ERP implementation. *Information & Management*, 46(6), 351–356. DOI:10.1016/j.im.2009.06.003.

Thompson, R., Compeau, D. & Higgins, C. (2006). Intentions to use information technologies: an integrative model. *Journal of organizational and end user computing*, 18(3), 25–46. DOI: 10.4018/joeuc.2006070102.

Tornatzky, L., & Fleisher, M. (1990). *The process of technology innovation.* Lexington: Lexington Books.

Venkatesh, V. & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision sciences*, 39 (2), 273–315. DOI: 10.1111/j.1540-5915.2008.00192.x.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), 186–205. DOI: 10.1287/mnsc.46.2.186.11926.

Venkatesh, V., Morris, M. G., Davis, F. D., Davis, G. B. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–479.

Yi, Y. M., Fiedler, K. D. & Park, J. S. (2006). Understanding the role of individual innovativeness in the acceptance of IT-based innovativeness: comparative analyses of models and measures. *Decision Sciences*, 37(3), 393–426. DOI: 10.1111/j.1540-5414.2006.00132.x.

Youngberg, E., Olsen, D., & Hauser, K. (2009). Determinants of professionally autonomous end user acceptance in an enterprise resource planning system environment. *International Journal of Information Management*, 29(2), 138–144. DOI:10.1016/j.ijinfomgt.2008.06.001.

Facebook as a Learning Tool - Students' Perspective

Mirela Mabić, Dražena Gašpar
University of Mostar
Faculty of Economics
Matice hrvatske bb, 88 000 Mostar, BiH
{mirela.mabic, drazena.gaspar}@ef.sum.ba

Abstract. *Popularity and adoption of social networks among young people, especially Millennials have been in full swing during the last decade. Many researches have shown that social networks, Facebook especially, can be successfully used for educational purpose, not just for leisure. The authors are interested in exploring the situation related to the use of Facebook in higher education in Bosnia and Herzegovina. Research is aimed to explore what students think about and whether they will accept and use Facebook in their education. The paper presents the results of research conducted among students of the Faculty of Economics at the University of Mostar in Bosnia and Herzegovina. The results confirm that students do not have a negative attitude and aversion to the use of Facebook in education, although the educational usage of Facebook is not everyday practice at the Faculty of Economics at the University of Mostar. The findings show that Facebook is already in use for the exchange of materials and information and that the students are ready for active implementation of Facebook for educational purposes, especially in communication with teaching staff, which would benefit for both sides.*

Keywords. Social network, student, higher education, quality in higher education

1 Introduction

Today's higher education institutions (HEIs) are faced with a great challenge – how to prepare students for entry into highly competitive, globalized, dynamic, high-tech, complex and interdisciplinary business environment. HEIs can respond to that challenge by equipping their students with appropriate skills, knowledge, values, and attributes. There is a strong drive to build and create knowledge together with understanding a working life and reformulating the concept of knowledge in education institutions. Namely, HEIs are bound to provide quality teaching that leads to learning outcomes and, above all, added value for their students (Gašpar & Mabić, 2015). In reaching that goal, social networks could be a useful tool. Social networks (SNs) could play a valuable role during the problem-based learning process, serving as a critical tool for information searching, organizing and analyzing data, and presenting solutions.

Many researches confirm the usefulness of social networks in higher education (Jones, 2015; Schroeder and Greenbowe, 2009; Chawinga, 2017; De Wever, Hämäläinen, Voet, & Gielen, 2015; Terrell, Richardson, & Hamilton, 2011; Waycott et al., 2010; Wheeler, 2010). Some of the research is focused on main drivers of social networks in higher education (Dumpit & Fernández, 2017; Tess, 2013), while others analyze the effects of social networks on the academic performance of university students (Lau, 2017). Some of the authors (Manca &Ranieri, 2016; Dzvapatsva, Mitrovic & Dietrich, 2014) have researched the potentials and obstacles of social media in higher education. Thus, Barczyk and Duncan (2011) explained that critics of social networks as an educational tool in HEI mainly stressed that social networks mostly offer poor reference material generated by unreliable sources. Some teachers saw social networks such as Twitter and Facebook as distracters to learning (Galagan, 2010), while others (Barczyk & Duncan, 2011; Harris & Rea, 2009) added challenges like absence of computing resources, disruption of web-based resources and plagiarism because of contents' openness and easiness of students' copy and paste. However, the most of authors think that it is more important to focus on how best social networks can align with lesson objectives or curricula (Kietzmann, Hermkens, McCarthy & Silvestre, 2011; Szapkiw & Szapkiw, 2011; Wheeler, 2010). Different authors (Chawinga, 2017; Tang & Hew, 2017; Menkhoff et al., 2014) analyzed the potential of Twitter in the teaching process.

Many authors have studied the application of Facebook for educational purposes. Thus, Chen (2011) presented how Facebook provides a common interaction environment for problem-solving and reasoning the people, which reflects the user satisfaction of Facebook and improved academic performance of the students. The study also presented that students found Facebook simple, suitable, user-friendly and effortless for academic discussion. The

study by Madge et al. (2009) proposed that Facebook could act as an important tool to assist students in settling in the classroom. They also proposed that Facebook could improve their teamwork and organizational skills. It is clear that the application of social networks and their integration into the educational process is nothing new in the world, but in Bosnia and Herzegovina (BiH) it is not yet in full swing. The reasons are manifold, ranging from missing infrastructure to personal views and beliefs. Infrastructure is the factor that can be reasonably improved with some efforts, but people's views and opinions are what is much more difficult to change. In order to improve the state, it is necessary first to investigate the current habits and attitudes of both students and teaching staff.

The main aim of the research was to explore what students at the Faculty of Economics, University of Mostar, think about and whether they will accept and use SNs for educational purposes. In addition, the aim was also to give a brief insight into their habits of using the Internet and social networks.

2 Methodology

The empirical study was conducted at the end of 2016 and early 2017 at the Faculty of Economics, University of Mostar. The convenient sample consisted of students of all study programs and all study years. Students included in research had experience related to the use of Facebook during some courses. Namely, some of the teachers at the Faculty of Economics have already created Facebook groups for their courses. They used Facebook for sharing learning materials, encouraging students' discussions about some themes related to courses. At the beginning of the survey, it was explained to students what the objective of the survey was and for what purposes the obtained data would be used. The survey was voluntary (students decided whether to participate) and anonymous. In total, 197 students agreed to participate in the survey.

The questionnaire consisted of questions on computer usage habits (how often, what activities, etc.), knowledge of SNs (what networks they use, how often, what they do on them, etc.), possible experiences in the application of Facebook during their education. In addition, students were offered a set of statements (Table 1) related to their views on the contribution of Facebook to the experience of education - in their opinion, how Facebook influence the communication process with faculty staff and quality of education. Students evaluated statement from 1 (completely disagree) to 5 (completely agree). At the end of the questionnaire, the student wrote gender and age. The questionnaire was designed by authors.

Table 1. The set of statements

(Code) Statement
(S1) Facebook contributes to the quality of studying
(S2) Facebook makes it easier to study
(S3) Facebook makes it easier to communicate with other colleagues
(S4) Facebook makes it easier to communicate with professors and instructors
(S5) Facebook makes it easier to exchange materials (text, pictures, videos, etc.)
(S6) Facebook encourages knowledge sharing
(S7) Materials available on Facebook are accurate
(S8) Materials available on Facebook are usable
(S9) Facebook should be extensively used during the study
(S10) Facebook should be the primary medium of communication with teaching staff
(S11) Facebook allows students to participate in the creation of educational contents
(S12) Facebook makes it easier to carry out teaching on schedule
(S13) Facebook stimulates the creativity of students
(S14) Facebook is a good source of information
(S15) Facebook contributes to a better performance of students
(S16) Facebook makes communication simple
(S17) Facebook makes passive participants active
(S18) Facebook expands views on education (during the study)
(S19) Facebook "pulls down" the walls of classrooms/buildings
(S20) It is good to share information on Facebook
(S21) The use of Facebook during the study cannot do any harm to me
(S22) I would recommend everyone to use Facebook
(S23) Facebook makes studying more interesting and active
(S24) I would recommend everyone to use Facebook during studies
(S25) Facebook should be actively used
(S26) Facebook has more positive than negative sides
(S27) Facebook can help me more than hinder me

Source: prepared by the authors

After technical and logical analysis, 185 questionnaires were accepted for further analysis; 8 questionnaires did not have any ratings with the set of statements and 4 questionnaires have grade 3 for all statements, so these questionnaires were excluded from further analysis. Data from the completed questionnaires were entered into a database in the SPSS for Windows (17.0, SPSS Inc. Chicago, Illinois, USA) which was used for data analysis. Descriptive analysis included absolute and relative frequencies,

measures of central tendencies (mean, mode) and measures of variability (standard deviation, the coefficient of variation).

Analysis of students by age and gender showed that the participants were 31.4% (n1 = 58) men and 68.6% (n2 = 127) women; considering that the survey was conducted among students, their age ranged between 19 and 25 years (4 students did not report their age, and 13 reported being more than 25 years old).

3 Results

Table 2 gives results about students computer usage habits (how often, what activities, etc.), knowledge of SNs (what networks they use, how often, what they do on them, etc.), and possible experiences in the application of SNs during their education.

Table 2.

		No (%)
How often do you use a computer?		
	every day	121 (65.4)
	several times a week	44 (23.8)
	less often	20 (10.8)
For what purposes do you use a computer?		
	entertainment	154 (83.2)
	learning	31 (16.8)
How often do you use the Internet?		
	every day	182 (98.4)
	< 1 hour	5 (2.7)
	1 – 3 hours	71 (39.0)
	3 – 5 hours	63 (34.6)
	> 5 hours	43 (23.6)
	several times a week	3 (1.6)
Everyday activities on the Internet		
	e-mail	60 (32.8)
	Facebook	164 (90.1)
	writing/reading in the forums	31 (16.9)
	online gaming	11 (6.1)
	reading news	111 (60.3)
	watching movies	6 (3.3)
	listening to music	132 (72.5)
	chat	168 (91.8)
	searching for information for the study	64 (34.6)
	reading professional texts	13 (7.1)

Source: author's calculations

The most prevalent programs (applications) that students most frequently use on their computers are various Internet browsers and the Microsoft Office suite, with the emphasis on Word. The response "I use Facebook" had an exceptionally high rate (to this particular question answer is given by more than 90% of students).

Personal profiles on Facebook have 96.8% students, 65.4% on Instagram, and 45.4% on YouTube. Other popular SNs (Twitter, LinkedIn and Google+) are slightly less common among students. Certainly, having a profile on SNs does not mean actively using it and. Since it was assumed that all 96.8% of students were also in the group of active Facebook users, a question about it was asked too. It is determined that only one of the respondents reported having a Facebook profile but not actively using it, so active users make 96.2% of the sample. It is similar with other SNs, with the percentage of those actively using them being slightly less than the percentage of students having profiles on them. In addition to the said networks, several students also reported using some other SNs not being offered in the responses, but they did not specify which social networks these were.

The students who actively use Facebook also briefly described their habits. Their results are presented in Table 3.

Table 3.

		No (%)
How long do you use the Facebook?		
	a year	3 (1.7)
	1-2 years	3 (1.7)
	2-3 years	12 (6.7)
	more than 3 years	161 (89.9)
Facebook friends		
	<100	7 (4.0)
	100-200	21 (11.9)
	200-500	82 (46.3)
	500-1000	46 (26.0)
	>1000	21 (11.9)
How often do you use the Facebook (post, share, comment, like, etc.)		
	every day	172 (93.0)
	< 1 hour	49 (28.7)
	1 – 3 hours	87 (50.9)
	3 – 5 hours	21 (12.3)
	> 5 hours	14 (8.2)
	several times a week	8 (4.3)
The primary reason for using Facebook		
	data and information gathering	78 (43.6)
	education	25 (14.0)
	self-promotion (personal data and photos posting)	5 (2.8)
	connect with old contacts	49 (27.4)

	maintain contacts	126 (70.4)
	connect with new contacts	51 (28.5)
	planning and organization of social events	42 (23.5)
	gathering general information	49 (27.4)
	trendy, cool	21 (11.7)
	Other	5 (2.8)

Source: author's calculations

The leading answers to the question "what do you do most often on Facebook" are “I chat,“ “I like posts,“ “I follow other's posts.“

After the questions on the Internet and Facebook habits, students evaluated a set of statements on the contribution of Facebook to the experience and quality of the educational process (all students evaluated them regardless of their habits of using Facebook). A descriptive analysis of the statements is shown in Table 4.

Table 4. Descriptive statistics for the set of statements

	N	[min-max]	D	M±SD	CV (%)
S1	185	[2-5]	4	3.87±0.69	17.9
S2	185	[3-5]	5	4.13±0.86	20.7
S3	185	[3-5]	5	4.76±0.50	10.5
S4	185	[1-5]	5	3.97±1.04	26.2
S5	185	[3-5]	5	4.77±0.49	10.0
S6	184	[2-5]	4	3.92±0.86	21.8
S7	185	[1-5]	3	2.98±0.80	27.0
S8	184	[1-5]	4	3.61±0.80	22.0
S9	184	[2-5]	5	3.97±0.96	24.3
S10	185	[1-5]	3	3.02±1.11	36.7
S11	185	[1-5]	3	3.16±1.08	34.2
S12	185	[1-5]	3	3.45±1.02	29.6
S13	185	[1-5]	3	3.48±1.01	28.1
S14	184	[1-5]	4	3.59±0.88	24.4
S15	185	[1-5]	3	3.17±1.04	32.7
S16	185	[3-5]	4	4.35±0.58	13.3
S17	184	[1-5]	3	3.46±0.97	28.0
S18	185	[2-5]	4	3.59±0.80	22.4
S19	183	[1-5]	4	3.77±0.76	20.2
S20	185	[1-5]	3	3.03±1.07	35.2
S21	185	[2-5]	4	3.70±0.75	20.4
S22	185	[1-5]	3	3.56±0.93	26.1
S23	185	[2-5]	4	4.01±0.86	21.4
S24	185	[2-5]	4	3.89±0.72	18.5
S25	183	[1-5]	4	3.72±0.77	20.6
S26	185	[1-5]	4	3.58±1.02	28.6
S27	185	[2-5]	4	3.75±0.87	23.2

N – the number of the answer, D – mode, M±SD – mean±standard deviation, CV - coefficient of variation

Source: author's calculations

Most of the statements have the value of the coefficient of variation, which confirms that they have representative mean. The exceptions are statements S10, S11 and S20, but since their coefficient of variation is slightly above 33%, the mean is also used for their analysis.

Since almost all students reported usage of Facebook, they were asked to answer several questions about their current use of Facebook for educational purposes. Thus, it is determined that 94.9% of respondents reported usage of Facebook as an aid during the study, 91.3% for some formal and informal groups whose main topic is studying, and 79.1% of them believe that Facebook can be used as an aid in learning, specifically through various group discussions.

Irrespective of the offered statements, in the part of the questionnaire on Facebook habits, students were asked to specifically express (yes or no) their views on the application of Facebook for educational purposes. Thus, it is determined that 85.8% of students believe that Facebook would be good for exchanging information related to the university, 89.1% believe that Facebook is good for sharing of teaching materials, and 60.8% feel that Facebook improves their study experience, while 24.4% of students do not know how to answer the question.

4 Discussion

The results show that more than a half of students use computers every day, one quarter does that several times a week, while others reported doing that less often (Table 2). As for the reasons for their use, the students offered only two reasons: entertainment and learning, with entertainment being significantly higher (83.2% vs. 16.8%).

Certainly, the use of computers does not show the frequency of the Internet use, the latter being available on other devices, particularly on mobile devices among students. That is why the question about the frequency of the Internet use had considerably different results. Almost all of the students answered the question as "every day." In this matter, ¾ of the students reported spending on the Internet up to 5 hours a day, and the rest more than 5 hours a day.

As for the activities on the Internet, the answers indicate that correspondence and use of Facebook are the most common activities (with about 90% of students being engaged in it on a daily basis). Reading news and listening to music are practiced every day by about ¾ of students. The least common extracurricular activities on the Internet include writing/reading in the forums, online gaming and watching movies, while the correspondingly

uncommon study-related activities include searching for information for the study and reading professional texts.

However, if the results of the computers use are compared with the results of the Internet use, it is evident that computer is not the dominant medium for the Internet use and online activities, although it is relatively frequent.

The results show that students use Facebook, but not primarily for educational purpose. However, the fact that they use Facebook can be a leading promoter for Facebook usage in education. Namely, the previous research (Sánchez, Cortijo & Javed, 2014) showed that Facebook adoption positively influences the educational usage of Facebook. According to the results, students are relatively active on SNs, especially on Facebook that is used by almost all students.

Most students have Facebook profiles longer than 3 years, almost ½ of them have between 200 and 500 and ¼ between 500 and 1000 friends, every day they post something, comment, like, etc. and spend between 1 and 3 hours a day on Facebook (Table 3). According to students' responses, the primary reason for using Facebook is to maintain contacts and to connect with old and new contacts. Only 14% of students reported education as the primary reason for using Facebook.

The best-rated statement according to mean grades is the statement S5: "*Facebook makes it easier to exchange materials (text, pictures, videos, etc.)*," closely followed by the statement S3: "*Facebook makes it easier to communicate with other colleagues*" (Table 4). None of the students rated these claims lower than 3, and the variations in ratings are relatively small too, which shows that the attitudes of students on these issues are relatively uniform. In addition to these, the claims S16, S2, and S23 also have means greater than four, with grades 4 and 5 dominant in students' responses, which reflects the agreement of students. Certainly, the agreement is not full because a part of the students does not have the opinion (grade 3). It shows that students feel that Facebook greatly facilitate communication, regardless of whether it is direct communication through correspondence or indirect communication accomplished through the exchange of different materials. That implies that students recognize and use the essential characteristics of Facebook, and these are strong interactivity and collaboration that generally characterize Web 2.0 technologies which Facebook belong. Of course, the communication is live and bidirectional, and it is no wonder that students feel that such communication makes studying more interesting and active. These results are consistent with the findings of Irwin et al. (2012). They analyzed student's perception using a designated Facebook page as a learning resource in courses at the Griffith University in Australia. The results of their research showed that students are receptive to incorporating Facebook into their courses. Students perceived the benefits of using Facebook through enhanced communication, interaction, and flexibility in course content delivery (Irwin et al., 2012).

The statement S7 was rated the lowest: "*Materials available on Facebook are accurate*." (mean 2.98 with dominant grade 3). When interpreting this result, one should take into account the assumption that students consider materials to be all materials available on Facebook. In this case, this result is acceptable because it shows that students are aware that only authorized materials and materials prepared by professionals can be considered accurate, high-quality and credible. That implies sharing of materials in groups created with the specific purpose and for a select group of people.

The results in Table 4 generally show that students have a relatively positive opinion on the application of Facebook during their education (studying). They find that Facebook significantly facilitates communication, primarily with colleagues, and then with teaching staff too. They believe that Facebook can improve the experience of studying, and a part of students feel that they can significantly contribute to a better performance of students by activating students in the exchange of data, information, ideas, opinions, positions, and knowledge as well as in the creation of new and innovative content for teaching. In addition, a significant number of students perceive Facebook as a tool for encouraging students to think outside of the box, which extends education beyond the classroom and outside of the bounds of what teachers transfer through various forms of teaching. Therefore, it is no wonder that students believe that Facebook should be extensively used during the study and would recommend everyone to use them. In addition, students agree more than they disagree that the advantages and benefits of Facebook overcome their negative sides and that their use can be beneficial more than frustrating. These results are similar to findings of Madge et al. (2009) and Chawinga (2017). Madge et al. (2009) proposed that Facebook could improve students' teamwork and organizational skills and act as an important tool in assisting students in the educational process. Chawinga reported that if appropriately deployed, Twitter and blogs emerged students to share and discuss course materials, post their course reflections and interact amongst themselves and with their teacher 24/7.

Despite the relatively positive attitudes of students towards the application of Facebook in the educational process, it should be noted that, in their opinion, Facebook should not be the primary medium for communication with teaching staff. Namely, new methods and media of communication, which are directly determined by the rapid and continuous development and impact of information technology, should complement the existing formal communication.

As it is presented in Table 4, the statement directly implying that Facebook contribute to the quality of studying (S1) is also offered. The mean grade of this statement is 3.88, with the range from 2 to 5 and four as the dominant grade. Not a single student reported being completely opposed; there are skeptics, but these results indicate the fact that students are very favorably disposed towards the application of Facebook during studies, which greatly enriches their study experience.

Almost 95% of respondents stated that they use Facebook for educational purposes; concretely each student is a member of at least one informal Facebook group in which discuss the themes related to study. More than ¾ of students thinks that the use of the Facebook is welcomed at the faculty, especially in the context of information and teaching materials sharing. Additionally, more than half of students think that Facebook improves their study experience.

5 Conclusion

So, students, could Facebook contribute to the quality of higher education?

The results show that there is a part of students who do not agree with it or at least do not agree with all of the given claims, but even these responses are sufficient for a positive answer to the question. Namely, students have recognized the significance of Facebook in promoting communication and sharing information. Additionally, the students' answers show that the use of Facebook during their study is not a novelty for them, because they already use Facebook for sharing useful information and teaching materials. Certainly, although the questionnaire was focused on the positive aspects of Facebook, students are evidently aware of their negative sides too. However, since the advantages of Facebook prevail, their use can be beneficial more than frustrating.

Based on the results of this study, it could be concluded that teaching staff at universities should take significant steps to increase the use of Facebook, as well as other social networks, at work. Certainly, the goal is not to adjust the educational process to SNs but to take the full advantage of SNs as a tool that can provide satisfaction and improve the experience of all interested parties, which will significantly reflect on the quality of the educational process.

One of the limitations of this research is the size of the sample. All the students who participated in this research come from the same institution – the Faculty of Economics at the University of Mostar, so findings cannot be generalized. Future research should include participants from different faculties and universities in Bosnia and Herzegovina.

Another limitation is a focus on specific SNs - Facebook. There are many other SNs, and their use and impact on educational process could differ.

However, since this research has not looked into attitudes of teaching staff, but only into views of students, that is what further research should do. It would be necessary to investigate the views and opinions of teaching staff on Facebook in general, their activities on Facebook and the extent to which they support the idea of using Facebook as a medium that can help improve and facilitate studying, but also bring improvements in the field of quality of education.

Similarly, the future research should also focus on negative aspects of the application of Facebook, as well as other social networks, both in everyday life and in the sphere of education.

The future research could be conducted among alumni, or people who studied at the time when there were no SNs and information and communication technologies were not used to such an extent. It would be interesting to know their opinions on past and present possibilities and ways of sharing information, materials, and knowledge. Their answers could be compared with the answers of today's students, which would yield clues about the influence of information and communication technologies on the educational process, and it would be possible to identify the changes they made in education.

References

Barczyk, C. C., Duncan, D. G. (2011). Social networking media as a tool for teaching business administration courses. International Journal of Humanities and Social Science, 1(17), 267–276.

Chawinga W.D. (2017). Taking social media to a university classroom: teaching and learning using Twitter and blogs, International Journal of Educational Technology in Higher Education, 14(1).

Chen, Y.-c. (2011). Learning styles and adopting Facebook technology. Proceedings of PICMET '11- Technology Management in the Energy Smart World (PICMET).

De Wever, B., Hämäläinen, R., Voet, M., Gielen, M. (2015). A wiki task for first-year university students: The effect of scripting students' collaboration. The Internet and Higher Education, 25, 37–44.

Dumpit D.Z., Fernández C.J. (2017). Analysis of the use of social media in higher education Institutions (HEIs) using the Technology Acceptance Model, International Journal of Educational Technology in Higher Education, 14(1).

Dzvapatsva, G. P., Mitrovic, Z., Dietrich, A. D. (2014). Use of social media platforms for improving academic performance at Further Education and Training Colleges. South African Journal of Information Management, 16(1).

Galagan, P. (2010). Burp, chatter, tweet: New sounds in the classroom. T + D, 64(7), 26–29.

Gašpar, D., Mabić, M., (2015): Student engagement in fostering quality teaching in higher education, Journal of Educational and Research, MSCER Publishing, Rome, Italy, 5(1), S1, 147-154.

Harris, A. L., Rea, A. (2009). Web 2.0 and virtual world technologies: a growing impact on IS education. Journal of Information Systems Education, 20(2), 137–144.

Irwin, C., Ball, L., Desbrow, B., Leveritt, M. (2012). Students' perceptions of using *Facebook* as an interactive learning resource at university. Australasian Journal of Educational Technology, 28 (7), 1221-1232.

Jones, A. (2015). How Twitter saved my literature class: a case study with discussion. In C. Wankel, M. Marovich, K. Miller, & J. Stanaityte (Eds.), Teaching Arts and Science with the New Social Media, 91–105. Bingley: Emerald Group Publishing

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. Business Horizons, 54(3), 241–251.

Lau W.F. (2017). Effects of social media usage and social media multitasking on the academic performance of university students, Computers in Human Behavior, 68, 286-291.

Madge, C., Meek, J., Wellens, J., Hooley, T. (2009). Facebook, social integration and informal learning at university: 'It is more for socialising and talking to friends about work than for actually doing work'. Learning, Media and Technology, 34 (2).

Manca S., Ranieri M. (2016). Facebook and the Others. Potentials and obstacles of social media for teaching in Higher Education, Computers and Education, 95, 216-230.

Menkhoff, T., Chay, Y. W., Bengtsson, M L., Woodard, C. J., & Gan, B. (2014). Incorporating microblogging ("tweeting") in higher education: lessons learned in a Knowledge Management Course. Computers in Human Behavior. https://www.sciencedirect.com/science/article/pii/S0747563214006815. Accessed 29 May 2018.

Sánchez, R.A., Cortijo, V., Javed, U. (2014). Students' perceptions of Facebook for academic purposes. Journal Computers & Education, 70, 138-149.

Schroeder, J., Greenbowe, T. J. (2009). The chemistry of Facebook: using social networking to create an online community for the organic chemistry laboratory. Innovate: Journal of Online Education, 5(4), 1–7.

Szapkiw, A., Szapkiw, M. (2011). Engaging higher education students through tweeting. In S. Barton, J. Hedberg, &K. Suzuki (Eds.), Proceedings of Global Learn Asia Pacific, 360–364. Melbourne: Global Learn Asia Pacific.

Tang Y., Hew K.F. (2017). Using Twitter for Education: Beneficial or simply a waste of time?, Computers and Education, 106, 97-118.

Terrell, J., Richardson, J., & Hamilton, M. (2011). Using Web 2.0 to teach Web 2.0: a case study in aligning teaching, learning and assessment with professional practice. Australasian Journal of Educational Technology, 27(Special issue), 846–862.

Tess, P. A. (2013). The role of social media in higher education classes (real and virtual)–A literature review. Computers in Human Behaviour, 29(5), A60–A68.

Waycott, J., Gray, K., Clerehan, R., Hamilton, M., Richardson, J., Sheard, J., & Thompson, C. (2010). Implications for academic integrity of using web 2.0 for teaching, learning and assessment in higher education. International Journal for Educational Integrity, 6(2), 8–18.

Wheeler, S. (2010). Open content, open learning 2.0: using wikis and blogs in higher education. In changing cultures in higher education, 103–114. Heidelberg: Springer Berlin.

Tangible User Interfaces for Enhancement of Young Children's Mathematical Problem Solving and Reasoning: A Preliminary Review of Relevant Literature

Lea Dujić Rodić
University of Split
Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture
R. Boškovica 32, 21 000 Split, Croatia
dujic@fesb.hr

Andrina Granić
University of Split, Faculty of Science
Department of Computer Science
R. Boškovica 33, 21 000 Split, Croatia
andrina.granic@pmfst.hr

Abstract. *The aim of this paper is to explore how Tangible User Interfaces (TUIs) have been used in enhancement and learning activities that support mathematical learning of young children. The research is based on the existing scientific literature, models and frameworks. Conducted research implies that TUIs can support mathematical problem solving and mathematical reasoning, but further theoretical and empirical research is needed in order to identify specific properties of TUIs that benefit learning.*

Keywords. Tangible User Interface, Education, Mathematics, Young children

1 Introduction

In the 21 century it is necessary to provide a firm mathematical knowledge and competence to young children so that they may be successful in STEM (Science, Technology, Engineering and Mathematics) disciplines. The growing impact of technologies on our everyday life has reshaped children's models, methods, forms and frames of learning. For this reason, many EU countries have recognized the importance of technologies in education and have renewed the core curricula for education, starting from primary education. For instance, Finland has renewed the national core curricula to update the comprehensive school system to the 2020 requirements to make Finland the number one country of inspiring learning and education (Kimmo, 2017). The reform focuses on three things:

- new pedagogy,
- new learning environments and
- digital learning.

Croatia has also recognized the importance of technologies and their use in early education. *The National Curricula for Early Childhood and Preschool Education* provides a framework to enhance all developmental domains in accordance with each child's abilities (Slunjski *et al*., 2014). The emphasis is on development of competences for lifelong learning which include mathematical competence and digital competence. In this paper we will consider young children to be children aging from 0 to 11 years in accordance to Piaget's stages of cognitive development (Piaget, 1964.). To explore possible learning benefits of learning with and about technologies for young children, we must consider cognitive theories and pedagogical practices which imply what are the most suitable forms of interaction.

Interaction with technology should enhance development of abstract mathematical concepts. One such form of interaction can be, for example, manipulation of objects. We must also examine practical application of above mentioned theories and practices in technology. Most common application are so called Tangible User Interfaces (TUIs), which will be discussed with regards to their use in learning.

In this paper we will review relevant scientific publications that consider the use of TUIs for enhancement of young children's mathematical problem solving and mathematical reasoning. TUIs might prove to be the most beneficial for learning in this domain, because trough linking of physical materials with digital information, children can explicitly see the relationship between concrete and abstract concepts.

2 Cognitive theories and pedagogical practices

If we are to design and create technologies for children, we must consider researches and theories on child development. Jean Piaget's theory of cognitive development is one of the major contributions of the 20th century to developmental psychology and education. Piaget (1964) distinguishes four stages of cognitive development:

1. sensory motor stage (from 0 to 2 years),
2. preoperational stage (from 2 to 7 years),
3. concrete operations stage (from 7 to 11 years),
4. formal operations stage (from 11 to 15 years).

Piaget thought that learning occurs through a process of children's adaptation to their environment. He considered adaptation as an active process which occurs by children's interactions with the world around them; with interactions children gain experience and construct knowledge (Hourcade, 2015).

The idea and the approach claiming that is important to learn through experience rather than being told about it, was laid down by educator Maria Montessori (1912). In her work she observed children in their activities and it helped her design special didactic materials. The materials are learning objects, such as wooden blocks, which allow purposeful learning activities and multi-sensorial interactions. The goal of those learning objects is to maximize children's learning potential (Montessori, 1912). Nowadays, these objects are called *Manipulatives* (Antle, 2013) - physical objects specifically designed to foster learning.

The physical manipulation of objects requires not only physical but also mental activities; in addition, it plays a crucial role in the development of thinking skills in general. Manipulation lightness the cognitive load by simplifying abstract concepts and makes them more accessible to young children (Antle and Wise, 2013).

A modern approach to Montessori didactic materials is provided by (Zuckerman *et al*., 2005) offering a new classification of Manipulatives - *Montessori-inspired Manipulatives* (MiMs). They argue that MiMs foster modelling of more abstract structures. The research suggested that *Digital MiMs*, shown in Figure 1, are engaging learning environments despite them being abstract, and they give children an opportunity to interact with dynamic behaviour at the symbolic level rather than the example level. They showed that Digital MiMs promote group interaction and discussion.

Figure 1. Illustration of MiMs

Consequently, we may assume that concrete physical manipulation of objects might support children's effective or natural learning thus allowing them to focus on the core of the problem that needs to be solved (Marshall, 2007).

3 Tangible User Interfaces and learning

Modern technologies such as mobile phones or computers are "paths" which lead us into the digital world allowing us to seek and give information. However, they are commonly designed for adult users and thus, not always suitable for young children. In the light of the previously mentioned possible benefits of touching and manipulating objects for learning enhancement, we might consider technologies that provide same models of interaction.

Tangible User Interfaces (TUIs) are interfaces that give physical form to digital information. The term TUIs was proposed by Ishi and Ullmer in 1997, who took their inspiration from the ancient counting tool, abacus (Ullmer, 2002). With TUIs we have a seamless integration of corresponding digital representation and digital control (Antle, 2013). TUIs provide the opportunity to reshape educational technologies in accordance to cognitivist theories of learning and can be used to support different types of learning (Markova et at., 2012). Marshall (2007) gave six perspectives on how to regard the use of TUIs with respect to learning: possible learning benefits, typical learning domains, exploratory and expressive activity, integration of representations, concreteness and sensory directness as well as effects of physicality. However, there is still little empirical work that provides evidence to claim that TUIs enhance learning, and moreover, there is a lack of theoretical framework that outlines how different features of TUIs affect learning outcomes, especially when designing guidelines (Antle and Wise, 2013).

A review of TUIs for learning was done by Markova. She proposed a classification framework that attended the important aspects of TUIs such as type of interaction or type of object manipulation in addition to their use in learning, for example explicit learning of facts tended to be mostly achieved using TUIs (Markova, 2012). However, she did not give a particular framework that considered children's use of TUIs for learning.

The model that focusses especially on TUIs for children is the *Child Tangible Interaction* (CTI) *framework* by Antle from 2007. The CTI framework is a conceptual framework for the design of tangibles and interactive spaces which support schemata level knowledge acquisition in children (Antle, 2007, p. 2). This framework focuses primarily on children above the age of four and under the age of twelve and is presented in five themes: Space for Action, Perceptual Mappings, Behavioural Mappings, Semantic Mappings and Space for Friends. These five themes of the CTI framework define vertical research areas for tangible and spatial interaction and children.

With regards to Marshall perspective, Antle and Wise presented the *Tangible Learning Design Framework* in the 2013. This framework is compiled from a taxonomy of five elements that need to be considered when relating TUIs features, interactions and learning. Specifically, those elements are physical objects, digital objects, actions on objects, informational relations and learning activities. For each element, guidelines for design are also provided. Altogether, the taxonomy and the guidelines constitute the *Tangible Learning Design Framework* (Antle and Wise, 2013).

Studies are needed that will explore if the Tangible User Interfaces are truly beneficial for children's learning. It needs to be explored how different interaction styles facilitate the development of children's problem solving skills and, if done in groups, how communication skills improve while they solve problems (Antle, 2013). It is also acknowledged that there is a clear requirement for researches to focus on the long-term effects of learning with TUIs in the classroom settings (Markova, 2012).

4 Research

4.1. Motivation

Mathematics enables problem solving in various areas of science and real life. Traditional maths teaching focuses more on giving procedural knowledge and less on applications of these knowledge in real world (Volk *et al*.,2017). In the last decade the use of technology to support learning has increased. The National Council of Teachers of Mathematics of the USA emphasized the importance of technology in teaching and learning mathematic, since it influences the way maths is taught and enhances learning. (Moller, 2015).

Tangibility and tangible interactions and real life observations seem to play an important role in mathematics, but we still lack formal evidence that tangible enhances learning (Marichal, 2017).

In theory TUIs could be beneficial because children are not explicitly taught about the link between abstract or symbolic content and its concrete physical manifestation (Moller, 2015). We need to understand relations between physical actions and cognitive processes, the link between physical and digital elements through actions, the system feedback and the impact of these elements on the problem solving processes (Marichal, 2017). To better understand these relations and give future guidelines for design of TUIs that might enhance mathematical learning and problem solving, we have conducted a review of prior relevant literature since this is a crucial feature of any scientific study.

4.2. Method

In order to identify relevant scientific publications, a focused structured approach following the suggestions of Webster and Watson (2002) was adopted:

1. search of the set of keywords,
2. refinement of publications by title,
3. quick scan of selected publications by abstracts and
4. detailed analyses of full texts.

The search of the literature was done through the search of the Web of Science database and the ACM Digital Library. It was carried out during a month period from April until May 2018. Set of keywords consisted of words *child/children, tangible/tangibles, touch, interface/interfaces, interaction, and math/mathematics* that were joined with AND and OR operators giving the final search phrase:

"child*"AND[("tangibl*"OR"touch")AND ("interaction"OR"interface*")AND("math*"]

The truncation method was used to cover all variations of keywords; for example, child* was used to search for literature that included the word child or children, while math* was used to search for literature that included the word math or mathematics. The time span of the search was not limited, so we took all time span, and as the first result 80 publications were selected. Figure 2 offers an illustration of the search for the Web of Science database.

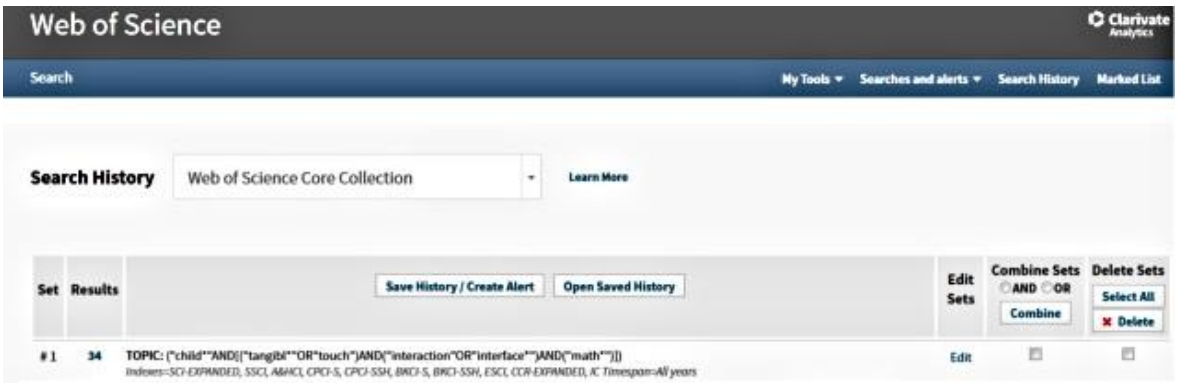


Figure 2. Illustration of the Web of Science search

Publications	Research method(s)	Main findings and conclusions
Saavedra A, Shoemaker A. (2017). DiMBI: An Interface to Connect People to Math's Big Ideas of Patterns and Relations. *In Proceedings of the 2017 Conference on Interaction Design and Children (IDC '17). ACM, New York, NY, USA,* 721-724.	With the intention of promoting math's big ideas in formal and informal settings, we created DiMBI (Discovering Math's multimedia platform developed in Processing that uses reacTIVision to read users' interactions with selected tangible regular polygons (a triangle and a square). The corners of each polygon are linked with specific features that aim to promote transfer learning (colours and music). The features change as the user manipulates the objects through symmetric actions and permutation of corners.	This first version of DiMBI promotes the successful accomplishment of the learning goals. In the short term, we plan to find other physical mechanisms similar to those within a Rubik's cube to promote our assisted flexibility principle. We believe this added structure might assist users in developing and testing meaningful schemata, thereby facilitating deeper understanding of the learning objectives. Once we have this mechanism, we plan to conduct user test sessions in formal and informal learning environments during three stages: interviews, think aloud, and focal groups.

Figure 3. Example of an information related to the selected publication

The publications were selected by title and then further analysed through a detailed process of reading abstracts and full texts. To ensure that the results were up to-date, daily e-mail alert was activated about new entries for the saved search. Some of the publications were excluded and the main criteria were the following: if publications were focused on different age groups of children (children older than 11 years), if they were phycology or socially oriented, if they regarded children with special needs, if they did not emphasize the use of TUIs in mathematical education, if they were not in English, if they were mainly technological publications in the sense that their main focus was engineering, or if they were focused on more traditional use of technology such as traditional input devices.

Finally, 16 scientific publications were selected for this review, providing a time span from 2001 up to year 2017. Publications were then organized in a single table, offering insight into relevant information of the publication itself (author(s), title, journal/conference info, publication year), applied research methods along with main findings and conclusions (see Figure 3). Such organization if selected publications enabled the creation of the main research focus centred table.

5 Results and discussion

The 16 selected publications were organised in a main research focus table, presented in Table 1. In the table every publication appears only once, although some of them could be focus on various concepts. We analysed full texts and based on their aspects of use of TUIs that support young children's mathematical problem solving and mathematical reasoning, we were able to distinguish four main research focus areas:

- design and/or implementations of a learning system or application,
- behavioural or cognitive change,
- enhancement of numerical or arithmetical abilities and
- theoretical review.

Table 1. Main research focus table

Main research focus	Authors
Design and/or implementations of a learning system or application	Marichal et al. (2017) Kubicki et al. (2016) Barendregt et al. (2012) Leong and Horn (2011) Scarlatos and Landy (2001) Khandelwal and Mazalek (2007) Zanchi et al. (2013) Bumbacher et al. (2013) Masood and Hoda (2014) Saavedra and Shoemaker (2017)
Behavioural or cognitive change	Mock et al. (2016) Jong J-T et al. (2013)
Enhancement of numerical or arithmetical abilities	Volk et al. (2017) Sedaghatjou and Campbell (2017) Roberto and Teichrieb (2012)
Theoretical review	Moeller et al. (2015)

Most of the publications focused their research on *design and/or implementations of a learning system or application* (10 out of 16) which makes 62.5% out of all selected publications. Among them, five (Marichal *et al.* (2017), Kubicki *et al.* (2016), Barendregt *et al.* (2012), Leong and Horn (2011) and Khandelwal along with Mazalek (2007)) focused on design and implementation, four focused only on the design (Zanchi et al. (2013), Bumbacher *et al.* (2013), Masood and Hoda (2014) plus Saavedra and Shoemaker (2017)), while just one (Scarlatos and Landy (2001)) focused only on the implementation. This finding implies that researchers tend to focus on a design *and* implementation of a learning system or application thus offering insight into the overall development process.

Among the publications that are focused on the design and implementation, the one that is explicitly related to *Tangible Learning Design Framework* (Antle and Wise, 2013) will be introduced in the following. Specifically, Marichal *et al.* (2017) presented CETA (Ceibal Tangible), a mixed-reality system with tangible interaction for 5-6 year old children. This mixed reality environment for

mathematical learning is inspired in OSMO, a mixed-reality play system for iPads. The authors discussed the design of CETA system in terms of the five element taxonomy proposed in the *Tangible Learning Design Framework*.

CETA is composed of an Android low cost tablet, a mirror, a holder and a set of wooden blocks which play the role of manipulatives. The goal of the game is to learn the concepts of additive composition and the number line representation (*learning activity*). The game narrative is about a robot called Bruno that needs to collect some screws appearing at a certain distance from it. Using the blocks, children must compose the number that matches this distance. Once they put the blocks on the table, the robot will perform an action to pick the screw. *Physical objects* are wooden blocks that become digital manipulatives through markers; each physical block is virtually represented through a virtual block (*digital object*) with the same color and shape on the screen. The most relevant digital object is the main character of the game, robot Bruno. Children control Bruno's actions and movements combining the blocks as illustrated in Figure 4.

As for the *actions on the objects*, children can move the blocks freely, although not all sensible actions for them are sensible or desirable for the system. Regarding *informational relations*, the mappings between physical objects, digital objects and actions, which can be perceptual (physical objects representing digital objects) or behavioral (specific actions on physical objects impacts on digital objects), are considered.

Figure 4. Children playing with CETA.

Going back to the overview of 16 selected papers, with regards to the frequency of publications by year, Figure 5 shows a number of articles that were published in a specific year.

We can notice that the most prominent year was 2017. Five publications were published from 2001 to 2012 and 11 from 2012 to 2017. The number for the latter period is more than 2 times higher than that of the former period indicating an increasing interest on this topic in the last five years.

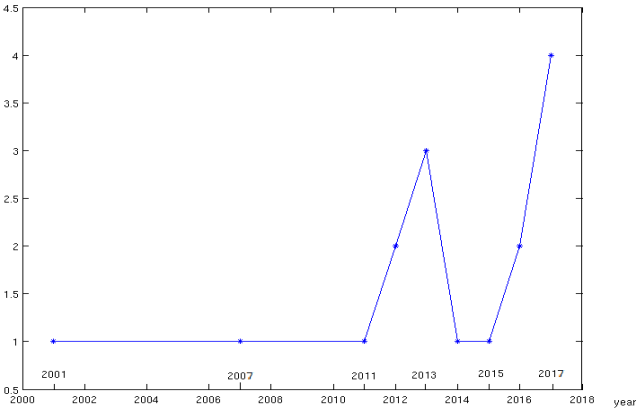


Figure 5. Frequency of publications

Further analyses of selected publications, enabled us to outline three factors that distinguish selected research: *learning topic*, *environment of the conducted experiment* and *form of the tangible object*.

Bearing in mind the *learning topic*, selected publications considered the following five topics: algebra, arithmetic, number line and/or cardinality, time and orientation as well as geometry. Table 2 shows the cross-analysis results of the research focus and learning topic.

Table 2. Frequencies of main research focus and learning topic

	Learning topic				
Main research focus	**Algebra**	**Arithmetic**	**Number line and/or cardinality**	**Time and orientation**	**Geometry**
Design and/or implementations of a learning system or application	**1**	**4**	**2**	**0**	**3**
Behavioural or cognitive change	**0**	**2**	**0**	**0**	**0**
Enhancement of numerical or arithmetical abilities	**0**	**1**	**1**	**1**	**0**
Theoretical review	**0**	**0**	**1**	**0**	**0**

From the achieved results we may conclude that *arithmetic* outnumbers all other learning topics and it is the most frequent one with respect to design and/or implementation of a learning system or application. In contrast, *algebra* and *time and orientation* are learning topics that are least frequent. Reason for such small frequency probably lies in the complexity of algebra. However, algebra is one of the fundamental mathematical branches because it includes everything from elementary equation solving to the study of

highest abstracts. Algebra is a unifying thread of almost all of mathematics and there is a need for further research involving this learning topic with respect to all research focuses.

Among the selected publications, 13 of them have conducted experiments and 11 involved young children. However, in two publications (Masood and Hoda (2014) and Bumbacher *et al*. (2013)) authors pointed out that due to the lack of time they were unable to conduct empirical research with children although initially it was planned. The experiments that involved young children had two distinguishing environments: *formal school environment* and *informal environment*. The experiments that were done in *formal school environment* were Kubicki *et al*. (2016), Barendregt *et al*. (2012), Leong and Horn (2011), Scarlatos and Landy (2001), Mock *et al*. (2016), Jong J-T *et al*. (2013), Volk *et al*. (2017). Experiments done in *informal environment* were Marichal *et al*. (2017), Khandelwal and Mazalek (2007), Sedaghatjou and Campbell (2017) and Roberto and Teichrieb (2012). The results imply that experimental work tend to be conducted in the *formal school environments*. We would like to point out that in the aforementioned publications the authors did not explain why a particular environment was selected for the experiment.

Regarding the *form of the tangible object*, three forms by which the children interacted could be differentiated in selected publications: *manipulatives*, *tablet* and *tabletop* (see Figure 6). The form of the tangible object is closely related with the learning topic of the conducted research implying that it shapes the form of the tangible object. Table 3 shows a cross-analyses of frequencies of use of the form of the tangible object with respect to the learning topic.

From the results we may conclude that tablets and tabletops outnumber the use of manipulatives when learning topic is arithmetic. However, manipulatives seem to have a broader range of use, since they are evenly used in all learning topics and are most frequent form of the tangible object. This is consistent with previously mentioned theories of Piaget and Montessori because of physical and spatial affordances of manipulatives. As stated before, children develop cognitively from physical engagement in reasoning with materials in real world settings.

Table 3. Frequencies of use of form of the tangible object with respect to learning topic

	Learning topic				
Form of the tangible object	**Algebra**	**Arithmetic**	**Number line and/or cardinality**	**Time and orientation**	**Geometry**
Manipulatives	**1**	**1**	**2**	**0**	**2**
Tablet	**0**	**3**	**1**	**1**	**0**
Tabletop	**0**	**3**	**1**	**0**	**1**

6 Conclusion

Finally, we may draw several implications on the influence of Tangible User Interfaces (TUIs) on young children's mathematical problem solving and reasoning. Research indicate that digital systems and game based learning motivates children, engages them into collaboration as well as enhances successful task completion. (Marichal *et al*. 2017), (Volk *et al*., 2017], (Jong *et al.,* 2013). Moreover, the research indicates that children achieve higher learning performances while using TUIs compared to other forms of interaction (Jong *et al*., 2013). One of the advantages of TUIs is that they can provide an external record of previous states and actions and may provide a huge potential for enhancing numerical learning and should thus be explored in future studies (Moeller *et al*., 2015).

However, there is a lack of empirical work which can provide evidence for enhanced learning in mathematics by young children (Chaliampalias, 2016) In a more general note, we conclude that researches and learning specialists need to further their research considering TUIs for learning in order be used in formal school environments (Markova, 2012).

Furthermore, there is a need for a long term exploration of benefits that TUIs may have for enhancement of young children's mathematical problem solving and reasoning. There is also a need for a concrete design framework for the use of TUIs in math education. Such a framework should provide the designers and researchers with design guidelines from cognitive theories and pedagogical practices. These guidelines should point out the appropriate form of tangible object and models of interaction that will facilitate better learning outcomes.

Figure 6. Examples of different forms of a tangible object: **a)** manipulatives, **b)** tablet, **c)** tabletop

References

Antle, A.N. (2007). The CTI framework: informing the design of tangible systems for children. *In Proceedings of the 1st international conference on Tangible and embedded interaction (TEI '07). ACM, New York, NY, USA, 195-202.*

Antle, A.N. (2013) Exploring how children use their hands to think: an embodied interactional analysis, *Behaviour & Information Technology, 32:9, 938-954.*

Antle, A.N., Wise, A. F. (2013). Getting Down to Details: Using Theories of Cognition and Learning to Inform Tangible User Interface Design. *Interacting with Computers, Volume 25, Issue 1, 1 January 2013, Pages 1–20.*

Barendregt, W., Lindström, B., Rietz-Leppänen, E., Holgersson, I., Ottosson, T. (2012). Development and evaluation of Fingu: a mathematics iPad game using multi-touch interaction. *In Proceedings of the 11th International Conference on Interaction Design and Children (IDC '12), 204-207.*

Bumbacher E. et al. (2013). BeatTable: a tangible approach to rhythms and ratios. *In Proceedings of the 12th International Conference on Interaction Design and Children (IDC '13). ACM, New York, NY, USA, 589-592.*

Chaliampalias, R., Chronaki, A., Kamea, A. (2016) Tangible User Interfaces in early year mathematical education: An experimental study. *Hellenic Conference on Innovating STEM Education (2016)*

Hourcade, P.H. (2015). Child-Computer Interaction. *CreateSpace Independent Publishing Platform; First Edition*

Jong J-T., Hong J-C., Yen C-Y. (2013). Persistence temperament associated with children playing math games between touch panel and embodied interaction. *Journal of Computer Assisted Learning (2013), 29, 569–578.*

Khandelwal M., & Mazalek.A. (2007). Teaching table: a tangible mentor for pre-k math education. *In Proceedings of the 1st international conference on Tangible and embedded interaction (TEI '07). ACM, New York, NY, USA, 191-194.*

Kimmo, K. (2017). Finland: Country Report on ICT in Education. European Schoolnet (EUN)

Kubicki, S., Pasco, D., Hoareau, C., Arnaud, I. (2016). Using a tangible interactive tabletop to learn at school: empirical studies in the wild. *Proceeding IHM '16 Actes de la 28ième conférence francophone sur l'Interaction Homme-Machine, pages 155-166.*

Leong, Z. A., & Horn, M. S. (2011). Representing equality: a tangible balance beam for early algebra education. In Proceedings of the 10th International Conference on Interaction Design and Children (IDC '11). ACM, New York, NY, USA, 173-176

Marichal, S. et al. (2017). CETA: designing mixed-reality tangible interaction to enhance mathematical learning. *In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17). ACM, New York, NY, USA, Article 29, 13 pages*

Markova, M.S., Wilson, S., Stumpf, S. (2012). Tangible user interfaces for learning. *Int. J. Technol. Enhanc. Learn. 4, 3/4 (January 2012), 139-155*

Marshall, P. (2007). Do tangible interfaces enhance learning?. *In Proceedings of the 1st international conference on Tangible and embedded interaction (TEI '07). ACM, New York, NY, USA, 163-170.*

Masood, Z, & Hoda R. (2014). Math Tutor: An Interactive Android-Based Numeracy Application for Primary Education. *In Proceedings of the*

Fifteenth Australasian User Interface Conference (AUIC2014).CRPIT Volume 150

Mock, P. et al. (2016). Using touchscreen interaction data to predict cognitive workload. *In Proceedings of the 18th ACM International Conference on Multimodal Interaction (ICMI 2016). ACM, New York, NY, USA, 349-356*

Moeller, K., Fischer, U., Nuerk, H.C., Cress, U. (2015). Computers in mathematics education – Training the mental number line. *Comput. Hum. Behav. 48, C (July 2015), 597-607.*

Piaget, J. (1964). Cognitive Development in Children Development and Learning. *Journal of Research in Science Teaching, 2, 176-186.*

Roberto, R. A., & Teichrieb, V. (2012). ARBlocks: A projective augmented reality platform for educational activities. *2012 IEEE Virtual Reality Workshops (VRW), Costa Mesa, CA, 2012, pp. 175-176.*

Saavedra, A, & Shoemaker, A. (2017). DiMBI: An Interface to Connect People to Math's Big Ideas of Patterns and Relations. *In Proceedings of the 2017 Conference on Interaction Design and Children (IDC '17). ACM, New York, NY, USA, 721-724.*

Scarlatos, L.L., & Landy, S.S. (2001). Experiments in using tangible interfaces to enhance collaborative learning experiences. *In CHI '01 Extended Abstracts on Human Factors in Computing Systems (CHI EA '01). ACM, New York, NY, USA, 257-258.*

Sedaghatjou, M., & Campbell, S.R. (2017). Exploring cardinality in the era of touchscreen-based technology. *International Journal of Mathematical Education in Science and Technology, 48:8, 1225-1239.*

Slunjski, E., et al. (2014). *The National Curricula for Early Childhood and Preschool Education*, MZOŠ, Zagreb.

The Montessori Method by Maria Montessori (1870-1952). Translated by Anne Everett George (1882). New York: Frederick A. Stokes Company, 1912.

Ullmer. B.A. (2002). Tangible Interfaces for Manipulating Aggregates of Digital Information. *Ph.D. Dissertation. Massachusetts Institute of Technology, Cambridge, MA, USA*

Volk, M., Coti, M., Zajc, M., Istenic Starcic, A. (2017). Tablet-based cross-curricular maths vs. traditional maths classroom practice for higher-order learning outcomes. *Comput. Educ. 114, C (November 2017), 1-23.*

Webster, J., & Watson, R. T. (2002). Analyzing the Past to Prepare for the Future: Writing a Literature Review. *MIS Quarterly, (26: 2) pp. xiii – xxiii*

Zanchi, C., Presser, A.L., Vahey, P. (2013). Next generation preschool math demo: tablet games for preschool classrooms. *In Proceedings of the 12th International Conference on Interaction Design and Children (IDC '13). ACM, New York, NY, USA, 527-530.*

Zuckerman, O., Arida, S., Resnick, M. (2005). Extending tangible interfaces for education: digital Montessori-inspired manipulatives. *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05). ACM, New York, NY, USA, 859-868.*

Cluster Analysis of Student's Activities from Logs and Their Success in Self-Assessment Tests

Danijel Filipović, Igor Balaban, Dijana Oreški
University of Zagreb
Faculty of Organization and Informatics
Pavlinska 2, 42000 Varaždin, Croatia
{dfilipovi, ibalaban, dijoresk}@foi.hr

Abstract. *Many educational institutions, especially the ones in higher education, implement Learning Management System (LMS) to assist teachers in teaching and aiding the students in learning. Some of such features available to teachers include logging student's activity and creating custom tests. This paper aims to: (1) Identify relationship between the student's activity on the course and their success on the self-assessment tests; and to (2) Create behavioural profiles of students. In order to do so, the correlation analysis and cluster analysis are performed on a data set retrieved from a course implemented on the Moodle LMS at the University of Zagreb, Faculty of organization and informatics. The research results indicated the relationship between students´ activities and their performance, as well as the profiles of students based on their activities and success on tests.*

Keywords. Cluster analysis, student activity, student success, logs, LMS

1 Introduction

The LMS allows teachers to customize their online or blended courses according to their needs. Those systems are able to record every student's activity, as well as their progress on the course (e.g. their success on tests). "When accessing an LMS with their personal account, students create a digital profile that is saved in the LMS log files" (Kadoić and Oreški, 2018). A log is a list of students' events where a single line represents a timestamp and fields give information about the activity performed (Romero et. al, 2013). These data represent a valuable source for various research activities. In order to gain an insight into such data, the data analytics is performed.

The research is based on the data derived from the blended course "Business Informatics". This course is being conducted at the University of Zagreb, Faculty of Organization and Informatics, at the undergraduate vocational study programme PITUP - Information Technology in Business Application. Business Informatics is taught within four study centres in Croatia: PITUP Varaždin, PITUP Križevci, PITUP Sisak and PITUP Zabok. The data is from the academic year 2017/2018.

The aforementioned course contains several self-assessment tests that students can take. The tests are optional and can be retaken. These tests merely serve as tools that students can use to assess their current knowledge on the subject as a complement to their learning or to improve possible gaps in their knowledge.

We will combine the activity data from students' log files and their success on the self-assessment tests to identify a correlation between students´ activities and their results. Furthermore, we will apply cluster analysis on the data set in order to explore profiles of students' behaviour.

The paper is organized as follows. In the second section we provide a brief review of the previous research on the given topic. Section 3 explains the methodology and section 4 describes the data used in the research. In the Section 5 we present research results and interpretation. We conclude in Section 6.

2 Related work

The LMS provides opportunities for an enhancement of student learning and, consequently, can impact students' final grades (Pislaru and Mishra, 2009). Various authors investigated a relationship between students' LMS activities and their performance.

Asif et al. (2017) used clustering to analyse the students' academic progression in a 4-year bachelor's degree programme. Students were assigned into three clusters based on their average marks: *Low*, *Intermediate* and *High*. They also used decision trees to predict the performance of their students at the end of the semester. Their likely performance, or their final grade, is predicted based on the grades achieved in the first two years.

Alfan and Othman (2005) analysed the performance of students from the University of

Malaysia that took the courses in business and accounting programme. In order to analyse students` performance, the authors performed various analyses. Their aim was to answer several research questions: does previous knowledge of the subjects taught at the study programmes affect their final CGPA, is there a difference in performance between male and female students and whether the students' performance is dependent on their race – Chinese, Indian and Malay?

Bouchet et al. (2013) analysed students' interaction with MetaTutor, the multi-agent intelligent tutoring system. Students were randomly assigned into two test groups - *Prompt and Feedback*, where the students were prompted by the MetaTutor to use specific strategies and were immediately given feedback, and *Control*, where students did not receive any prompts or feedback. The interaction data MetaTutor stored was extracted and clustered by applying Expectation-Maximization (EM) algorithm in Weka 3.

Saarela and Kärkkäinen (2015) analysed the performance of students at Department of Mathematical Information Technology (DMIT) at the University of Jyväskylä in Finland. They specifically focused on performance in Computer Science curriculum. First, they performed a correlation analysis to see if students' grades in certain courses affect their overall success. Furthermore, they performed a cluster analysis and analysed the clusters based on the average grade per course and the average credit score per cluster. Finally, they performed predictive analysis to infer which courses have the highest influence on students' performance.

Talavera and Gaudioso (2004) used clustering to obtain several behavioural profiles of students based on their log files recorded by the LMS. They applied the EM algorithm to dataset which generated 6 different clusters and, to an extent, 6 different behavioural profiles of students.

Gašević, Dawson, Rogers, and Gasevic (2016) explored the extent to which students´ activity influence the prediction of academic success in a blended learning model. LMS data included the usage of the Moodle features: forums, course logins, resources, assignments, book, quizzes, feedback, lessons and chat.

Wang, Lv, Cao and Biao (2017) collected the log data generated by students in the self-learning platform named "Engineering Mechanics Experiment" Autonomous Learning Platform which was designed by their own institution. In their paper, they analysed two set of factors: factors influencing students' landing behaviour and factors influencing students' resource browsing behaviour.

Cantabella et al. (2018) conducted a case study at Catholic University of Murcia in which they analysed the student behaviour in three different modalities (online, on-campus and blended) in the following academic years: 2012/2013, 2013/2014, 2014/2015 and 2015/2016. The analysis was performed with the help of a framework that is built with big data technologies – Apache Hive for storing the student data and Apache Hadoop for performing various statistical analyses. The data was collected from the Sakai LMS, specifically the events triggered by students where extracted and stored in Apache Hive. First, they ranked the tools the students used in Sakai LMS, specifically they measured how much each tool was used for each academic year for each modality. Second, they ranked the total amount of events the students triggered for each modality. Third, they analysed the relationships between the events for each modality. In this instance, the relationship indicates the combination of events students triggered in the same session (Cantabella et al., 2018, page 22). Lastly, they searched for monthly and weekly connection trends - the number of times students visited Sakai LMS - across all years for each modality.

Estacio and Raga Jr. (2017) tried to show whether students' learning behaviour can be extracted from logs recorded by Moodle LMS and visualized accordingly. They also tried to determine if the aforementioned logs can give insight into students' course performance and if their demographic profile affects their level of activity on Moodle LMS. The interesting part of their research is the application of Vector Space Model algorithm to extraction and visualization of students' learning behaviour.

Based on the results of previous research, we have defined the following research goals and research questions. The aim of this research is three-fold:

(i) to identify a relationship between the students' activities within the course and their success on the self-assessment tests
(ii) to create behavioural profiles of students that took the blended course
(iii) to investigate the differences between students from different study centres.

The following research questions were set up:

RQ1: Is there a correlation between students' activities and their success on self-assessment tests?

RQ2: What are the profiles of students' behaviour on the course?

3 Methodology

In order to answer the first research question (RQ1), the correlation analysis was performed on the full dataset, which included students from all PITUP centres. The correlation analysis answered which attributes are connected to the student performance.

With the aim to answer the second research question (RQ2) we applied the cluster analysis. Clustering is a process which groups objects into classes, or groups, of similar objects (Romera & Ventura, 2007). Clustering is a type of unsupervised learning algorithm. According to Baker (2010), clustering is one of five general methods that can be used in educational data mining (EDM). The literature

review revealed the applicability of cluster analysis in analysing students' behaviour and progression. Bovo et. al. (2013) proved that cluster analysis is a great tool for profiling students based on LMS data.

Based on the log data of students' activities from all PITUP centres, we separated the data into four datasets (one for each PITUP centre) and applied the clustering for each one.

Open-source data mining software Weka 3 was used for clustering. As for the algorithm used in the clustering, a simple *k*-Means algorithm was chosen.

4 Data description

In this section, we explain the data used in the research, which was collected from Moodle course log in an actual class.

4.1 Log data

The students' log data was retrieved from the Moodle LMS implementation of the blended course Business Informatics. The original dataset, which was used to extract a new dataset for analysis, contains the attributes described in Table 1.

The log data contains activities from 356 students, of which 33 belong to PITUP Križevci, 74 belong to PITUP Sisak, 124 belong to PITUP Zabok and 125 belong to PITUP Varaždin. In total, this dataset contains 667 174 instances of log data. The earliest instance was created on October 2017, and the latest instance on May 2018.

Table 1. Attributes of the original log dataset

Name	Description
Time	Date and time the instance of log data was created
Full name	First and last name of the user that triggered the event that created the new instance of log data. The "user" can be either student, teacher or the system. In case of a system, a dash symbol (-) is used
Affects user	First and last name of the user on whom the specific event that triggered the event that created the new instance of log data affects. For example, if user *John Doe* views the profile of user *Jane Doe* this attribute will contain value *Jane Doe*. Otherwise, a dash symbol is used (-)
Context	Label of the specific Moodle page of the course that the user was viewing where the event that triggered the creation of the new instance of the log data occurred. It's usually the title of a lesson, title of a test, name of a file that was downloaded, etc.
Component	Category to which the *context* belongs, i.e. *Lesson*, *Test*, *File*, etc.
Name	Name of the event that triggered the creation of the new instance of log data
Description	Detailed description of the event that triggered the creation of the new instance of log data
Source	Source of the event. Only contains the value *web*
IP address	IP address of the user that triggered the event that created the new instance of the log data

Values from the attribute *Component* were used as a set of new attributes for the new dataset. The idea was to count the number of times each value of the attribute *Component* appears for each distinct student. This is where the attribute *Full name* is used – to group the total number of each *Component* by student's name. The instances of log data made by teachers and the system were omitted.

To actually group values of the attribute *Component* by students' names we used PivotTable functionality of Microsoft Excel. Thus, we got one part of the complete dataset – a student's activity dataset.

4.2 Success on self-assessment tests

The students' success on self-assessments tests data was also retrieved from the Moodle LMS implementation of the blended course. The system allows teachers to generate reports for any valid context: attendance, project grades, exam grades, success from online tests, etc.

For this research, we generated a report on each student's success on every self-assessment test, of which there are nine, and their average success on all of the tests combined. Thus, we got a second part of the complete dataset – a students' success on self-assessment tests dataset – whose attributes are explained in Table 2.

Table 2. Attributes of the students' success on self-assessment tests dataset

Name	Description
Student	Contains student's first and last name
SA 1-9	Contains student's success for each individual self-assessment test. Each of the self-assessment test is one attribute in the dataset (hence the 1-9 in the name).

	Contains nominal values ranging from 0,0 to 1,0
SA AVG	Contains student's overall success in all of the self-assessment tests. Contains nominal value ranging from 0,0 to 1,0

4.3 Final dataset

As explained, the dataset derived from the log data, as explained in section 4.1, and the dataset explained in section 4.2 are two parts of a complete dataset that was used in this research. To analyse the data, the datasets where converted into CSV files where the values are separated by semi-colons (;). Then, a simple Python script was written that merged the columns by student's full name and wrote it into a third CSV file. That CSV file is now the complete dataset used in the research.

However, due to the correlation analysis results explained in section 5, we removed the attributes *Selection*, *Folder*, *Records* and *SA 6* because they had no major effect on other attributes.

The attributes of the complete dataset are described in Table 3.

Table 3. Attributes of the complete dataset

Name	Description
Student	Contains student's first and last name.
File	No. of times the student downloaded a file or viewed it in-browser.
Forum	No. times the student viewed the forums section of the course.
Student report	No. times the student viewed his/her or other student's Moodle LMS profile.
Lesson	No. of times the student viewed lessons on the course.
File upload	No. times the student uploaded a file.
Link	No. of times the student clicked on an outgoing hyperlink.
Overview report	No. of times the student viewed their overview report in the gradebook.
Page	No. of times the student viewed any of the course's pages
System	No. of other general user activities
Test	No. of times viewed or took the online test
Homework	No. of times the student uploaded a homework file. In the Moodle LMS, this could relate to an actual homework (seminar papers, source code, etc.) or a file the student created as part of an exam
SA 1-5, 7-9	Contains student's performance for each individual self-assessment test. Each of the self-assessment test is one attribute in the dataset (hence the 1-9 in the name). Contains nominal values ranging from 0,0 to 1,0.
SA Average	Contains student's overall success in all of the self-assessment test. Contains nominal value ranging from 0,0 to 1,0.

4.4 Datasets per PITUP centre

For this research we wanted to perform the cluster analysis for each PITUP centre individually. That means separating the complete dataset into four smaller datasets which correspond to one of the PITUP centres.

To do this, we used the Moodle LMS to extract the list of students for each centre and created another Python script that takes the complete dataset and copies each row of data into one of four corresponding new datasets. After that, the data was prepared for the analysis.

5 Research results

The most important results from the log file analysis are presented in this section.

5.1 Correlation analysis

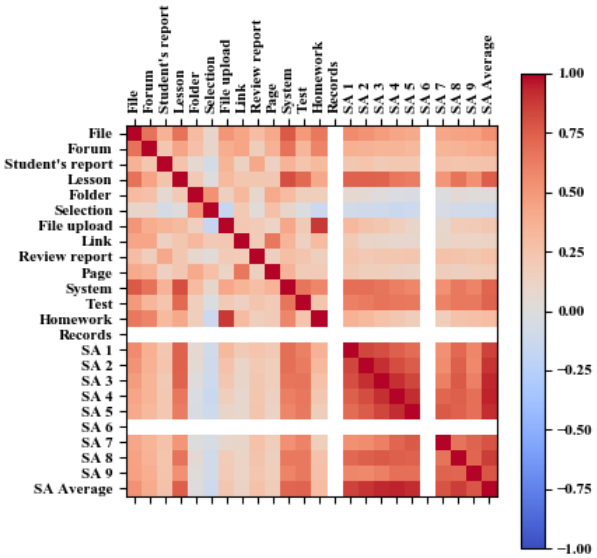


Figure 1. Correlation matrix

To determine whether there is a relationship between students' activities and their success on self-assessment tests, we computed a correlation matrix of the complete dataset. For the computation, another Python script was created, with the help of *Numpy*, *Pandas* and *Matplotlib* libraries, that generated a visualisation of the correlation matrix (Figure 1).

The correlation matrix shown in Figure 1 uses a heatmap to show direction and strength of the relationship and to display how much the two attributes correlate to each other. The redder the point (or square) of intersection is, the greater the correlation is. Opposite of that, the bluer the point of intersection is, the lower the intersection is.

From the correlation matrix we can see that attributes *Records* and *SA 6* seem to be excluded from the computation of the matrix. This is because for every instance in the dataset the values of those attributes were zero. Thus, we removed those attributes from the further analysis since they do not provide any information.

For other attributes regarding students' activities we can see that there are both weak and strong correlations with the attributes regarding students' success on the self-assessment tests. The attributes *Folder* and *Selection* have close to zero correlation with the success attributes, but due to their nature we determined that they can be ignored and removed from the complete dataset. On the other hand, attributes *Lesson*, *Test* and *System* have strong correlations with the attributes which indicate student's success.

High correlation of attributes *Lesson* and *Test* with the success attributes can be explained as follows:

- While studying, students use the self-assessment tests to test their knowledge. If they answered a question wrongly or didn't know the answer, they would recheck the lessons to see what the correct answer was and fill the gaps in their knowledge
- Attribute *Test* basically counts the number of times a student took the test. Since the tests can be repeated the number is higher and therefore the stronger the correlation with the success attributes.

We concluded that there is a correlation between students' activities and their success on the self-assessment test.

5.2 Cluster analysis

In order to answer the second research question, we have performed cluster analysis. Each dataset that corresponds to its PITUP centre was imported into Weka and clustered using the *k*-Means algorithm. At first, we performed the clustering multiple times with a different number of clusters each time to determine the optimal number.

It was decided there should be 4 clusters. It was observed that every additional cluster just seemed to provide nearly the same data as one of the existing clusters. Clusters will be labelled *A*, *B*, *C* and *D*. Cluster A would point towards the best cluster, and Cluster D would point towards the worst cluster.

5.2.1 PITUP Križevci

Table 4 shows the number of students for each cluster of PITUP Križevci dataset as well as the percentage they take from the total number of students for this centre. Table 5 shows the centroids for each cluster dataset. This dataset has the least number of students in all datasets.

Table 4. Number and percentage of students per cluster from PITUP Križevci

Cluster	Count	%
A	5	15.15%
B	5	15.15%
C	8	24.24%
D	15	45.45%
Total	33	100.00%

Table 5. Centroids for PITUP Križevci

Attribute	Clusters			
	A	B	C	D
File	12.00	6.80	4.75	2.13
Forum	13.80	9.80	11.75	5.67
Student report	4.80	4.00	3.88	0.40
Lesson	1979.2	1625.8	840.00	211.73
File upload	0.00	0.0	0.00	0.00
Link	0.40	0.20	0.75	0.47
Overview report	0.20	0.00	0.38	0.00
Page	2.60	2.00	3.00	1.67
System	359.60	271.00	145.13	43.93
Test	448.80	237.20	54.00	0.53
Home-work	0.00	0.00	0.00	0.00
SA 1	0.81	0.70	0.52	0.00
SA 2	0.89	0.81	0.08	0.00
SA 3	0.83	0.73	0.00	0.00
SA 4	0.82	0.53	0.00	0.00
SA 5	0.82	0.52	0.00	0.00
SA 7	0.87	0.00	0.00	0.00
SA 8	0.80	0.40	0.21	0.00
SA 9	0.43	0.0	0.09	0.00
SA Average	0.70	0.41	0.10	0.00

As it is shown from the data in the presented tables, we can see that only a small number of students

(clusters A and B) actively took self-assessment tests while the rest (clusters C and D) tried a few times or did not try at all.

Students from cluster A had a clear lead in their activity on Moodle LMS and their success on self-assessments tests. It is worth noting that *SA 9* seems to be ruining their *SA Average*.

Students from cluster B were almost as active as cluster A, but their average success on self-assessment tests is mediocre at best. They tried their best on the first three self-assessment tests, but slowly stopped trying for the other five.

Students from cluster C weren't as active as students from clusters A and B, but are still more active than the students from cluster D. However, their success on self-assessment tests were quite low. They tried with the first test but seemed to give up on every other test except for self-assessment test 8 (*SA* 8).

Students from cluster D were the least active students and they haven't even tried taking the self-assessment tests. Students from this cluster were also in the majority of this PITUP centre. In numbers, 15 out of 33 students from PITUP Križevci dataset had no interest in taking self-assessment tests.

5.2.2 PITUP Sisak

Table 6 shows the number of students for each cluster for PITUP Sisak, as well as the percentage they take from the total number of students from this centre. Table 7 shows the cluster centroids. This dataset has second lowest number of students in all datasets

Table 6. Number and percentage of students per cluster from PITUP Sisak

Cluster	Count	%
A	12	16.22%
B	13	17.57%
C	13	17.57%
D	36	48.65%
Total	74	100.00%

Table 7. Centroids for PITUP Sisak

Attribute	Clusters			
	A	B	C	D
File	17.42	10.31	8.69	2.61
Forum	49.33	9.00	10.23	7.75
Student report	10.75	3.31	3.85	0.97
Lesson	3226.6	2448.6	1242.4	190.28
File upload	0.00	0.00	0.00	0.00
Link	0.67	0.23	0.85	0.17
Overview report	0.42	0.00	0.00	0.00
Page	3.92	1.77	2.46	0.56
System	638.67	295.69	179.77	49.36
Test	764.58	264.85	53.85	0.08
Home-work	0.00	0.00	0.00	0.00
SA 1	0.81	0.69	0.47	0.00
SA 2	0.85	0.83	0.32	0.00
SA 3	0.84	0.83	0.05	0.00
SA 4	0.81	0.54	0.00	0.00
SA 5	0.87	0.47	0.00	0.00
SA 7	0.91	0.00	0.00	0.00
SA 8	0.73	0.53	0.04	0.00
SA 9	0.70	0.16	0.00	0.00
SA Average	0.72	0.45	0.10	0.00

The results indicated that students from cluster A were most active and most successful on the self-assessment tests. An interesting thing to note was the value of their attribute *Test*. It was almost 3 times bigger than the value in cluster B, which was the second largest value of all clusters. A possible interpretation could be that students have been continuously retaking the self-assessment tests until they had a high enough score, either as a proof to themselves that they learned the specific topic or simply as a type of self-accomplishment.

Students from cluster B were somewhat less active than the students from cluster A, but they still had some activity. As for their success in self-assessment tests, they were mediocre at best. Similarly to PITUP Križevci, they started strong but gradually receded.

Students from cluster C had a decent activity. They were more active than their PITUP Križevci counterparts. However, their success on self-assessment tests was weak.

Students from cluster D were the least active. From their success on self-assessment tests one could assume they haven't given those tests any thought. Almost half of students from PITUP Sisak dataset belonged to this cluster.

5.2.3 PITUP Zabok

For PITUP Zabok dataset, Table 8 shows the number of students for each cluster as well as the percentage they take from the total number of students from this centre, and Table 9 shows the centroids for the computed clusters. This dataset has the second greatest number of students in all datasets, only one less than PITUP Varaždin dataset.

Table 8. Number and percentage of students per cluster from PITUP Zabok

Cluster	Count	%
A	11	8.87%

B	15	12.10%
C	14	11.29%
D	84	67.74%
Total	124	100.00%

Table 9. Centroids for PITUP Zabok

Attribute	Clusters			
	A	B	C	D
File	11.91	9.67	12.29	4.50
Forum	30.45	14.20	19.43	8.75
Student report	4.73	3.13	4.8571	1.21
Lesson	2591.3	2210.1	1745.1	243.85
File upload	0.00	0.00	0.00	0.00
Link	0.45	0.67	1.21	0.29
Overview report	0.36	0.27	0.21	0.06
Page	0.91	1.40	2.21	1.00
System	361.45	277.00	268.71	62.54
Test	422.36	257.67	91.57	2.58
Home-work	0.00	0.00	0.00	0.00
SA 1	0.76	0.68	0.51	0.03
SA 2	0.82	0.87	0.53	0.00
SA 3	0.84	0.84	0.24	0.00
SA 4	0.82	0.69	0.03	0.00
SA 5	0.83	0.68	0.00	0.00
SA 7	0.71	0.06	0.00	0.00
SA 8	0.80	0.42	0.28	0.00
SA 9	0.52	0.10	0.00	0.00
SA Average	0.68	0.48	0.18	0.00

Again, students from cluster A had the most activity, but not by a wide margin when comparing them to students from cluster B. They were also most successful on self-assessment tests. Like in PITUP Križevci dataset, their *SA Average* was slightly ruined by *SA 9*.

Students from cluster B were slightly less active than students from cluster A, but their average success on self-assessment tests were mediocre just like in previous datasets.

Cluster C students had a quite decent activity, slightly less than students from cluster B. Their success on self-assessment tests is quite low but it can be seen that students tried in the first two test and gave up on the other ones.

Lastly, students from cluster D had the least activity and made almost no attempts to solve the self-assessment tests. Also, over half of students from PITUP Zabok dataset belonged to this cluster, which was much greater than in other datasets.

5.2.4 PITUP Varaždin

Table 10 shows number of students per cluster and the percentage they take from the total number of students from this centre. Table 11 shows the cluster centroids for this dataset. This dataset has the greatest number of students in all datasets.

Table 10. Number and percentage of students per cluster from PITUP Sisak

Cluster	Count	%
A	35	28.00%
B	26	20.80%
C	30	24.00%
D	34	27.20%
Total	125	100.00%

Table 11. Centroids for PITUP Varaždin

Attribute	Clusters			
	A	B	C	D
File	25.60	18.38	21.63	6.65
Forum	63.37	49.69	55.03	27.12
Student report	24.43	9.15	13.30	3.88
Lesson	3083.3	2716.9	1591.7	424.82
File upload	4.69	6.62	7.13	2.94
Link	1.17	0.81	1.33	0.79
Overview report	0.49	0.19	0.17	0.03
Page	2.63	1.62	2.63	0.88
System	572.63	406.23	373.23	111.79
Test	470.46	325.12	122.77	18.65
Home-work	35.37	34.92	42.40	15.03
SA 1	0.80	0.72	0.56	0.10
SA 2	0.91	0.84	0.41	0.00
SA 3	0.89	0.83	0.16	0.00
SA 4	0.82	0.58	0.00	0.00
SA 5	0.87	0.39	0.00	0.00
SA 7	0.90	0.00	0.00	0.00
SA 8	0.80	0.47	0.13	0.00
SA 9	0.68	0.21	0.03	0.00
SA Average	0.74	0.45	0.14	0.01

Cluster A contains the most active students which were also most successful on self-assessment tests.

Similar to previous datasets, the value in attribute *SA 9* is slightly ruining the value of attribute *SA Attribute*.

Students from cluster B were only slightly less active comparing to students from cluster A. But, as it was the case with B clusters in previous datasets this cluster also had a mediocre average success on self-assessment tests. Students from cluster B seemed to be the students that start strongly in first few tests, but then stop giving much effort into other ones.

While students from cluster C had the second least activity, they were still decently active. However, their successes on the self-assessment tests were weak.

As for the students from cluster D, it's interesting to note that they were the most active when comparing them to cluster D students from other datasets. But they still seemed to not give much effort to taking self-assessment tests as their successes on them were mostly zero.

What's even more interesting is that there were the same number of students belonging to clusters A and D. In previous datasets, there were just slightly-under-half or slightly-over-half of total number of students belonging to cluster D. Even PITUP Zabok datasets, which had the same number of students, had 50 more students in cluster D than this dataset.

6 Conclusion

From the data presented in this paper and its interpretation we can conclude that the more active a student is on the LMS the more likely he/she is going to take the self-assessment test with success. Also, with the cluster analysis we concluded that in PITUP centres Križevci, Sisak and Zabok most of the students are profiled as non-active students. This could mean that students' activities and performances on the self-assessment tests are also influenced by location and/or different studying terms in PITUP centres outside Varaždin.

However, we need to emphasize that students in Varaždin are given also live classes of 30 hours (2 hours per week), and are mostly full-time students. However, students from other three centres are all part-time students, most of them travel to get to classes, and for them traditional classes are held only twice (5 hours each).

Moreover, students in Varaždin are given scores for their activity in LMS and other activities in the course throughout the semester, while students in other centres are not rewarded for any of the extra activities (including self-assessment). Instead, the self-assessment tests are just a feedback mechanism that help them to prepare for the final exam.

There are several possible directions for future research. We could take students' performances on the self-assessment tests and their course grades to determine if the former influences the latter. Cluster analysis could also be applied for the next academic year, or next several academic years, and compare them to see if the next generations of students will be more, less, or equally active on LMS and successful on self-assessment tests. This would be beneficial in case of PITUP centres where less active students are a majority (Križevci, Sisak and Zabok).

7 References

Asif R., Merceron A., Ali S.A. & Haider N.G., Analyzing undergraduate students' performance using educational data mining, *Computers & Education* (2017), doi:10.1016/j.compedu.2017.05.007

Baker R.S.J.d. (2010). Data Mining. In McGaw, B., Peterson, P., Baker, E. (Eds.) *International Encyclopaedia of Education (3rd edition)*. Oxford, UK: Elsevier.

Bouchet, Harley, J., Trevors, G., & Azevedo, R. (2013). Clustering and Profiling Students According to their Interactions with an Intelligent Tutoring System Fostering Self-Regulated Learning. JEDM | Journal of Educational Data Mining, 5(1), 104-146. Retrieved from https://jedm.educationaldatamining.org/index.php/JEDM/article/view/32

Bovo, A., Sanchez, S., Héguy, O., & Duthen, Y. (2013, September). Clustering moodle data as a tool for profiling students. In e-Learning and e-Technologies in Education (ICEEE), 2013 Second International Conference on (pp. 121-126). IEEE.

Cantabella M., Martínez-España R., Ayuso B., Yáñez J. A., Muñoz A. (2018). Analysis of student behavior in learning management systems through a Big Data framework, *Future Generation Computer Systems*, 90, Pages 262-272, doi: https://doi.org/10.1016/j.future.2018.08.003.

Ervina Alfan, Md Nor Othman, (2005) Undergraduate students' performance: the case of University of Malaya, Quality Assurance in Education, Vol. 13 Issue: 4, pp.329-343, https://doi.org/10.1108/09684880510626593

Gašević, D., Dawson, S., Rogers, T., & Gasevic, D. (2016). Learning analytics should not promote one size fits all: The effects of instructional conditions in predicting academic success. The Internet and Higher Education, 28, 68-84.

Kadoić, N., & Oreški, D. (2018, January). Analysis of Student Behaviour and Success Based on Logs in Moodle. In 41st International Convention on Information and Communication Technology, Electronics and Microelectronics MIPRO 2018.

Pislaru, C., & Mishra, R. (2009, April). Using VLEs to support student centred learning in Control Engineering Education. In Proc. 5th Int. Conf on Multimedia and Information and Communication

Technologies (m-ICTE 2009), University of Lisbon, Portugal (pp. 22-24).

Romero, C., & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33, 135-146.

Romero, C., López, M. I., Luna, J. M., & Ventura, S. (2013). Predicting students' final performance from participation in on-line discussion forums. Computers & Education, 68, 458-472.

Rosalina Rebucas Estacio, Rodolfo Callanta Raga Jr, (2017) "Analyzing students online learning behavior in blended courses using Moodle", *Asian Association of Open Universities Journal, Vol. 12* Issue: 1, pp.52-68, https://doi.org/10.1108/AAOUJ-01-2017-0016

Saarela, M., & Kärkkäinen, T. (2015). Analysing Student Performance using Sparse Data of Core Bachelor Courses. JEDM | Journal of Educational Data Mining, 7(1), 3-32. Retrieved from https://jedm.educationaldatamining.org/index.php/JEDM/article/view/JEDM056

Talavera, L., & Gaudioso, E. (2004). Mining Student Data To Characterize Similar Behaviour Groups In Unstructured Collaboration Spaces. In *Workshop on artificial intelligence in CSCL. 16th European conference on artificial intelligence* (pp. 17–23).

The University of Waikato (2017). Weka (3.8.2.) [Data mining software]. Retrieved from https://www.cs.waikato.ac.nz/ml/weka/

Wang Jie, Lv Hai-yan, Cao Biao and Zhao Yuan, Application of educational data mining on analysis of students' online learning behavior, *2017 2nd International Conference on Image, Vision and Computing (ICIVC)*, Chengdu, 2017, pp. 1011-1015. doi: 10.1109/ICIVC.2017.7984707. Retrieved from https://ieeexplore.ieee.org/document/7984707/

Emerging Trends in ICT

Framing the Digital Transformation of Educational Institutions

Martina Tomičić Furjan, Vjeran Strahonja, Katarina Tomičić-Pupek
University of Zagreb, Faculty of Organization and Informatics
Department of Organization, Department of Information Systems Development
Pavlinska 2, Varaždin, Croatia
{martina.tomicic, vjeran.strahonja, katarina.tomicic}@foi.hr

Abstract. *Digital transformation, as a new trend of organizational development, changes the way work is being done from the technological and operational point of view. Educational institutions (EI) need to adapt to this new paradigm and other upcoming trends in order to give the new generations right knowledge and skills to cope with the new challenges that are coming with these trends as well as to make their own work easier, faster, more efficient and effective. Methodological frameworks should help guide organizations to digitally transform and they can be used with the aim of making the change right. In this paper we analyse two well- known digital transformation frameworks in order to frame the digital transformation process of EIs within them with the aim to guide them to achieve the goal of increasing their digital maturity.*

Keywords. Digital transformation, Methodology, Framework, Education

1 Introduction

Digital transformation is a trend that was already addressed in the 1990s and again in the mid-2000s, with a focus on the business process digitalization (Digital Transformation: History, Present, and Future Trends, 2016).

The increasing digitization of business processes by new ICT requires developing new business models for organizations to remain competitive on global markets. When applying this fact to the non-profit sector, especially to educational institutions, this means that EIs need to transform their activities in order to achieve goals related to the expectations: of their stakeholders in terms of realizing their mission, of providing high-quality public services, of delivering key performance indicators regarding the cost- mission implementation and of other non-profit specific targets.

In Croatia, the digital maturity of primary and secondary schools differs from region to region, from school to school, and its level can be measured by five dimensions: Leadership, planning and management, ICT in learning and teaching, Development of digital competences, ICT culture and ICT infrastructure (Begičević et al, 2017). With the aim to research the current state, introduce new EI processes and improve existing ones, introduce ICT related improvements related to the operations in schools as well as the teaching processes, including the support of contemporary ICT, a project e-Schools: Establishing a System for Developing Digitally Mature Schools (pilot project)" was performed from 2015 to 2018 (e-Schools, 2018).

The project had two main parts: the investigation and informatization of operational / business processes of schools and introducing an ERP system for better data processing and the investigation and informatization of teaching and learning processes including computer equipment and digital educational content, wireless Internet and the education of and support to teachers in the implementation of e-classes.

The goal of our paper is to investigate what are the main determinants of digital transformation and how EI's can implement those determinants in their digital transformation.

2 Insights about publications related to Digital transformation methodologies

This research reviewed available literature on Digital transformation methodologies in order to summarize existing findings in this research field. Special focus has been put on the methodologies published in scientific publications. The next section presents the methodology of literature gathering and conducting a short insights analysis.

The literature gathering process started with the identification of the relevant databases for this research. Relevant databases that were selected are Scopus and Web of Science (WoS). The search was conducted based on the combination of keywords "Digital transformation" and "methodology". The

search has been conducted in April 2018 and it received results containing 53 hits in Scopus and 27 hits in WoS. The results showed duplicated entries as 17 of 27 publications from WoS were found also in Scopus, i.e. in both databases. Main trends were recognized as same and since our intention is to show only trends in this field, we put our focus on Scopus publications due to the greater number of publications in total.

The results show that although digital transformation was used as a term since 1979, its current meaning of transforming businesses by relying on contemporary ICT has a growing trend since 2014.

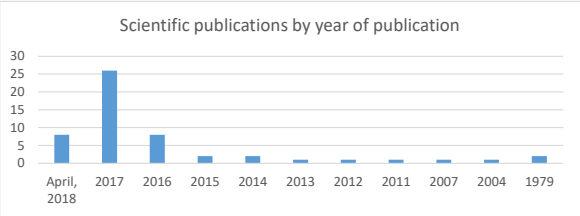


Figure 1. Digital transformation methodology publication by year of publication (SCOPUS)

This short publication gathering study also gave insights about subject area in which digital transformation is a topic of research. The following figure shows that the most publications were related to the computer science and to Business, Management and Accounting. This reveals the digital transformation domain, namely its orientation towards business and IT.

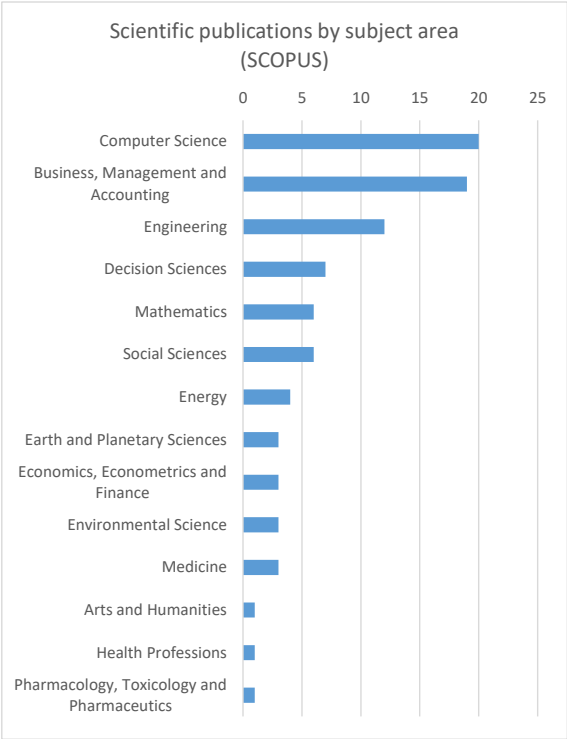


Figure 2. Digital transformation methodology publication by subject area (SCOPUS)

For a more detailed topic refining we analysed keywords and created a "wordcloud" visualization showing keywords that occurred 3 or more times (Figure 3). The keywords analysis implies the orientation of DT to various initiatives and can be used as a map or possible pool of targeting measures or goals confirming our understanding that digital transformation is not triggered exclusively by technological possibilities, but moreover it is being inspired by business needs and then supported by appropriate technologies afterwards.

Figure 3. Digital transformation methodology publication by keywords (SCOPUS)

Our short insights include limited literature insights in only one high-quality database. Due to that fact, this literature review cannot be considered as comprehensive but it shows indications about this trend of going digital. It also gives some light about the authors' understanding and interpretation of this new paradigm.

Also it shows that the number of publications has grown in the past few years which is an indicator that the field is still in its developing stage, meaning that the maturity is not yet reached. Further research in this field could investigate types of publications and the more detailed report on case studies about digital transformation.

3 Research Methodology

Digital transformation frames should help companies digitally change their work and / or the results of that work. Frames are methodologically developed by theoreticians and practitioners and their aim is to provide guidance and guidance through the transformation (Tomičić Furjan & Kuhar, 2018).

Schallmo et al. (Schallmo et al., 2017) defines digital transformation as follows: "The DT framework includes the networking of actors such as businesses and customers across all value-added chain segments, and the application of new technologies. As such, DT requires skills that involve the extraction and exchange of data as well as the analysis and conversion of that data into actionable information. This information should be used to calculate and evaluate options in order to enable decisions and/or initiate activities. For increasement of performance and reach of a company,

DT involves companies, business models, processes, relationships, products, etc."

Some methodological dimensions about how EIs could take part in digital transformation or guidelines on how such organizations could digitally transform their processes is in focus of existing research. Many concepts of digital maturity of EIs have been identified and translated into different frameworks (Begičević et al, 2017, p. 360) which are aimed to be implemented for achieving higher level of digital fluency in the fourth industrial revolution (WEF, 2017).

With the intention to encourage openness in education innovation, especially in introducing new approaches, models and technologies that support the increasement of digital literacy (WEF, 2017), we analyse two well-known digital transformation frameworks from the industrial sector. Replication of frameworks that have been developed at or well accepted in practice into the educational sector could reveal possible scenarios that are feasible for digital transformation of EIs.

The first selected framework is the Digital transformation compass, introduced by Westerman, Bonnet and McAfee (Westerman et al., 2014, p. 173). The Compass, with defined activities in areas that cover the whole life cycle of digital transformation, is based on findings from a global survey on hundreds of companies across different industries conducted by the authors, is very practice oriented, and therefore was chosen for analysis in this paper.

The second framework is the Business model canvas, developed by Osterwalder (Osterwalder et al., 2014, p. 16). Business model canvas as a tool is not exclusively reserved for framing the digital transformation. Its aim is to help develop a company in any form of digital change. This framework was selected because of its wide implementation in practice of over 5 million users (Business Model Canvas), as well as for the fact that authors of this paper have experience in the use of the Canvas in other projects (Digitrans method framework, 2018).

These frameworks will act as our frames for specifying EI's digital transformation by identifying key determinants aiming to achieve the goal of increasing EI's digital maturity. The two selected frameworks are presented in section 3.1 and 3.2, while the Framing of EI's digital transformation is given in section 4.

3.1 Digital transformation compass

The digital transformation compass, as a framework that should help organizations digitally transform, was developed by Westerman, Bonnet and McAfee (Westerman et al., 2014, p. 173). The methodology described in this framework was developed based on experiences and best practices established by the 'digital masters'. Digital masters are large companies in traditional industries that use digital processes to improve their business.

Figure 4. Digital transformation compass
Source: (Westermann et.al., 2014, p.174)

The compass, (shown in Figure 4) leads organization through steps to be performed, grouped in four areas, each one of them dealing with the different dimension of the transformation:

1. FRAME - Framing the digital challenge is the first area of digital transformation and includes activities of building and rising awareness, especially that of the leaders in the company, of digital potentials and of how they can translate the AS-IS state into TO-BE vision.
2. FOCUS - Focusing investment is the second area, and it means that the organization has to get the transformation funded and sponsored by the right people/sources and that it has to be defined by the right activities and governance.
3. MOBILIZE – Mobilizing the organization includes organizational and cultural change on all levels by all employees.
4. SUSTAIN – Sustaining the digital transition deals whit the question of how to stay on top as a digital master.

Each area has three steps which lead the company through the process of digital transformation. The companies may have already taken some of the digital initiatives, and they have to determine what they are still missing and concentrate on doing that.

3.2 Business model canvas

Business model canvas was developed by Osterwalder and it is based on his earlier work on Business Model Ontology. It is a template for capturing value for the organization (Osterwalder et al., 2014, p. 16).

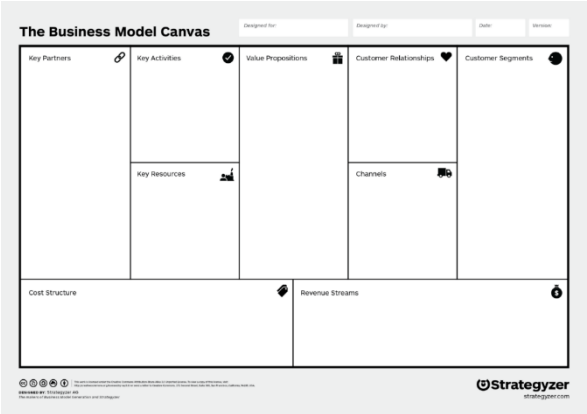


Figure 5. Business model canvas
Source: strategyzer.com

Business model canvas (shown in Figure 5) consists of nine elements that represent the companies' way of creating, delivering and capturing value (Osterwalder et al., 2014, p. 17).

4 Framing EI's digital transformation

In this chapter, previously demonstrated frameworks have been applied in the context of EI as our theoretical proposal of possible scenarios of DT. In Table 1. the analysis was made on how to use the digital transformation compass as a guide for educational institutions which want to digitally transform. The 12 steps in digital transformation compass are listed and explained in column 1, and how to interpret them and perform in EIs in column 2.

In Table 2. the analysis was made on how to create a new business model in educational institutions that want to digitally transform using the business model canvas. The 9 elements from the canvas are listed and explained in column 1, and what they represent in digitally transformed EIs in column 2.

Table 1: Framing EI's digital Transformation according to DT Compass

Step in the digital transformation compass (Westerman et al., 2014, p. 174)	**Implementation in EIs**
Build awareness of digital potential, threats and opportunities to the leaders in the organization – they have to feel the need to change	Principal, head of any organisational unit, a visionary teacher, EI advisory agencies, ministry or any educational support organization - whoever the initiator of the change is, needs to get the other leaders and decision makers on board. Organizing internal workshops, webinars, idea dissemination events and conferences is a good way to present the advantages of digital technologies when further implementing them to the educational sector.
Define your starting point related to strategic assets and digital competencies	The most important strategic asset of a school are teachers with their capabilities and competencies to implement the change, and together with the physical assets like appropriate classrooms, networks, computers etc., as well as the intangible ones, like organizational culture and fundamental values, they all build the starting point of the digital change. It is important to know where the EI stands now in order to determine where it wants and can go.
Create a shared vision, digital and known by all leaders in the organization – define the *what* you want to achieve	With the starting point in mind, the vision of the digitally changed EI has to be created and communicated to all participants already on board. Their need to understand, approve and live the vision every day, in order to distribute the idea further to other employees.
Translate your vision into action, by defining strategic goals and a roadmap of activities to be performed - define the *how* you want to achieve it	When the vision is created and the EI leaders know where they want to be in the future, it is time to define how to get there. This is done by operationalising the vision through strategies, goals and activities. The goals have to be related to EI's customer (students, their parents, future employees and the society) and by keeping its value change in mind as well as the improvement of operations which is done by EI's employees themselves. Both of the mentioned "stakeholders" are equally important, and therefore they should be involved in the process of defining goals and in the activities for achieving them.
Build your governance through mechanisms to steer digital transformation	Once the roadmap is set and everybody in the EI knows what to do, it is important to ensure that all the activities "lead" to the same direction and that they "stay" on the proposed path. For that purpose, it is important to define additional steering roles – responsible persons for the achievement of goals, as well as shared units that should help dealing with the joint challenges.

Fund the transformation by building a balanced portfolio of investments	Although EIs are non-profit organizations, they should shape the next generation society makers and therefore the education system improvement should not depend on costs, digital transformation always brings some costs. Usually, in EIs, the costs are related to the building of digital infrastructure, the education of teachers and administrative stuff to use new technologies, for digital and online teaching materials, as well as all related and occurring costs during the transformation. Due to that it is crucial to involve as many stakeholders as possible and to get them engaged in the transformation process regardless of their local, regional, national and international level of operation.
Signal your ambitions, define the advantage of going digital to every employee	The top team of leaders already on board has now a significant task to get everybody else on the "ship". This should be done at the earliest stage of the transformation as possible, through every available channel – meaning at school board meetings, internal educations and other formal events as well as through teachers' chats, during the morning coffee breaks and other informal happenings. Every teacher, accountant, sociologist, psychologist or other expert in the EI needs to know what is supposed to be done, understand what are the benefits for them personally from the new way of getting the job done and how their change contributes to overall digital strategy operationalization.
Earn the right to engage from the employees that co-create the solution and make it happen	All employees will be familiar with the new digital change, but that does not mean they all are going to follow and live the vision, strive to achieve the goals and perform the activities that are set. It is important to find the champions among employees, identify good examples and practices and encourage and motivate others to "be like them". EIs always have teachers who are building bridges with students, promoting a good atmosphere in classrooms and who are generally very well accepted by students, let others learn from them. Additionally to the best practices, it is important as well to give the opportunity to the "not that good" ones to talk about and deal with their issues and problems without any judgement or negative consequences.
Set new behaviours and evolve culture by using new technologies to change the organisational culture	In order to strengthen and increase the influence of good practices of the champions, new technologies need to be available and used by everyone so the "word can be spread" faster, reaching everyone. New technologies include new applications as well as new ways of using them, and they will be accepted and used by everyone only if they are well adopted to their purpose. The best content management system with the greatest number of possibilities for storing and sharing digital learning materials will not be accepted if it is complicated to use. Try to avoid contra productive activities such like these ones which are increasing instead of reducing operations time and costs.
Build foundations skills with a digital platform for in-house knowledge exchange and a competence development plan	Building skills that will enable employees to implement digital technologies in their everyday work starts with the determination of the difference or the gap between the ones they already use and the ones they need to acquire. A competence development plan for teachers as well as for the support processes staff, that includes internal and external education programmes, has to be created. A good digital platform can help exchanging the new knowledge and also give everyone access to up to date information about the progress of the transformation.
Align incentives and rewards for those who achieve the goals first	The best have to be rewarded, in order to motivate others to be more agile, adaptive and dedicated to their goal achievement. At this step a customer evaluation could be useful. The leaders should give the students the opportunity to validate their teachers work, and reward the best ones.
Measure, monitor and iterate the progress of digital transformation and adapt if needed	Measuring single employees' success can be motivating, but also it can be experienced as control and pressure, if it is not connected to the objective results of the work that is being done. It is important to set a measurement and monitoring system of results that lead to the accomplishment of the defined vision. The results should be measured periodically, and the activities defined in the roadmap of change should be revised if the goals are far from being achieved. It is important when everybody is on the "same ship" that is keeps sailing to the right direction.

Table 2: Framing EI's digital Transformation according to Business model canvas

Element in the business model canvas (Osterwalder et al., 2014, p. 16)	**Implementation in EIs**
Key partners include the network of partners that help getting the job done	Key partners of EIs are all decision making and funding institutions that in some way influence the work of EIs (like the ministry on the national level, the founders on regional or local level), suppliers (of student meals, teaching materials and books, assets – also IT assets, and services like transportation of students, accommodation etc.) and other support organisations. Key partners should share a common idea of EI's digital transformation.
Key activities that are the most important and have to be done good	The most important activities in EIs are teaching, learning for better teaching and evolving in performing educational processes for new trends and challenges that the digital era brings. It is important to encourage and motivate employees in order to become best, or near to the best.
Key resources that are required for performing the key activities	The most important resources of EIs are teachers and other support processes staff with their capabilities and competencies to perform the new activities and work on the development of their skills in order to implement the change. Skills that are missing have to be built through a detailed competence development plan for teachers as well as for the support processes staff that is created and performed. Other, physical assets (appropriate classrooms, networks, computers etc.) and intangible assets (suitable organizational culture and corresponding fundamental values) are required as well in order to ensure that the new knowledge and skills can be used for the transformation.
Value propositions include products and services that create value for the customer	The services which are provided by EIs are important to the society because knowledge and skills help students to cope with the new challenges that are coming with new trends. The change is affecting EI employees as well, by making their own work easier, faster, more efficient and effective. These values are of strategic importance to every EI and they have to be assessed. Every teacher, accountant, sociologist, psychologist or other expert in the EI needs to know what is supposed to be done and what is the new or better value for the students that comes with the transformation.
Customer relationships defines the type of relationship as well as the acquisition process of new customers	Relationships that are established between the EIs and their primary customers, the students, as well as their parents, are assuming trust in the fact that they are doing their jobs in the best possible way. In the primary education there is no acquisition process, but in the secondary education the digital maturity of the school, its efforts in applying new technologies and new ways of teaching can bring a competitive advantage for the potential students and build a stronger relationship to the current ones.
Channels include the way of communicating the value propositions to the customers	The value proposition should be communicated as soon as possible, through every available channel – meaning at school and student board meetings, in classes and other formal events as well as through informal happenings. Organizing external workshops, webinars, idea dissemination events and conferences is a good way to present the communicating the value propositions to the customers from other segments.
Customer segments that the organization is aiming to reach with the new product or service	EIs primary customer segment are students. Other customer segments are related to the wider society participants like future employers, higher education institutions, and other stakeholders.
Cost structure including all costs that can occur in the transformation process	In EIs there are several cost categories: related to the building of digital infrastructure, the education of teachers and administrative stuff to use new technologies and for digital and online teaching materials. Besides these direct costs indirect costs are also relevant to overall transformation costs (like costs of promoting and communicating the change, infrastructure maintenance costs and other).
Revenue streams that are supposed to get the new value financed	Public EIs are mainly nationally financed from the state budget, but there are possibilities to enrich the funds for digital transformation by public or private funds, donations and sponsorships. It is possible to get additional funds from institutions on the local, regional and national level, as well as to apply to international funds if possible.

Remaining question for further research and analysis is how to ensure sustainability of digital transformation of EIs in terms of ensuring funding and maintenance regarding costs that will occur throughout the transformation process. These issues need special attention and efforts to address them.

The European Commission funds a variety of activities on research and innovation for digital learning under several programmes, including Horizon 2020, previously funded through areas of Technologies for learning and skills (2016), Technologies for better human learning and teaching (2015) and Advanced digital gaming/gamification technologies (2014), (Research and Innovation for ICT in education).

5 Discussion

Previous tables 1 and 2 show that concepts of digital transformation that are well known in various business industries can be mapped and applied to an educational environment. Operatively, educational institutions can replicate implementation efforts and guidelines from the business/ industry sector for supporting processes that are performed in a similar way in educational institutions. This research was conducted with the goal of developing possible scenarios for digital transformation of educational institutions within the project "e-Schools: Establishing a System for Developing Digitally Mature Schools" (e-Schools, 2018) for the purpose of planning how to continue the digital transformation of educational institutions in Croatia.

By framing EI's digital transformation by two selected digital transformation frameworks we gained a good perspective of how to approach the transformation process itself. Both frameworks show the linkage of strategy and operational processes as well as the importance of resource capability assessment in form of employee skills and motivation to carry out the digital transformation.

The digital transformation compass recognizes the significance of organizational factors like leading and cultural influencers. Business model canvas framework emphasizes the customer centricity dimension as a leading determinant of setting goals related to key products and services.

None of two compared frameworks require components of infrastructure or technology related to digital transformation. This also confirms our understanding and interpretation of digital transformation paradigm that it is not all about technology, but it is being inspired by business needs firstly and then supported by appropriate technologies secondly.

6 Conclusion

Digital transformation as a paradigm of changing the way organizations run their business is often misunderstood and mistakenly considered as only another informatization initiative or a buzzword. Our short insights about publications related to DT methodologies indicate increasing interest of researchers for the topic and its subject area analysis shows that DT is more than just another technology initiative. Digital society as context in which educational institutions are providing public services is challenging, so EIs need to change and adapt themselves.

In order to support EIs in this transformation process, we investigated methodological frameworks and took two well- known frameworks in order to frame the digital transformation process of EIs within them. This has led us to confirming our presumption that the leading issue in digital transformation are not digital technologies, but challenges of digital societies that urge organizations to set their goals with the emphasis on business needs and customer expectations which then can be supported by new technologies.

References

Begičević Ređep, N. Balaban, I., Žugec, B. Klačmer Čalopa, M. & Divjak, B. (2017). Framework for Digitally Mature Schools, *Proceedings of the European Distance and E- Learning Network 2017 Annual Conference* (pp. 360-371). Volungeviciene, Airina; Szűcs, András – Jönköping (Eds.), European Distance and E-Learning Network, Sweden.

Business Model Canvas. Retrieved on April 28, 2018 from: strategyzer.com/canvas/business-model-canvas

Digital Transformation: History, Present, and Future Trends. Retrieved April 28, 2018. from: auriga.com/blog/digital-transformation-history-presentand-future-trends/

Digitrans method framework, Retrieved August 22, 2018. from www.interreg-danube.eu/approved-projects/digitrans/section/digitrans-method-framework

e-Schools, Retrieved April 28, 2018. from: www.e-skole.hr/en/e-schools/project-description/

Osterwalder, A., Pigneur, Y., Bernarda, G. & Smith, A. (2014). *Value proposition design*, Hoboken, New Yersey: John Wiley &Sons

Research and Innovation for ICT in education. Retrieved on August 21, 2018 from: ec.europa.eu/digital-single-market/en/ research-and-innovation-ict-education

Schallmo, D., Williams, C.A. & Boardman, L. (2017). Digital transformation of business models — best practice, enablers, and roadmap. *International Journal of Innovation Management* 21(8), pp. 17. doi: 10.1142/S136391961740014X

Tomičić Furjan, M. & Kuhar, P. (2018) Comparative analysis of digital transformation frameworks. In prof. dr. sc. Vladimir Rosić (ur.). *Radovi 7. međunarodne znanstveno-stručne konferencije: Liderstvo kroz partnerstvo.* Visoka poslovna škola PAR, Rijeka.

Westerman, G., Bonnet, D. & McAfee, A. (2014). *Leading Digital – turning technology into business transformation*. USA: Harvard business review press.

World economic forum - Realizing Human Potential in the Fourth Industrial Revolution, An Agenda for Leaders to Shape the Future of Education, Gender and Work, Retrieved August 22, 2018. from: www3.weforum.org/docs/WEF_EGW_ Whitepaper.pdf.

CRM Solutions and Effectiveness of Sales Processes in Export Organizations

Bojan Krajnc
LOPIS
Thalerjeva ul. 10, 2212 Šentilj v Slov.goricah
bojan.krajnc@lopis.si

Polona Tominc
University of Maribor
Faculty of Economics and Business
Razlagova ul. 14, 2000 Maribor
polona.tominc@um.si

Ruben Picek
University of Zagreb
Faculty of Organization and Informatics
Pavlinska 2, 42000 Varaždin
ruben.picek@foi.hr

Simona Sternad Zabukovšek
University of Maribor
Faculty of Economics and Business
Razlagova ul. 14, 2000 Maribor
simona.sternad@um.si

Abstract. *The main purpose of this research study was to analyse the impact of CRM solutions on the efficiency of sales processes in the export organisations. The occasional sample of export organisations was obtained using the sampling frame of the SloExport database. Despite surveyed companies' expressed opinion, that CRM supported sales processes are the most important for implementation of key criteria of efficiency, the results revealed that the use of CRM solutions in organisations is rather limited: in most cases companies have not supported all sales process phases, and they are often also using only selected functionalities of CRM solutions.*

Keywords. CRM, sales process, sales automation, key performance indicators, export organisations

1 Introduction

In the last years, there were intensive discussions about Customer Relation Management (CRM) solutions in the academic and business world (Wahlberg et al., 2009; Krajnc, 2016; Wan & Xie, 2018). Modern trends in business are deregulation, globalisation, convergence, mobility and fast growth of internet of things (IOT) which have changed the way companies do business; it forced them to find new business models, products and sales orientations (Wahlberg et al., 2009; Krajnc et al., 2018). On the other side, companies benefit by gaining more attention and focus on the customers, improving customers' relationships, satisfaction, loyalty, thus improving profitability and successfulness of business (Krajnc, 2016).

Strategic CRM is becoming one of the most critical and strategic activities for companies regardless of the size, industry and the mode of operation. In the highly complex and heterogeneous environment of the 21st century, companies have realised that creating and maintaining a quality long-term relationship with the customers is more important than just focusing on gaining new customers (Agapitou et al., 2017).

In these specific circumstances, CRM solutions have a more important impact on business, as compared to other business applications (Gartner, 2018). Companies need quality information support and better control over the sales processes with which they can increase the success rate of closed sales opportunities (Gartner, 2017). Collecting, storing and using information about existing and potential customers have never been as efficient as in the time of IT. It is clear that the work of salespersons is changing (Yerpude & Kumar Singhal, 2018). They have to appreciate database as a precious tool which shortens the time for searching for the potential customers and increases the time available for gaining the customers (Krajnc et al., 2018). The nature of strategic sales approach demands fusion of business and selling processes and total support for customers. This demands that companies' processes are supported with an IS which assures clarity, traceability and fast response. Automation of sales processes and integration of sales methodology and processes in CRM is becoming an imperative without which the realisation of strategic sales approach cannot be assured (Krajnc et al., 2018). That is why this research is focused on CRM solutions as information tools which enable salespersons to perform more tasks at a higher quality in a very short time and so help to achieve better efficiency of sales processes.

The paper aims to research the impact of CRM solutions on the efficiency of sales processes. Companies included into the present research were among companies mainly involved in international business in Slovenia: exporters, which achieved more than 50 % of their income by export (1.365 companies) and all international companies operating in Slovenia (634 companies). The sample includes 101 companies.

2 CRM Solutions

CRM is a synonym that is used for practices, strategies and technologies that organizations use to manage and analyze customer interactions and data through customer lifecycle (Brassington & Pettit, 2000; Ahn et al., 2003; Wahlberg et al. 2009; Krajnc, 2016; Starzyczna et al., 2017). It can be defined as an integrated approach to managing customer relationships through a combination of three important components: people, processes and technology (Agapitou et al., 2017). Wahlberg et al. (2009) pointed out that CRM researches were focused on the integration of business processes in an organisation, a matter of customer-focused business strategy and a matter of customer knowledge management. Tightly connected to this last perspective is CRM as a technology-enabled customer information management activity (Wahlberg et al., 2009). CRM solutions combine customer data and documents into a single CRM database, where an employee can more easily access and manage it (Krajnc et al., 2018). It includes four areas (Wahlberg et al., 2009; Krajnc, 2016; Starzyczna et al., 2017; Yerpude & Kumar Singhal, 2018):

- Strategic CRM, where the main focus is on the customer and to implement this focus as a CRM strategy, including emphasising the systematic analysis and use of customer information as a platform for marketing and management.
- Analytical CRM includes the customer data analysis, evaluation and prediction of future customer behaviour and identification of key customer groups and their preferences. The dashboards are drawn to enable the facilitation of maximum customer value.
- Operational CRM includes business processes that support front office activities, including sales, service and support. The focus is on automating the customer-facing processes through three main components: sales force automation (SFA), marketing automation and service automation. It generates enormous data, which together with the market intelligence data coming through different channels and are gathered into a single database. From there it is received to derive meaningful business analytics in analytical CRM.
- Collaborative CRM enables various departments (sales, technical support, and marketing) to share any information they collect from interactions with customers via different communication channels such as websites, emails, phone calls, live chat, marketing materials and the social media.

Today's customer is more educated, informed and highly demanding due to the knowledge and alternatives available in the market (Yerpude & Kumar Singhal, 2018). Therefore, an effective CRM solution is about acquiring, analysing and sharing knowledge about the customers and also sharing information with the customers for the quick and timely service to the customer. CRM solution provides an integrated view of customer interactions starting with software applications that capture these interactions (operational CRM) and with the effective analyses of the data (analytical CRM) to reveal the hidden and important information required for improving the relationship of companies with the customers (Davood Karimzadgan et al., 2013). The main reasons for the use of the CRM system are the effort to reduce costs, streamline processes, increase the number of customers and maintaining active relationships with customers, increase repeated purchases and increase turnover, increasing customer loyalty, and expanding customers' lifetime value (Brassington & Pettit, 2000; Ahn et al., 2003; Starzyczna, Pellešova & Stoklasa, 2017) by applying coordinated and customized marketing, sales and services concepts with modern information and communication technology (Zaby & Wilde, 2018). Krajnc (2016) pointed out that there are two main reasons for the increased interest in the field of CRM solutions: the development of information technologies (internet of things - IoT, big data, social media, mobile devices, artificial intelligence – AI, etc.) and the increased importance of orientation to the customers as the basic business philosophy.

After implementing CRM solutions, the improvement of its effectiveness is the constant endeavour of organisations. CRM solutions held out a lot of promise in the mid-1990s, but a considerable number of failures caused concern about its usefulness (Kaushik & Kundan, 2009). An important factor for this could be that the management in the organisation still often treats the technological aspects as the most important and not enough attention is focused into processes and employees (Vella & Caruana, 2012). Research results show, that the technology is important, but is not the only and sufficient factor for the success of the implementation of the CRM system (Mary Lou et al., 2005; Krajnc, 2016). The majority of problems in the implementation of CRM are not of the technological nature, but organisational and include organisational changes and disorders, different views regarding customers' data and changes in business processes (Wahlberg et al. 2009; Finnegan & Currie, 2010; Zaby & Wilde, 2018). That is why the comprehensive research of company's orientation towards organisational factors that have the impact on the effective use of CRM solutions is needed. Researchers are analysing the effectiveness of the implementation of CRM solutions (McGill & Bax,

2007) from the different viewpoints. Researchers often base their work on TAM (Technology acceptance model) and extend it with factors associated with the level of individuals at the different working positions (Avlonitis et al., 2005), as well as at the level of the whole organisation. Some are also oriented towards research of factors influencing the usefulness of IT solutions and technologies for clients on the market (Giovanis et al., 2012). For the organisations to achieve positive results with CRM solutions, the employees have to understand the critical success factors (CSF) of CRM solution use. The researchers emphasise that organisational factors (management support, training and motivating the employees, together with the organisational structure/processes) play a key role in the implementation of CRM solutions in different business areas (Padilla-Melendez & Garrido-Moreno, 2014; Riyad & Hatem, 2014).

Organizations are constantly dealing with the question of how to obtain the highest quality information about their customers, which would help them make key business decisions as well as establish long-term and profitable relationships with customers, and thus at the same time indirectly boost the chances of business success (Gneiser, 2010; Agapitou et al., 2017). In 2013 CRM solutions were the first most commonly used management tool worldwide (Rigby & Bilodeau, 2013). Even more, at the end of 2017, worldwide CRM solutions revenue overtook other systems, making CRM the largest of all software markets, according to Gartner (2018). Worldwide CRM software revenue amounted to $39.5 billion in 2017. In 2018, CRM software revenue will continue to take the lead in all software markets and be the fastest growing software market with a growth rate of 16 % (Gartner, 2018). The most popular CRM software worldwide are (Capterra, 2018): Salesforce, Zoho, Odoo, Microsoft Dynamics, OroCRM, Hubspot, SAP, Sugar CRM, Sage etc. Each of them has specific features. However, Gartner (2018) pointed out that no company will achieve its sales, marketing, customer service, field service and digital commerce business objectives with just one CRM vendor, but they will implement multiple technologies from multiple vendors through CRM marketplaces (such as Salesforce's AppExchange, SAP's Hybris Extend Marketplace, Microsoft AppSource, Oracle Cloud Marketplace etc.).

With the development of technology new forms of CRM solution (e.g. electronic CRM – eCRM, mobile CRM – mCRM, the social CRM - sCRM) were developed that facilitate the collection of customer data with the widest possible perspective, the creation of large-scale data warehousing and data mining as an important infrastructure of analytical CRM solution (Wan & Xie, 2018). The advanced analytical functions become more important in the actual context of the economy (Gartner, 2017; Zaby & Wilde, 2018). The database with customers and the analytical functions associated with the database could make the difference between a winner and a loser in the game of the economy (Furtuna & Barbulescu, 2012). Analytical CRM solutions, which can discover information and enhance the knowledge hidden in the huge amount of data, play a crucial role in decision support (Xie et al., 2008).

CRM solution has, therefore, become a key enabler to personalise customers' experience and improve customer satisfaction and customer retention. However, it has been found that some organisations do not exploit the full potential of CRM solution and the majority of organisations are trying the enhance its' successful implementation (Krajnc, 2016). Research findings show that investment in technology is a necessary but not a sufficient condition for achieving positive results with CRM solutions (Padilla-Melendez & Garrido-Moreno, 2014; Agapitou et al., 2017). Wahlberg et al. (2009) exposed on their extended literature review that under researched areas in the body of CRM research are research within the area of operational CRM and collaborative CRM, where important part is also sales area with sales processes. CRM solutions are the heart of the modern sales processes that depends on capturing, analysing and using data to make decisions (Gartner, 2018). CRM vendors add functionalities all the time. Fig. 1 shows an expectation of use sales functionalities in companies over time (Gartner, 2017).

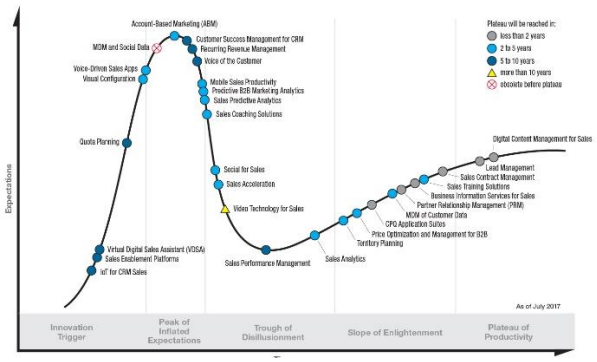


Figure 1. Hype Cycle for CRM Sales, 2017 (Gartner, 2017)

Fig. 1 shows that plateau of productivity (last column) will have reached functionalities digital content management for sales and lead management in less than two years. Slope of enlightenment (fourth column) will reached functionalities such as sales contract management, business information services for sales, partner relationship management (PRM) and CPQ application suites in two years; and functionalities such as territory planning, price optimization and management for B2B, MDM of customer data and sales training solutions in two to five years. New functionalities which is in the phase of innovation (first column) are IoT for CRM sales, sales enablement platforms, virtual digital sales assistant (VDSA), quota planning, visual configuration, voice-driven sales apps etc. (Gartner, 2017). Future trends of CRM will be in three areas (Keenan, 2017; Nicuta et al., 2018): (1)

personalization (organizations, departments, team and client ability to define personal standards, personal processes, specific data and reports and personal CRM experience), simplification (determinates extended adoption and lower prices of ownership) and alternative user experiences (voice-based interactions using CRM systems and artificial conversational entities, i.e. chatbots).

3 Research Study

The main purpose of the research was to analyse the impacts of CRM solutions on the efficiency of sales processes in companies. The key criteria, according to which we have measured the efficiency of sales processes, were:

- time duration of the sales cycle,
- the number of closed sales opportunities,
- the number of gained sales opportunities,
- the value of sales income.

Based on the literature review (see Krajnc, 2016), we hypothesised, that the efficiency of sales processes depends on the extent to which CRM solutions are integrated into the business processes within the sales function; the indicator of this characteristic in this empirical research was:

- the number of phases in sales process supported by the CRM information solution.

Hypothesis H1 was tested:

H1: The efficiency of the sales process depends on the extent to which CRM solutions are integrated into the business processes of the sales function.

In a research study, we have used database SloExport. The database of Slovene exports is a project of Chamber of Commerce and Industry of Slovenia. Database SloExport contains data of more than 4.500 Slovene exports companies (SloExport, 2014). The occasional sample of export organisations was obtained using the sampling frame of the SloExport database; the web-based survey was conducted. Questionnaires were addressed to the management teams, directors and managers of companies. Companies included into the present research were among companies mainly involved in international business in Slovenia: exporters, which achieved more than 50 % of their income by export (1.365 companies) and all international companies operating in Slovenia (634 companies). We received 101 completed questionnaires. Most companies (63.4 %) reported they use a CRM solution. Fig. 2 shows that the most companies (34.9 %) use Microsoft Dynamics CRM solution for supporting the sales activity efficiency in sales process. In the second place, there is SAP CRM solution, being used in 17.5 % companies. Intrix CRM (local solution) is used in 15.9 % of companies. Less frequently used are: 11.1 % of companies use CRM modules of existing business information system (ERP), 4.8 % use SalesForce CRM, 3.2 % use Sugar CRM. Other solutions are used by 12.7 % of companies. Where 98 % companies used them in the sales area, 57,8 % in the marketing area, 45,3 % in the service area, 40,6 % in the finance area and less than 20 % in the purchase, accounting and HR areas.

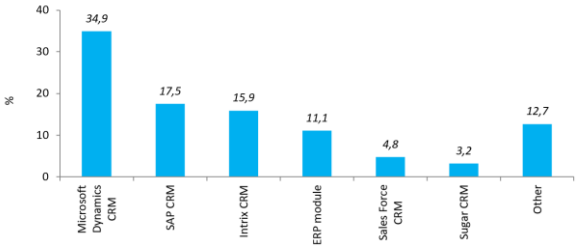


Figure 2. Usage of CRM solutions by companies

Almost all companies which use CRM solution, mainly use it for managing costumers' contacts and other customers' data (93.7 %). More than four-fifths (81.0 %) of companies use it for detecting leads (potential sale opportunities), 77.8 % for analytics or reports and 68, 3 % for managing workflows of the sales process. Other less frequently used areas of CRM solutions are: customers' orders (57.1 %), financial information about customers (55.6 %), dashboards (55.5 %), qualification of leads in sales opportunities (54 %), planning and implementation of marketing campaign (49.2 %), "360 degrees" view on the customer (47.6 %). Less than 40 % of companies use CRM solution for customer segmentation (38.1 %), for SFA (38.1 %), for management of procedure of complaints (34.9 %), for measuring efficiency of sales employees (31.7 %), for customer profitability analysis (28.6 %) and as knowledge base (27 %). Only 15.9 % of companies use CRM solution for measuring customer satisfaction. The research shows that not all functionalities of CRM solutions are implemented and used as it might be expected.

CRM solutions are workflow-based solutions and could provide support for sales process workflows of different complexity (Krajnc, 2016; Krajnc et al., 2018; Zaby & Wilde, 2018). We have analysed the characteristics and complexity of defined workflows within CRM solutions used by the companies. In almost all companies (93.0 %) the step named "offer or cooperation offer" is well defined and implemented. On the second place is the step named "closing an opportunity", which is defined and implemented in 80 % of companies, on the third place is the step named "the first contact with the client" which is implemented in 78 % of companies. Only three other steps are implemented in more than half of the companies: "current analysis of customers" in 62 % of companies, "lead or potential sales opportunity" in 60 % of companies and "negotiation" in 55% of companies. Workflow steps which are at least implemented are: "qualification of customers based on defined measures" in 45 % of companies, "problem solving" in 43 % of companies and the step named "additional -

cross-selling" only in 35 % of companies. There are many even less presented steps of sales process workflow – only 11 % of companies have implemented the step "solving complaints". The reason could be that CRM processes (also sales processes and workflows) are not optimal (Zaby & Wilde, 2018). They added that they should be optimised continually using the latest knowledge generated from operational data.

Nearly all companies (92.7 %) replied that CRM solutions are an important source for ensuring and increasing the competitiveness and they contribute to faster and better access to internal data and better control, coordination and transparency of the sales activities in the company.

A lot of companies also replied (88.5%) that using CRM solution lead to better sales effectiveness and better resource controlling in sales. 82.3 % of companies answered they have a better response time. Furthermore, 77.1 % of companies quoted better sales satisfaction of clients, more efficient segmentation of customers and creating a better qualitative long-term relationship with customers. Only a few companies quoted shorter sale cycles, increasing sales incomes and the volume of sales opportunities (76.0 %), greater flexibility of buying process (75 %) and lower processing customer costs (70.8 %) Other areas of CRM solutions usage identified by companies are: differentiation from other competitors (65.6 %), ensuring greater and more unique added value for costumer (62.5 %), increased reputation for the company (54.2 %), the improvement of strategic position on the market (51.0 %), better profitability (50.0 %) and bigger market share (50.0 %).

With the purpose to test H1, we first employ the factor analysis (Tabachnick & Fidell, 2013), to form the multidimensional indicator(s) of sales process efficiency.

Results of factor analysis revealed that the use of factor analysis is justified (KMO = 0.506; Bartlett's test of sphericity is significant; $p<0.05$) and that the two-factors explain 63.93 % of the variance for four measured variables. Commonalities, explaining the proportion of variance explained for each of four measured variables, as well as the factor loadings are presented in Table 1.

Table 1. Factor loadings – sales process efficiency

Measured variable	Factor 1	Factor 2	Commu-nalities
Time duration of sales cycle	0.636	0.246	0.464
The number of closed sales opportunities	0.653	-0.240	0.484
The number of gained sales opportunities	0.846	-0.038	0.716
The value of sales income	-0.021	0.944	0.892

Based on the factor loadings, factor 1 is named "Organisational efficiency" and factor 2 "Value efficiency".

In the next step the simple regression was used; due to the two-factor results for the dependent variable, two sub-hypotheses were tested:

H1a: The "Organisational efficiency" of the sales process depends on the extent to which CRM solutions are integrated into the business processes of the sales function.

H1b: The "Value efficiency" of the sales process depends on the extent to which CRM solutions are integrated into the business processes of the sales function.

Results for the two simple regression models are presented in Tables 2 and 3.

Table 2. Regression analysis results – testing H1a

	Regression coeff.	t	Sig
Constant	-0.257	-1.673	0.099
The number of phases supported by the CRM	0.066	2.546	0.013
	Rsquare = 0.088 ($p<0.05$)		

Table 3. Regression analysis results – testing H1b

	Regression coeff.	t	Sig
Constant	0.059	0.367	0.715
The number of phases supported by the CRM	-0.015	-0.0558	0.579
	Rsquare =0.005 ($p>0.05$)		

Regression results revealed that the number of phases supported by the CRM solutions has a significant positive impact on the Organisational efficiency of the sales function. Therefore hypothesis H1a is confirmed. The low coefficient of determination is expected and suggests that there is also another factor, which influences the organisational efficiency of sales functions, but the purpose of the regression analysis was not the identification of all possible factors affecting the efficient organisational aspects of the sales function.

Hypothesis H1b is not accepted since the regression coefficient is not significant – we cannot confirm that the impact of the number of phases supported by the CRM solutions, has a significant positive impact on the Value efficiency of the sales function.

The main result of this part of the research is that the extent to which CRM solutions are integrated into the business process within the sales function, has the important impact on the organisational aspect of efficiency in sales function, but that there are other factors, that influence the efficiency measured in terms of value of sales income.

4 Discussion

Based on conducted survey research study and based on interviews with selected companies which we studied as cases (Krajnc, 2016), it may be concluded, that in most companies a high share of CRM implementation initiatives starts without defining a business case or blueprint which will define what the company wants to accomplish. Similarly, potential barriers which the company needs to overcome are not defined properly. The sales (CRM solution use) vision too often relies on the short-term results, despite it should be developed from the long-term vision of creating the value for the customers and other stakeholders of the company (Krajnc et al., 2018). Our research showed that 74.3 % of companies measure efficiency in the sales process, while 22.8 % of companies do not. The first part in the process of CRM development strategy is to revise the business strategy of the company. The detailed understanding of business strategy is crucial for the appropriate implementation of customer relationship strategy. CRM solution implementation activities have to be aligned with the business strategy of the company.

Many research surveys in the past years show that the consistent sales success is closely connected with CRM solutions used in the company and its alignment with the sales methodology (Krajnc, 2016). The implementation of CRM solution in the company is not a guarantee for the complete exploitation of CRM concept and expected benefits necessary for successful business in a long term. Our research shows that only a quarter of companies share the opinion (25.7 %) that sales process is entirely unsupported with CRM solutions, while 8.9 % of companies share the opinion of being poorly supported. Only 27.7 % of companies think that they are partially supported, and it is surprising that the same percentage of companies thinks that they are well supported. Very surprising is that only one-tenth of them (9.9 %) think that sales processes are entirely supported with CRM solution. It should be pointed out that because of new technologies some processes can become standardised and consolidated in time, wherein other areas of new functionalities are appearing (Gartner, 2017). Because of that, organizations need to manage a portfolio of different CRM applications with a different level of maturity (see Fig. 1).

The key is the orientation to costumers in all areas of company operations, and full awareness of these processes, before the company even starts with CRM solution implementation. We recommend to organizations that they conduct the preliminary analysis of CRM solutions acceptability by employees. At that point, there is always good to refer to the examples of good practice where the increase of the sales efficiency was achieved. SFA systems as part of CRM were developed to help companies in the areas of leading and tracking of sales activities and also to assure the sales forecasting and support. In practice, a great number of CRM projects fail because they are not efficient and do not have enough tools and resources to support sales employees. Salespersons in most cases do not get the promised value from the CRM solutions. SFA systems are mostly used for routine data entry. Collected data is usually used by management, but directly the sales do not benefit from it.

On the other hand, it often happens that the employees are not at all or are only poorly qualified for costumer conduction, contacts improvement and long-term relationship building. If salespersons become more experienced users of CRM solutions and have appropriate CRM oriented skills, their customer service is better (Agapitou et al., 2017). But management expects that the implementation of CRM solution in the company will solve all problems. Although the companies think they are customer oriented, what the company is willing to show in public, we often detect, that this is not always true. Most companies usually still possess the predomination of short-term vision of making business or gaining benefits instead of the long-term strategic vision (Krajnc, 2016; Krajnc et al., 2018). Companies which succeed in CRM alignment to the corporate strategy and culture, claim higher satisfaction from CRM implementation (Agapitou et al., 2017).

Companies need more proper information support and precise control over sales process which will allow them to increase the success of closing sales opportunities. Collecting, saving and using information about the existing and potential customers could never be as efficient as now, in the era of IT (Gartner, 2017). The salespersons have to consider the use of the database as a crucial tool which shortens the time for the customer seeking and enlarges the time available for customer contact.

The nature of strategic sales approach demands fusion of business and sales processes, and complete customer support. This point of view demands that the company develops information supported sales processes which ensure visibility, traceability and fast response. SFA and integration of sales methodologies and processes within CRM solutions is becoming an imperative without which it is practically impossible to ensure the realization of strategic sales approach that would consequently enable company to adapt to the specific customer needs. Gartner (2018) research exposed that sales teams should concentrate on four major topics: sales enablement platforms, mobile sales productivity, sales predictive analytics and sales performance management.

Sales process provides information about the daily activities of salespersons when they are gaining new customers, qualifying sales opportunities, preparing offers, dealing or closing the open opportunities. It is a sequence of steps which salespersons are systematically carrying out with the one goal - successfully close the sales. All these phases are supported if they are included in CRM solution workflow, and that support provided by CRM solution

is regularly used. More than 30.6 % of researched companies are not practising this. Only 21.4 % of companies have defined basic three steps in their CRM workflow, 18.4 % of companies have defined four or five steps in CRM workflow, 16.3 % have defined six or seven steps in their CRM workflow, and only 13.3 % are using CRM with the very elaborated workflow of more than seven steps. Our research also shows that a number of phases supported by CRM solutions has a significant positive impact on time of duration of the sales cycle, the number of closed sales opportunities and the number of gained sales opportunities. Therefore, good leaded sales process helps the company to improve the visibility, predictability, forecasting, consistency, scalability and sustainability; this is usually reflected in the stronger company's competitive advantage on the market.

5 Conclusion

Sales methodology is a concretisation of sales process steps since it precisely illustrates the concrete action of salespersons' activity in the sales processes. Both, sales processes and sales methodology are extremely powerful and efficient in the company when they are supported or integrated with adequate CRM solution. CRM solution could be a beneficial system for management as well as for the salespersons, while it focuses on the automation of sales efficiency and therefore influences sales processes as well as sales methodology supported (Krajnc et al., 2018). In that case, salespersons data capturing is used for sales management, as well as for sales process itself. Salespersons will rather use the tool or application which will help them to sell more and enables them to be more successful.

Our research results revealed that the companies on average share the opinion, that CRM solution supported sales processes are the most important for implementation of key criteria of efficiency. It is interesting that 74.3 % of companies' measure their efficiency in sales processes and, for that, they use key performance indicators which are bound to accomplished activities or achieved results. Furthermore, we have found that the number of information supported phases as well as the number of information supported areas of sales processes partially affect the efficiency of sales processes, namely the organizational efficiency. The same relates to the proportion of users of CRM solution supported sales processes which do not have any influence. On the other side, we have established that companies' that are users of CRM solutions give more importance to CRM information supported sales processes for ensuring or increasing the competitive advantages on the market in international business, as compared with companies not using CRM solutions (Krajnc, 2016).

The most important for companies are: faster and better access to information, better control, coordination and visibility over the sales activities, bigger operative efficiency in sales and better resource controlling in sales, faster time response and better customer satisfaction with sales service, more efficient segmentation of customers and creating qualitative long-term relationships at higher level. CRM solutions, which are intended for sales support and are supporting different business areas in the company, can be treated as investment in business process, which brings several benefits and extra value in different phases of the sales process. Because customer behaviour and competitive environments are changing rapidly today, the CRM processes have to be adapted frequently to stay with the goals (Zaby & Wilde, 2018).

Our empirical research is limited to Slovenian export organizations, so the extension of the research to all organizations in Slovenia and broader, offer the opportunity for further research.

References

Agapitou, C., Bersimis, S. & Georgakellos, D. (2017). Appraisal of CRM implementation as business strategy option in times of recession: The role of perceived value and benefits. *Int. Journal of Business Science and Applied Management*, 12(2), 18-31.

Ahn, Y.J., Kim, K. S., & Han S. K. (2003). On the design concepts for CRM systems. *Industrial Management and Data Systems,* 103 (5), 324-331.

Avlonitis, G., Nikolaos, J., & Panagopoulos, G. (2005). Antecedents and consequences of CRM technology acceptance in the sales force. *Industrial Marketing Management,* 34 (4), 355-368.

Brassington, F. & Pettit, S. (2000). *Principles of Marketing*. London: Prentice Hall.

Capterra. (2018). The top 20 most popular CRM Software. Retrieved from https://www.capterra.com/customer-relationship-management-software/#infographic

Davood Karimzadgan, M., Davood, V. & Rahebeh, A. (2013). Investigating Factors that Affect CRM Success with Using Structural Equation Modeling. *International Journal of Academic Research in Economics and Management Sciences,* 2(2), 160-168.

Finnegan, D. J. & Currie, W. L. A. (2010). A multi-layered approach to CRM implementation: An integration perspective. *European Management Journal*, 28, 153-167.

Furtuna, T. F. & Barbulescu, A. (2012). The usage of data mining techniques in analytical CRM. *International Conference on Informatics in Economy*, 10-11, 352-355.

Gartner. (2017). 4 Insights From Gartner Hype Cycle for CRM Sales, 2017. Retrieved from https://www.gartner.com/smarterwithgartner/4-insights-from-gartner-hype-cycle-for-crm-sales-2017/

Gartner. (2018). Gartner Says CRM Became the Largest Software Market in 2017 and Will Be the Fastest Growing Software Market in 2018. Retrieved from https://www.gartner.com/en/newsroom/press-releases/

Giovanis, A., Binioris, S. & Polychronopoulos, G. (2012). An extension of TAM model with IDT and security/privacy risk in the adoption of internet banking services in Greece. *EuroMed Journal of Business*, 7 (1), 24-53.

Gneiser, M.S. (2010). Value-Based CRM. The Interaction of the Triad of Marketing, Financial Management, and IT. *Business & Information Systems Engineering*, 2, 95-103.

Kaushik, M. & Kundan, S. (2009). CRM: A Strategic Approach. *Journal of Management Research*, 8 (2), 65-82.

Keenan, M. (2017). 3 Big CRM Trends to Watch in 2017. Retrieved from https://www.destinationcrm.com/Articles/ReadArticle.aspx?ArticleID=115675

Krajnc, B. (2016). Poslovni učinki uporabe CRM informacijskih rešitev v mednarodnem poslovanju: magistrsko delo. Maribor: Krajnc.

Krajnc, B., Bobek, S., Tominc, P., & Sternad Zabukovšek, S. (2018). Vpliv uporabe CRM informacijskih rešitev na učinkovitost prodajnega procesa. In S. Sternad Zabukovšek et al. (Eds.), Dinamika sprememb informacijske tehnologije in informacijskih rešitev ter spremembe v poslovnih modelih podjetij (pp. 6-43). Maribor: Bio energija.

Mary Lou, R., Liu, R.R. & Hazard, K. (2005). Strategy, technology and organisational alignment: Key components of CRM success. *Database Marketing & Customer Strategy Management*, 12 (4), 315-326.

McGill, T. & Bax, S. (2007). From Beliefs to Success: Utilizing an Expanded TAM to Predict Web Page Development Success. *International Journal of Technology and Human Interaction*, 3 (3), 36-53.

Nicuta, A.-M., Luca, F.-A. & Apetrei, A. (2018). Innovation and trends in CRM. *Network Intelligence Studies*, VI(11), 21-25.

Padilla-Melendez, A. & Garrido-Moreno, A. (2014). Customer relationship management in hotels: examining critical success factors. *Current issues in tourism*, 17(5), 387-396.

Rigby, D. & Bilodeau, B. (2013). Management tools & trends 2013. Retrieved from http://www.bain.com/Images/BAIN_BRIEF_Management_Tools_%26_Trends_2013.pdf

Riyad, E. & Hatem, E.-G. (2014). Testing and Validating Customer Relationship Management Implementation Constructs in Egyptian Tourism Organizations. *Journal of Travel & Tourism Marketing*, 31(3), 344-365.

SloExport. (2014). Podatkovna baza slovenskih izvoznikov. Retrieved from: http://www.sloexport.si

Starzyczna, H., Pellešova, P. & Stoklasa, M. (2017). The Comparison of Customer Relationship Management (CRM) in Czech Small and Medium Enterprises According to Selected Characteristics in the years 2015, 2010 and 2005. *Acta Univ. Agric. Silvic. Mendelianae Brun.*, 65, 1767-1777.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. Pearson: Boston, Columbus etc.

Vella, J. & Caruana, A. (2012). Encouraging CRM systems usage: a study among bank managers. *Management Research Review*, 35(2), 121-133.

Wahlberg, O., Strandberg, C., Sunberg, H. & Sandberg, K. W. (2009). Trends, topics and under-researched areas in CRM research. *International Journal of Public Information Systems*, 3, 191-208.

Wan, J. & Xie, L. A. (2018). Bibliometric Review of Research Trends in Social CRM. Association for Information Systems, 2018. *WHICEB 2018*, 63. Retrieved from http://aisel.aisnet.org

Xie, Y., Xiu, L. & Weiyun, Y. (2008). A Process Driven Architecture of Analytical CRM Systems with Implementation in Bank Industry. I*nternational Colloquium on Computing, Communication, Control, and Management* 3, 57-61.

Yerpude, S. & Kumar Singhal, T. (2018) Internet of Things based Customer Relationship Management – A Research Perspective. *International Journal of Engineering & Technology*, 7(2.7), 444-450.

Zaby, C. & Wilde, K. D. (2018). Intelligent Busines Processes in CRM – Exemplified by Complaint Management. *Business & Information Systems Engineering*, 60(4), 289-304.

Guidelines for GDPR Compliant Consent and Data Management Model in ICT Businesses

Dijana Peras
Faculty of Organization and Informatics
University of Zagreb
Department of Information Systems Development
Pavlinska 2, 42000 Varaždin, Croatia
dijana.peras@foi.hr

Abstract. *The purpose of this paper is to set guidelines for managing consent and personal data in ICT businesses taking into account the provisions of the General Data Protection Regulation (GDPR). The analysis of previous studies on consent management models and GDPR requirements, as well as the comparison of five data management models was made. Based on the analysis, guidelines for the framework of GDPR compliant Consent and Data Management Model in ICT businesses were proposed. The result of the study can help data controllers to improve the integration of consent and data management and to demonstrate compliance with GDPR.*

Keywords. Consent management, data management, GDPR, informed consent, framework of consent

1 Introduction

According to the General Data Protection Regulation (Regulation (EU) 2016/ 679 2016), consent for the processing of personal data for one or more specific purpose is one of six legal basis of lawfulness of processing. It can only be an appropriate lawful basis if user is offered control and a choice with regard to accepting or declining the terms offered or declining them without harm (Article 29 Data Protection Working Party, Guidelines on Consent under Regulation 2016/679 2016). However, in practice consent is the most common basis of data processing (Hunton and Williams 2016). GDPR imposes strict rules on obtaining consent on personal data processing, where data controller must be able to proof the validity of consent. Data controllers are expected to be prepared for demonstrating compliance, especially in case they are dealing with various users and using multiple data sources where there are a lot of complexities involved in building and maintaining a complete model of how personal data is used (Fatema et al. 2017). ICT businesses have different approaches on how to deal with consent, and they usually provide limited solutions. Furthermore, complexity in providing consent and expressing privacy preferences can negatively impact data subject's willingness to disclose personal information (Mont et al. 2009). Furthermore, ICT businesses contain a large number of confidential and sensitive data, which should be protected from malicious activities. Therefore, ICT businesses should cover all the areas that represent a security risk Forming a consent model is an important step in ensuring compliance with the GDPR that will help data controllers to meet the requirements related to the specificity and uniqueness of the data.

This paper will identify important concepts of consent, present guidelines for the framework of managing consent in line with GDPR and propose guidelines towards the successful data management model in line with GDPR. It consists from following sections: an introduction is presented in Section 1. Section 2 describes State of the Art. The research method is presented in Section 3. Sections 4 and 5 are focused on identifying important concepts of consent, describing framework for consent management model and components of GDPR compliant data management model for ICT businesses. Conclusions with Future research are presented in Section 6.

2 State of the Art

The concept of consent originated in the field of medicine. In framework for enforcing consent policies for healthcare systems based on workflows, consent had a central role for assigning permissions to subjects that access patients' medical data (Russello, Dong, and Dulay 2008). In biomedical research (Vayena and Blasimme 2017), the notion of control was divided along three dimensions that debate on data protection: control over data access, control over data uses and control through governance. Other report (Coiera 2003) has outlined several possible models (e-Consent) for determining that patient consent exists prior to

allowing access to health information, which will contain the specific conditions under which the data to which it is attached can be retrieved. Electronic consent (Rowan et al. 2017) was also explored on a Health Social Network to improve the form and accessibility of information presented to users. This study suggested returning control over private health information back to the users, in line with the GDPR. The outline of a new model for informed consent used for personal genome testing, which can meet the norm of providing sufficient information and the norm of providing understandable information, was presented by Bunnik et al. (Bunnik, Janssens, and Schermer 2014). The CMA framework described by Hyysalo et al. provided a simple, general purpose consent management framework and architecture that conforms to the GDPR, with main focus on consent and the use of consent for enabling secure transactions, for authorizing data access to services and for health-related personal data management and processing (Hyysalo et al., 2016).

In order to examine the use of consent in an online environment, where the individuals are able to control the collection, use and dissemination of their personal data, researchers have adopted Faden's and Beauchamp's theory of informed consent (Agrafiotis, 2012). Conceptual model of informed consent based on disclosure, comprehension, voluntariness, competence and agreement was provided (Friedman, Felten, and Millett 2000), which examined how these components play out in a wide range of online interactions. Another model for organizations (Mont et al. 2009), enabling capturing consent, managing and enforcing it, along with revocation, included a set of basic requirements of relevance for organizations: Personal Consent & Revocation Assistant, Data Registry, Consent and Revocation Provisioning, Privacy-aware Policy Enforcement, Disclosure and Notification Manager, Audit and Risk Assurance. Furthermore, authors (Karjoth, Schunter, and Waidner 2003) described the Platform for Enterprise Privacy Practices (E-P3P), which defined technology for privacy-enabled data management and introduced separation of duty between the privacy officer, the security officer and the customers. Architecture for a privacy-enhanced database management system was described and algorithms for privacy constraint processing were discussed (Thuraisingham 2005). Bertino et al. discussed requirements towards the development of privacy-preserving database management systems (Bertino, Byun, and Li 2005), who presented two initial solutions dealing with purpose meta-data and their use in access control. Informed consent online was assessed according to set criteria, and it was examined how cookie technology and Web browser designs have responded to concerns about informed consent (Millett et al., 2001). The rule types, which are the essence of consent model and enable the expression of actions associated with obtaining and revoking consent for the use of personal data, were analyzed (Casassa Mont et al., 2011). Consent management model resulting from the focus group sessions with experts in the field of privacy and consent described three main categories for which it is necessary to require consent (Agrafiotis, 2012): collection of personal data, use of personal data and sharing of personal data. MyData was designed as a framework and model for a user-centric approach for managing and processing personal information in the context of online services (Rissanen, n.d.). A provable expression of consent available for aggregated personal data was offered in a form that allows passing it on, that can be retained, and that remains verifiable by third-parties (Pöhls, 2008). Finally, consent and data management model (Fatema et al., 2017) addressed the lifecycles of consent and data along with the various interactions between their stages and between consent and data lifecycle states due to change of context.

3 Research method

This paper aims to set guidelines for managing consent and personal data in ICT businesses taking into account the provisions of the GDPR. For the purpose of defining framework of GDPR compliant consent management, important concepts of consent were defined. The analysis of previous studies on consent management models and GDPR requirements was made, which resulted with general guidelines for obtaining the consent and processing personal data in line with GDPR.

The comparative analysis of five data management models proposed by Thuraisingham, Mont et al., Rissanen, Hyysalo et al. and Fatema et al. was made. Models were compared based on their elements and components. The presence of the following elements was examined: a) legal basis for personal data processing, b) data controller, c) data subject, d) control through governance, e) data processes, f) privacy levels, g) authorization, h) foundation, i) policy layers, j) policy rules, k) possibility of consent revocation, l) context dependency, and m) validity checking. Elements of Consent and Data Management Models are listed in Table 1. Models were further examined based on described data management model components. Collected type of information was listed for each consent management model, and in case the model did not collect examined information, the n/a mark was assigned. Based on their functions, data management model components were grouped into five units: Contact Interface, Consent Management, Data Management, Origin Management and Context Management. Components of Consent and Data Management Models are listed in Table 2.

Guidelines for the framework of GDPR compliant consent management were then used to describe those data management model units. Furthermore, a scenario of obtaining consent for processing personal data for the marketing purposes in ICT businesses was created.

Table 1. Elements of Consent and Data Management Models

	Thuraisingham Model	**Mont et al. Model**	**Rissanen Model**	**Hyysalo et al. Model**	**Fatema et al. Model**
Legal basis	consent	consent	consent	consent	consent
Data controller	privacy controller	organization	operators	operators	organization
Data subject	users	data subjects	user	account owners	user
Control through governance	governments	regulators	n/a	n/a	n/a
Data processes	collect, process, store, share, delete/modify	collect, process, store, share, delete/modify	collect, move and process	protection, authorization, control	collect, use, store, archive, share, delete
Privacy levels	public, private, or highly private	n/a	n/a	n/a	consent permissions, obligations and validity
Authorization	required	required	required	required	required
Foundation	privacy policy	privacy policy	predefined rules and policies	privacy policy	access control policy, participants policies, GDPR
Policy layers	n/a	legal, business, process, application, information, system, network	n/a	n/a	n/a
Policy rules	n/a	notification, access control, update, protection, obligation	n/a	n/a	n/a
Consent revocation	n/a	yes	yes	yes	yes
Context dependency	yes	yes	yes	yes	yes
Validity checking	no	no	yes	yes	yes

Table 2. Components of Consent and Data Management Models

	Thuraisingham Model	**Mont et al. Model**	**Rissanen Model**	**Hyysalo et al. Model**	**Fatema et al. Model**
Contact Interface	User Interface	Personal Consent & Revocation Assistant	Single Point of Contact	Data Account	User Interaction Handler
Consent Management	Constraint Manager	Privacy-aware Policy Enforcement	Data Sources	Data Source	Consent Manager
Data Management	DBMS and Database Design Tool	Data Registry, Risk Assurance, Disclosure and Notification Mng.	Data Sinks that use personal data	Data Sink	Data Manager
Origin Management	n/a	Audit	n/a	n/a	Provenance Manager
Context Management	Query Processor and Update Processor	Consent and Revocation Provisioning	n/a	n/a	Context Handler

4 Guidelines for the framework of GDPR compliant consent management

This chapter will focus on identifying the important concepts of consent, as well as on requirements for the framework of consent management.

4.1 Concepts of consent

According to the analyzed models (Thuraisingham 2005) (Millett, Friedman, and Felten 2001) (Casassa Mont et al. 2011) (Rissanen n.d.) (Fatema et al. 2017), the structure of the consent can generally be described as follows:

a) consent form, filled by data subject or his representative,
b) context of a consent, which usually contains data about time, location and relevant information communicated between data subjects and data controllers, and which can be modified by data controller, data subject or environment,
c) permissions set by data subject, such as the validity period, allowed party, data format and category, prohibited and permitted actions and their purpose, as well as the conditions for given permissions, if they exist, and obligations that result in certain activity or event.

Furthermore, five conceptual components of consent were detected (Friedman et al., 2000): disclosure, comprehension, voluntariness, competence and agreement. This means that the data subject should know the purpose and benefits of disclosure, as well as potential harms. He should be able to interpret accurately Terms of Use, Privacy Statement and purpose of data processing, as well as the consent form through which the consent is being obtained. However, there is no guarantee that all data subjects will completely understand all aspects of the consent. Voluntariness means that the data subject is not forced or manipulated to give a consent. Data subject has to be mentally and physical competent to give the consent. People who lack those competences (e.g. children under age of 16, mentally ill persons and similar) need to have their representative since they cannot reliably determine the appropriateness of the information they choose to disclose. Agreement means the data subject is given a choice to accept or decline the consent. Furthermore, it means he can choose among different options, and decide to approve or decline them without losing the right to service. He can also choose to withdraw his consent at any time. Stated criteria have to be satisfied in order to obtain the valid consent.

Onwards, the requirements related to consent provided by GDPR (Regulation (EU) 2016/ 679 2016) were collected, and they are as follows:

- consent should be freely given, specific, informed and unambiguous,
- consent should be given by a written or an oral statement,
- consent should cover all processing activities carried out for the same purpose or purposes,
- consent request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided,
- the principles of fair and transparent processing require that the data subject is informed of the existence of the processing operation and its purposes,
- the data subject shall have free choice and be able to refuse or withdraw consent,
- protection of personal data requires setting out of the rights of data subjects and the obligations of those who process personal data,
- modalities should be provided for facilitating the exercise of the data subject's rights, including mechanisms to request access to and rectification or erasure of personal data,
- the controller shall be able to demonstrate that processing is performed in accordance with GDPR.

4.2 Framework of Consent Management

Stated structure of consent, conceptual components of consent and GDPR requirements for obtaining the consent and processing personal data are the foundation of the proposed framework of consent management. Guidelines for the framework of consent management are listed in Table 3.

Table 3. Guidelines for the framework of GDPR compliant consent management

	Guidelines
1	The rights, responsibilities and obligations of data subjects, data controllers and allowed party should be identified.
2	The consent management modalities required for data processing should be defined.
3	Modalities for consent withdrawal should be provided.
4	Capability to adapt to changes of law and privacy policy requirements should be provided.
5	Transparent processing of personal data should be possible.
6	Data erasure should be enabled.
7	Data portability should be possible.
8	Data subject should be able to manage his personal data.
9	Privacy Policy should be understandable.
10	Appropriate technical and organizational measures shall be implemented to demonstrate that processing is performed in accordance with GDPR.

Guidelines for the framework of consent management listed in Table 3. were used to describe the data management model according to GDPR principles for ICT businesses. Thus, in the following chapter, harmonization of data management model according to GDPR principles will be presented.

5 Guidelines for GDPR Compliant Data Management Model

The structure of proposed model is based on the detected components of data management models stated in chapter 3, and amended with guidelines for the framework of GDPR compliant consent management defined in chapter 4. It is structurally most similar to the data management model defined by Fatema et al.

5.1 Components of Data Management Model

To meet the guidelines set out in Table 3, proposed data management model should consist of following five components:

Contact Interface. This component supports interaction between data controllers and data subjects, and serves for obtaining consent and exchange of consent related information. It manages all communication between the data subject and the data controller, and assists data subjects in expressing the consent by providing transparent privacy policy. According to Karjoth et al. (Karjoth, Schunter, and Waidner 2003), privacy policy describes what operations for which purpose by which data user can be performed on each personal data, and consists of three elements: a) header with information about privacy policy, b) declarations of policy and c) authorization rules. Declarations of policy consist of:

- data subjects' personal information,
- information about internal and external data users,
- information about purpose of data processing and subsets of the purpose (if applicable), which can be performed in case the main purpose is approved,
- information about data activities, such as collecting, processing, storing, archiving, transferring etc., and
- information on the essential activities used for defining authorization rules.

Contact Interface also handles all data subject actions through unified interface, and helps them to pursue their rights regarding access to personal data, rectification, erasure ('right to be forgotten'), restriction of processing, data portability, object and automated individual decision-making (Regulation (EU) 2016/ 679 2016). These rights refer equally to all data subjects and their application is independent of individual consent.

Consent Management. Consent permissions need to be obtained prior to collecting and processing the data. Consent Management records privacy instructions, including a description of different contexts under which a new consent is required, or conditions under which the data subject needs to be informed. It also records consent permissions related to termination of data processing, consent validity and communication of personal data breach, but stores only information relevant for current data processing. All other information related to context of the consent, but irrelevant for current data processing, are being stored in Origin Management. Before collecting or processing personal data, it is necessary to obtain user's consent through consent form. In order to ensure fair and transparent processing in respect of the data subject, consent form should contain following information (Regulation (EU) 2016/ 679 2016):

1) the identity and the contact details of the data controller and data protection officer,
2) the purposes and the legal basis for the processing,
3) the categories of personal data concerned,
4) the recipients of the personal data (including recipients in a third country or international organization),
5) the period for which the personal data will be stored,
6) the right to access, modify or erase personal data, restrict processing or object to processing, and right to data portability,
7) the right to withdraw consent at any time and to lodge a complaint with a supervisory authority,
8) the existence of automated decision-making, including profiling.

Consent form should clearly state the terms of consent, and it should contain a series of options for which the data subject can, but does not have to give a consent.

Context Management. It manages context, but also detects changes of context and shares them with Consent Management, Data Management and Origin Management. It collects information on modifications made by data controller (e.g. change of the purpose of the processing), data subject (e.g. data modification, consent withdrawal) or environment (e.g. expiration of consent or data, change of partner). It is important to notice the consent is given in a particular context, for a specific purpose. The purpose can change over time, which means the context of the consent will also change and the consent will no longer be valid. Change of Context can be driven by the following events: data modification, consent withdrawal, consent expiry, data breach, business acquisition, change of data processor etc.

Data Management. It handles data according to the given consent and provides personal data protection control (access control, anonymization, pseudonymization etc.). It performs following actions: collecting, processing, sharing (transferring), storing, deleting, archiving. To be able to manage the consent,

it has to know all the locations where data is stored. It also needs to know at any time where the data is located within the organization, the format in which data is stored and the information to whom the data has been disclosed. This component is critical and has to be secured and protected. The concept of control is observed through three dimensions that are related by a causal connection, which means that one control can affect other controls (Vayena and Blasimme 2017):

a) Control over data access is the basis which sets the conditions of data disclosure. Authorization rule specifies activities that may be executed on personal data by the data controller if the data subject has given him a consent for specific purpose. There are two types of authorization rules: authorization rule with a condition and authorization rule without the condition. If the authorization rule contains a condition, it applies only if the condition is met.
b) Control over data uses determinates who has the right to access personal data and the purposes for which personal data is used, and decides on the relevance of the purpose and its compliance with the interests and expectations of the data subject. Control over data uses is responsible for keeping unauthorized persons away from the personal data. If the person who tries to access personal data is not authorized to execute specific task, he will not be allowed to do it. Furthermore, it can limit the access of applications in case their request for accessing certain personal data is denied.
c) Control through governance is related to management structures, including GDPR, in which data subjects are put in the focus and are given back the control over their personal data

Origin Management. It keeps history of all activities related to consent, which could help identify origin and other relevant information in data activities. It records the location of data and tracks the data that were disclosed to third parties. With the introduction of GDPR, its utility will increase, since it is capable to prove whether the data was used in line with the consent. Furthermore, it can track activities related to change of consent, data breach, data modification etc. Its main purpose is to ensure compliance by providing evidence of taking proper actions at critical moments.

Components of GDPR compliant data management model and interactions between them are presented in Figure 1. The following subchapter will describe scenario of obtaining data subject's consent for the purposes of marketing in ICT businesses.

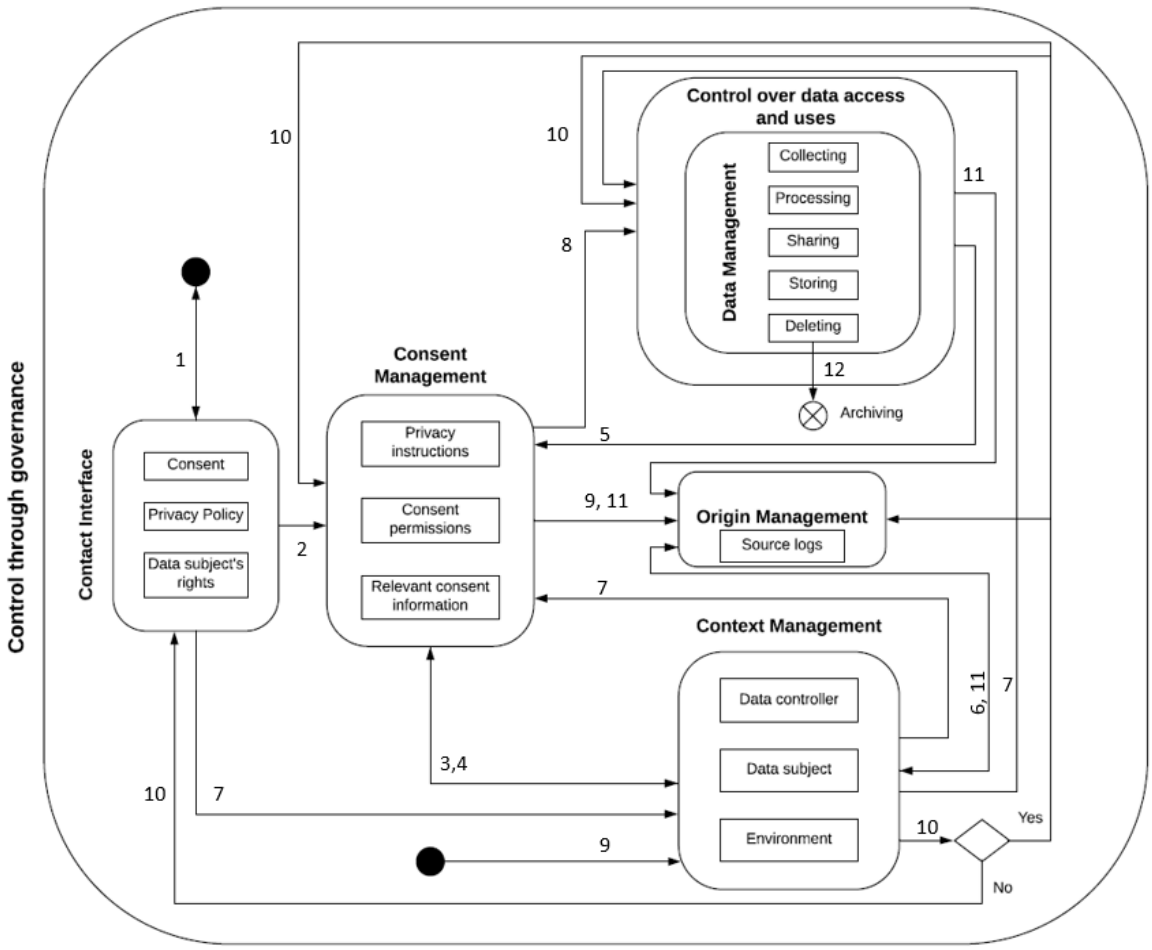


Figure 1. Components of GDPR compliant data management model and interactions between them

5.2 Scenario of obtaining consent for the purposes of marketing

In following scenario of obtaining data subject's consent for the purposes of marketing, which is described in Table 4, data subject can choose whether his data will be used for the purposes of marketing.

Table 4. Scenario of obtaining consent for the purposes of marketing

Scenario of obtaining consent for the purposes of marketing	
1	Data controller seeks data subject's consent for the purposes of marketing through Contact Interface. If the data subject decides to give a consent for the purposes of marketing, he can choose the preferred communication channel(s) of receiving notifications about special offers and novelties. There are multiple options to choose from: e-mail, SMS, telephone, Viber, Skype, Messenger, etc.
2	After the decision is made, the consent is stored in Consent Management. Consent management records the consent by extracting and storing relevant information on consent validity, obligations and permissions.
3	Consent management shares information about consent origin with Origin management and Context Management. Context Management registers the context of the consent.
4	If the data subject withdraws the consent for processing his personal data for the purposes of marketing, the new context is being generated and stored in Context management. Context Management initiates interaction with Data Management, which immediately stops processing and removes all personal data, and Consent Management, which updates the consent accordingly. If the consent validity has expired, Context Management receives notification from the Consent Management.
5	In case the data controller wants to notify data subject about special offers and novelties via Messenger, he first needs to check if he is allowed to do so. Data Management initiates the interaction with Consent management and checks consent permissions before making a decision. It the data subject has given a consent for communication via Messenger, he will receive notifications. If he did not give the consent, data controller can consult Consent Management about other options of notifying data subject (via e-mail, SMS etc.). If data subject has given the consent for receiving notifications via other communication channel, data controller can send him special offers and novelties, but if he didn't give the consent, the notification should not be sent.
6	Origin Management notes changes of context.
7	Data management stores permissions for the intended data processing, which are in line with obtained consent for the purposes of marketing. In case of change of consent permissions by data subject, Contact Interface informs Context Management about the context of changes that have occurred. Context management sends the information to Consent Management and Data Management. Consent Management updates the information on consent permissions to restrict the access to affected data, while Data Management checks the updated consent and accordingly adjusts the data collection process.
8	After it detects the changes of consent, Consent Management updates information related to data processing as mentioned in step 7 and sends new consent to Origin Management and Data Management. Data Management then stops data processing and checks updated permissions for further data processing. If the data subject has given permissions for processing personal data for the purposes of marketing, Data management will execute the task, otherwise it will not.
9	Context Management collects information on modifications made by data controller, data subject or environment.
10	Context Management then checks consent permissions in Consent Management to determine if there is a need to communicate with data subject. If the context of the consent has changed, it initiates interaction with Contact Interface in order to get the new consent from data subject. If there is no need for the new consent, it repeats step 6 and 7.
11	Origin Management keeps history of all activities related to consent, and records the process of obtaining consent from Contact Interface, consent information from Consent management, and all the activities that were using consent from Data management. It keeps track of when the request was received and what changes were performed. Origin Management also keeps records of archived consents as described in Step 7. In case personal data were used for the purposes of marketing before receiving the information on withdrawal of consent, it would be possible to prove the processing was valid at that moment.
12	If the user requires deletion of personal data related to purposes of marketing, Data Management deletes and archives the information, including data and consent origin. In case data subject requests termination of contract with data subject, any personal data that is no longer needed will be destroyed. Data origin and processing activities will be archived, since they are required for demonstrating compliance with GDPR.

6 Conclusions with future research

In this paper, state of the art of consent and data management was presented. Based on the analysis of GDPR requirements and existing data management models, guidelines for managing consent and data in ICT businesses were defined. Important concepts of consent were identified and the guidelines for the framework of consent management were presented. The data management model for ICT businesses, which meets the requirements of GDPR related to the specificity and uniqueness of the data, was proposed. Model was then explained by simple scenario of obtaining consent for the purposes of marketing. Given guidelines can help data controllers to demonstrate compliance with GDPR.

Further work should focus on defining solutions and technologies that could be used to implement the proposed model.

References

Article 29 Data Protection Working Party, Guidelines on Consent under Regulation 2016/679, (2016). https://iapp.org/media/pdf/resource_center/wp29_consent-12-12-17.pdf, accessed November 4, 2018.

Bertino, Elisa, Ji-Won Byun, and Ninghui Li, (2005). Privacy-Preserving Database Systems. *In* Foundations of Security Analysis and Design III. Alessandro Aldini, Roberto Gorrieri, and Fabio Martinelli, eds. Pp. 178–206. Berlin, Heidelberg: Springer Berlin Heidelberg. http://link.springer.com/10.1007/11554578_6, accessed March 23, 2018.

Bunnik, Eline M., A. Cecile J.W. Janssens, and Maartje H.N. Schermer, (2014). Informed Consent in Direct-to-Consumer Personal Genome Testing: The Outline of A Model between Specific and Generic Consent: Informed Consent in Direct-to-Consumer Personal Genome Testing. Bioethics 28(7): 343–351.

Casassa Mont, Marco, Siani Pearson, Sadie Creese, Michael Goldsmith, and Nick Papanikolaou, (2011). A Conceptual Model for Privacy Policies with Consent and Revocation Requirements. *In* Privacy and Identity Management for Life. Simone Fischer-Hübner, Penny Duquenoy, Marit Hansen, Ronald Leenes, and Ge Zhang, eds. Pp. 258–270. Berlin, Heidelberg: Springer Berlin Heidelberg. http://link.springer.com/10.1007/978-3-642-20769-3_21, accessed March 21, 2018.

Coiera, E. (2003).E-Consent: The Design and Implementation of Consumer Consent Mechanisms in an Electronic Environment. Journal of the American Medical Informatics Association 11(2): 129–140.

Fatema, Kaniz, Ensar Hadziselimovic, H. J. Pandit, et al. (2017). Compliance through Informed Consent: Semantic Based Consent Permission and Data Management Model. *In* 5th Workshop on Society, Privacy and the Semantic Web–Policy and Technology (PrivOn2017), C. Brewster, M. Cheatham, M. d'Aquin, S. Decker and S. Kirrane, Eds, CEUR Workshop Proceedings, Aachen Pp. 1613–0073.

Friedman, Batya, Edward Felten, and Lynette I Millett (2000). Informed Consent Online: A Conceptual Model and Design Principles. University of Washington Computer Science & Engineering Technical Report 00–12–2: 8.

Hunton and Williams, (2016) A Guide for In-House Lawyers. https://www.huntonprivacyblog.com/wp-content/uploads/sites/18/2015/06/Hunton-Guide-to-the-EU-General-Data-Protection-Regulation.pdf.

Karjoth, Günter, Matthias Schunter, and Michael Waidner (2003). Platform for Enterprise Privacy Practices: Privacy-Enabled Management of Customer Data. *In* Privacy Enhancing Technologies. Roger Dingledine and Paul Syverson, eds. Pp. 69–84. Berlin, Heidelberg: Springer Berlin Heidelberg. http://link.springer.com/10.1007/3-540-36467-6_6, accessed March 21, 2018.

Millett, Lynette I., Batya Friedman, and Edward Felten (2001). Cookies and Web Browser Design: Toward Realizing Informed Consent Online. *In* Pp. 46–52. ACM Press. http://portal.acm.org/citation.cfm?doid=365024.365034, accessed March 21, 2018.

Mont, Marco Casassa, Siani Pearson, Gina Kounga, Yun Shen, and Pete Bramhall, (2009). On the Management of Consent and Revocation in Enterprises: Setting the Context. HP Laboratories, Technical Report HPL-2009-49: 11.

Regulation (EU) 2016/ 679 (2016). Official Journal of the European Union L 119/1: 88.

Rissanen, Teemu (2016). Public Online Services at the Age of MyData: A New Approach to Personal Data Management in Finland: 12.

Rowan, W., Y. O'Connor, L. Lynch, and C. Heavin (2017.) Exploring User Behaviours When Providing Electronic Consent on Health Social Networks: A 'Just Tick Agree' Approach. Procedia Computer Science 121: 968–975.

Russello, Giovanni, Changyu Dong, and Naranker Dulay (2008). Consent-Based Workflows for

Healthcare Management. *In* Pp. 153–161. IEEE. http://ieeexplore.ieee.org/document/4556594/, accessed March 23, 2018.

Thuraisingham, Bhavani (2005). Privacy Constraint Processing in a Privacy-Enhanced Database Management System. Data & Knowledge Engineering 55(2): 159–188.

Vayena, Effy, and Alessandro Blasimme (2017). Biomedical Big Data: New Models of Control Over Access, Use and Governance. Journal of Bioethical Inquiry 14(4): 501–513.

Gamification in a Business Context: Theoretical Background*

Barbara Šlibar, Dijana Plantak Vukovac, Sandra Lovrenčić, Martina Šestak, Darko Andročec
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, Varaždin
{bslibar, dijana.plantak, sandra.lovrencic, msestak2, dandrocec}@foi.hr

Abstract. *Gamification is a recent research and practical concept for use of game design elements in non-game contexts. In this work, we study the gamification in business context. The mentioned context is very important, due to its usage by companies who are looking for new ways to engage, teach, reward and retain employees and customers. First, we list a state of the art of game studies and a transition from game design to gamification design. We also list main guidelines for application of gamification in business organizations. The effects of gamification and some ethical considerations are also tackled.*

Keywords. Gamification, gamification frameworks, gamification in business context, gamification design

1 Introduction

Having fun during work presents a great motivator for most of the people while they perform their daily job routines. The possibility of "playing" or "gaming" in the environments that have different purposes of those imposed by usual playground environment has led to emergence of a new research and practical concept called "gamification". Although the roots of gamification could be traced back to early 20th century in a form of a toy surprise in a sweets box, and applications of game elements could be found in the non-gaming project by Bartle or in a Malone's study on motivational impact of games (Werbach, n.d.), the term gamification in a current sense is associated with a computer game developer Nick Pelling (Pelling, 2011; Werbach, n.d.; Werbach & Hunter, 2012). Around 2002 / 2003 Pelling has coined the term to denote usage of game-like user interface design for electronic devices in order to make them fun and easy to use (Pelling, 2011). Later, the concept of *gamification* has found its application in many other areas, e.g., in education, health, lifestyle, marketing, business, etc.

Gamification has been identified as a promising concept that helps solve real business problems, so it was introduced to Gartner's Hype Cycle as a technology trigger in 2012 (Gopaladesikan, 2012). In 2017, gamification reached the slope of enlightenment for a digital workplace ("Gartner Releases 'Hype Cycle for the Digital Workplace", 2017), which denotes widespread use of technologies with mature products and acknowledgement of its benefits to the enterprise (Gartner Inc., n.d.).

Due to its broad scope, the term gamification has not been defined in academic literature until 2011, and the most common definition of gamification is "the use of game design elements in non-game context" (Deterding, Dixon, Khaled, & Nacke, 2011). This definition includes various contexts and enhances its relatedness to game studies. The business context of gamification is explained as "the application of gaming metaphors to real life tasks to influence behaviour, improve motivation and enhance engagement" (Marczewski, 2012b), which underpins motivational and behavioural change of a person (employee or customer), or by Gartner's redefined definition as "the use of game mechanics and experience design to digitally engage and motivate people to achieve their goals" (Burke, 2014), which limits the usage to digital environments.

Constant change of terminology is a common occurrence in every new emerging area. Besides, most of those who are in touch with new area are using the available terminology without additional categorical separation. Hence, the same issue is found when talking about gamification (Mora, Riera, Gonzalez, & Arnedo-Moreno, 2015).

Using gamification as innovative and entertaining way to engage employees and build their skills, is further highlighted by several Gartner predictions, of which the one saying that "80% of current gamified applications will fail to meet business objectives" by 2014 ("Gartner Says by 2014, 80 Percent of Current Gamified Applications Will Fail to Meet Business Objectives Primarily Due to Poor Design," 2012),

* This paper is published and available in Croatian language at: http://ceciis.foi.hr

highlights the importance of good gamification design to achieve gamification goals.

Therefore, this paper unveils the gamification concept related to the business context. Firstly, importance of game theory for gamification is highlighted, as well as transition from game design to gamification design. Secondly, the guidelines that should be considered before application of gamification in business environment are also emphasized in this paper. Thirdly, the paper points out the effects of gamification, which can be observed from different aspects. Finally, some ethical considerations regarding gamification are described.

2 Theoretical Background of Gamification

Gamification has its roots in games and game theory, video game industry, as well as in human-computer interaction (HCI) researches. Having said that, we should have in mind that early adoption of gamification was not related to digital but to military environment where the game elements such as badges and ranks were used as means of achievement (Dicheva, Dichev, Agre, & Angelova, 2015).

Various researchers (e.g. (Deterding et al., 2011; Groh, 2012; Huotari & Hamari, 2012; Salen & Zimmerman, 2004)) have elaborated the relation between gamification and games, explaining the difference between the *game*, as a set of structured activities defined by explicit rules to achieve defined goals, and the *play*, which usually comes in free, unstructured, expressive form.

Gamification has its foundations in games since it incorporates some of game design elements, but with different purpose, which is engagement and behavioural change of the user in non-gaming environment, in contrast to entertainment and enjoyment, which is the main purpose of video games (Groh, 2012). Further, gamification does not include the use of actual video game for serious purposes (which is recognized as "serious game"), but rather borrow principles of game design and apply it in non-game situations (Deterding et al., 2011; Robson, Plangger, Kietzmann, McCarthy, & Pitt, 2015).

Increasing popularity of video games has raised interest of HCI researchers to explore rationale behind designing enjoyable user interfaces by proposing heuristics for its design and methods to evaluate user experience. The same can be applied to gamified applications, which Deterding and his associates see as a re-purposed and new extension of games beyond entertainment (Deterding et al., 2011) .

Since gamification borrows its constructs from the game studies, the following chapters unveil concepts related to both game and gamification.

2.1 MDA and MDE Frameworks

MDA (Mechanics, Dynamics and Aesthetics) framework was developed by Hunicke, LeBlanc and Zubek (2004), and it is an approach for understanding games. The need for a formal and recognized proposal in the context of game design led to its development. It is trying to connect or shorten the difference between game design and development, game criticism and technical game exploration. According to MDA framework, games can be divided into three elements: rules, system and fun. These elements are translated to the following design components: Mechanics, Dynamics and Aesthetics. The Mechanics component is describing some specified game components at the level of data representation and algorithms. The Dynamics component describes the influence of behaviour of mechanics on the player inputs and each other's outputs during the execution time. Aesthetics is a third component describing desirable emotional response, which is induced in a player during interaction with a game (system). During the design process, the game components should be defined in the same order as they are listed (Hunicke, LeBlanc, & Zubek, 2004).

Robson at al. (2015) have adduced that the term emotions is more convenient for results of the synergy, which a company can obtain from relation employee – customer compared to the term aesthetics. Therefore, the MDE framework is more commonly used then MDA within gamification, where the letters MDE stand for Mechanics, Dynamics and Emotions (Robson, Plangger, Kietzmann, McCarthy, & Pitt, 2015). Emotions are player's state of consciousness. It can be said that emotions are results of player's tracking mechanics and generating dynamics.

2.2 DMC Pyramid

Werbach and Hunter have showed the structure of the game elements more precisely, commonly referred as the DMC Pyramid, and indicated that the basic elements of the game, as well as the gamification, are dynamics (D), mechanics (M) and components (C) (Werbach & Hunter, 2012):

- Dynamics - represent conceptual structures on which the game is based, the ones that shape the game itself. Additionally, they are the most abstract element of the gamification. Players feel their action, but they do not encounter them directly (e.g., narratives, relationships or constraints).
- Mechanics - can be described as processes that initiate action in the game. They are also defined as actions through which higher ranked dynamics are carried out and are reflected in lower ranked components. Typical mechanics are competitions, challenges, rewards, resource acquisition, etc.
- Components - represent specified game structures or instances of mechanics and dynamics. The most

common components are badges, levels, quests, points, achievements, leaderboards, avatars, etc.

Mccarthy and Gordon (2011) have mentioned that mechanics equalize the organizational systems and technologies, which managers can use to encourage some specific behaviour and achieve better results (McCarthy & Gordon, 2011).

As Robson et al. (2015) have described, mechanics are decisions that designers make to specify aims, rules, settings, context, type of interaction and the situation boundaries which would be gamified. They are known before the experience begins, they remain constant meaning that they should not be changed regardless of the player, and they should stay the same every time when the user is participating (Robson et al., 2015).

2.3 Importance of Participants in a Gamified Environment

One of the building blocks of the game and gamification are players whose participation should be voluntary (Huotari & Hamari, 2012). Players are the users who interact with the game or gamified application, so in order to create a satisfying user/player experience, characteristics of the players should be considered when designing gamified environment.

The most recognised taxonomy of the player types is the one proposed by Bartle (Bartle, 1996; Kumar & Herger, 2013), who classified the players into four categories. Players can be: 1) *Achievers*, who play to gain points and status, and their actions are directed towards that goal; they represent around 10% of the players, 2) *Explorers*, who love to discover new aspects of the game and figure out how things work; also represent around 10% of the players, 3) *Socializers*, who play to have fun while interacting with others, and to build inter-player relationships rather than to achieve points or finish the game; around 80% players falls into this category, and 4) *Killers*, who have similar goals as achievers, but find satisfying to see other players lose as a consequence of the killer's action; less than 1% players falls into this category.

However, Bartle himself pointed out that his taxonomy doesn't fit well into non-massively multiplayer online games or non-game related systems, including gamified ones. Marczewski has been exploring Bartle's taxonomy taking into account players' intrinsic and extrinsic motivation for better understand why and how people would use a gamified system, He proposed the Marczewski's Gamification User Types, a taxonomy for users in gamified systems (Marczewski, 2014). His taxonomy includes six types of users at a basic level. Four user types are intrinsic types motivated by their inner drive: *Achiever* (wants to learn new things), *Socialiser* (wants to interact with others), *Philanthropist* (wants to enrich the lives of others) and *Free Spirit* (wants to create and explore), whereas two user types are extrinsic types motivated by external incentives: *Player* (motivated by rewards) and *Disruptor* (motivated by change).

Those user types can be further categorized according to their willingness to play/participate, which is an important criterion when introducing gamification to a working environment. Only the *Player* is thoroughly motivated and happy to participate in gamification, in contrast to *Disruptor* who doesn't want to do anything with it. Other types are less willing to play, so gamification designers need to choose dynamics and mechanics that will encourage positive behaviour and best outcome of gamified system (Marczewski, 2015).

2.4 Gamification Frameworks

Marczewski has proposed a development process framework of gamification considering two parts. The first part consists of a set of questions, which should be asked when deciding if gamification should be used or not. The second one includes list of things which should be known about gamification and which should never be forgotten (Marczewski, 2012a).

Table 1. Marczewski's Gamification Framework

FIRST PART OF MARCZEWSKI'S GAMIFICATION FRAMEWORK	
QUESTION/STEP	EXPLANATION OF QUESTION/STEP
1. WHAT is gamified?	It should be clarified which activity or activities will be gamified.
2. WHY is gamified?	Important question to ask is what can be achieved with such/this project.
3. WHO are the users?	It should be known who are or who the users will be to make it easier to connect with them.
4. HOW is gamified?	Once the first three questions are answered, one can start thinking about the steps to be done in order to start gamification of the system.
5. Are the ANALYTICS set up?	The specific metrics and analytics should be defined because of the performance measurement, work control, reporting etc.
6. Is it TESTED on users?	Gamification should always be tested on the target group. It is the users who are going to be part of such system, not managers or designers.
7. Is it REACTED on feedback?	Feedback can be collected during testing. If some actions or changes are made regarding to feedback, then feedback collecting would make sense.
8. Is the solution RELEASED?	The new system should be announced before it is released in order to gain users even before they see the system.
SECOND PART OF MARCZEWSKI'S GAMIFICATION FRAMEWORK	
ABOUT GAMIFICATION	EXPLANATION

It should think like game designers	Since gamification is based on the game theory it is understandable that gamification designers rely on game design principle.
It should be created to be voluntary	Voluntary players are much better players than those who are not.
The plan for cheaters should be created	Cheating is a part of human nature, so it shouldn't be surprising if someone will try to cheat the system, particularly if extrinsic reward is included.
Intrinsic vs. Extrinsic	Intrinsic motivation is more powerful than extrinsic. Extrinsic motivation encompasses what has been done because of the extrinsic rewards - something tangible or material. For extrinsically motivated people, outcomes are important and not action or behavior. On the other hand, intrinsic motivation encourages behavior which results with intrinsic rewards like enjoyment, positive feelings, happiness, etc.
It should not be evil	Gamified system should not be created to exploit people, otherwise they will probably stop to use it.
The fun has not been forgotten	Even small amount of fun can make almost everything more endurable.
It should have social elements	Social mechanics are key for creating long-term engagement.

The questions asked in the first part in some way represent steps in the development process. Therefore, steps 6 and 7 can be repeated in the circle (as a loop) as many times as needed, and then steps 5 to 8 should be repeated. In order to keep the initial interest, it is necessary to collect feedbacks, to improve iterations, and to add new elements (Marczewski, 2012a).

During the implementation of a gamified system, the extrinsic rewards and intrinsic motivation should be taken into consideration. However, intrinsic motivation is the one from which longevity and genuine engagement would emerge.

Hence, this is why Marczewski has developed the RAMP (Relatedness, Autonomy, Master, Purpose), an intrinsic motivation framework that can be used as foundation for gamified systems. *Relatedness* is component, which describes the urge for connecting to other people, and is associated with the *Socialiser* user type; the *Autonomy* means as low level of control as possible, and is associated with the *Free Spirit* user type; the *Mastery* represents the never ending process of working on personal development, and is associated with the *Achiever* user type; and last but not least *Purpose* is there to bring the value to finished actions and is associated with the *Philanthropist* user type.

3 Principles and Strategies for Gamification

Before starting the implementation of gamification in a particular environment, variety of principles should be considered. Gamification designers could consult general principles for gamification, e.g., those that are based on self-determination theory which describes needs in intrinsic motivation and are further explained in (Groh, 2012), or principles for gamification of the working environment proposed in (Oprescu, Jones, & Katsikitis, 2014).

Groh (2012) has elaborated the principles that are based on: 1) *Relatedness*, the universal need to interact and be connected with others, 2) *Competence*, the universal need to be effective and master a problem in a given environment, and 3) *Autonomy*, the universal need to control one's own life. Those are similar to the already presented RAMP motivation framework (Marczewski, 2012a).

Oprescu et al. (2014) have elaborated ten principles that could facilitate gamification in everyday workplace processes. Among others, they put the accent on persuasive elements, learning orientation, amusement factors, personal and organizational wellbeing, knowledge-based, as well as adaptation to Y generation.

Whichever principles are adopted, they present the foundations for applying strategic decisions related to implementation of gamification in working environment.

3.1 Strategic Application of Gamification Principles in Business Context

In (Robson, Plangger, Kietzmann, McCarthy, & Pitt, 2016), the authors have described five guidelines, which can be used by managers and gamification designers when thinking about strategic application of gamification principles to engage employees and/or clients:

1. Before making a decision about gamification mechanics, it is necessary to understand the players, i.e., evaluate the type of players, and select the appropriate gamification mechanics,
2. Timing of rewards is key – the progression mechanics should reward player's good behavior (after his successful performance) as soon as possible to increase the player's motivation and the probability of him repeating the desired behavior,
3. New layers, tasks or players should be added only if necessary to keep the gamified experiences interesting and challenging for the players,
4. Managers must act as referees and monitor the experiences to prevent players from breaking the rules, and negatively affect other participants, and
5. Gamification mechanics should be used to keep track of the score by using appropriate metrics and

key performance indicators (KPIs) to measure the efficacy of gamification strategy applications.

On the other hand, Kumar and Herger (2013) have emphasized the importance of gamification mission. Each set of actions related to gamification, that is carried out in the business environment, and has the purpose of achieving goals, can be defined as the gamification mission. Defining the gamification mission is an extremely important step, because the properly defined mission of the game can greatly affect the performance of the game. The authors have defined the following three steps of specifying the gamification mission (Kumar & Herger, 2013):

1. Understanding the current situation (scenario) in the business environment leads to a better definition and understanding of the need for introducing gamification,
2. Understanding a target business situation in the business environment and the benefits that are to be achieved and expected from the introduction of gamification, and
3. Identification of the SMART mission (Specific, Measurable, Actionable, Realistic, Timebound) to define the mission based on the current and expected future situation (identified in steps 1 and 2).

For example, the authors have described defining a gamification mission to identify the employees. In this case, the current situation is a large company where employees don't know each other, while the expected situation is the one, in which they will increase the interconnectedness and co-operation of employees in the workplace. The gamification mission to achieve this is to help employees get to know more each other's in the next 3 months. Mechanisms that can be used for this are, e.g., that every user needs to identify a random employee based on his/her face when logging in with the company's system.

3.2 Design and Development of Gamification in Business Context

Marczewski has proposed a development framework for business gamification, which includes three phases: definition, design and improvement, where each phase contains a series of repeating steps (Marczewski, 2017a). In the definition phase, it is necessary to define the problem that is being solved by the gamification, the users involved in the process of gamification, and which final results will be considered successful (metrics for measuring these results need to be defined). The design phase includes activities related to the "user travel" design, which aim is to familiarize and understand the concepts of user experience, which users encounter during their "user travel". To do this, in the next steps you need to design and build mechanisms to achieve the desired user experience through designing the desired behaviour (What do we want users to do?), motivation (What motivates users?), emotion (How do we want our system users feel during its use?), and mechanics that implement previous designs and ensure that they are properly implemented. At the end of the improvement phase, actions are carried out to continually check the success of the steps taken and change, in case of need, undertaken activities.

Werbach and Hunter have proposed six steps for execution of the design phase of the business system gamification (Werbach & Hunter, 2012):

1. Define business objectives - it is necessary to have a clear definition of business objectives and objectives of the target system performance rather than organizational mission, etc. The definition includes listing goals in the form of a list, ranking goals, deleting mechanics, and justifying identified goals.
2. Describe target behavior - define what we want users to do and the associated metrics. Target behavior should be clearly and precisely described (for example, exercise for at least 30 minutes, sign up at the company's website, visit a restaurant, etc.), with a view to promote the achievement of the business goals defined in Step 1.
3. Describe players - it is necessary to define who will be the users of the system, who are the employees, who are the users (clients), what can motivate them within the system for its use, and categorize them as different types of players.
4. Create activity cycles that are used in the system to identify and represent user actions in a gamification system (for example, a user performs an action that results in a subsequent action, etc.).
5. Do not forget fun - Before implementing the system, it is necessary to make a last check of how much fun this system is, or whether it will be fun enough to motivate users to use it.
6. Use appropriate tools - In the implementation phase, it is necessary to use the correct mechanics and components and properly integrate them into the system. There are two options for system implementation at the technical level: a) it is possible to create a custom system implementation, b) it is possible to use platforms that offer software-as-a-service solutions or embeddable components (e.g., Freshdesk, GamEffective, Playvox, Badgeville, etc.).

In his work, Swacha (2016) has said that the success of implementing a gamification system of a business organization significantly increases if attention is paid to this process and if it is planned, i.e., according to the defined plan/procedure. Apart from the design process, Swacha has emphasized the importance of technology for the implementation of the gameplay system. Possible implementation approaches are that the system is implemented as a module within the Enterprise Information System, or as a separate software solution that integrates into the initial system (Swacha, 2016).

4 Effects of Gamification

Kappen and Nacke have presented a framework or kaleidoscope of effective gamification, where they propose a definition that "effective gamification is influencing human behaviour through engaging experiences, using game design principles in decision-making applications and services" (Kappen & Nacke, 2013). This kaleidoscope has several layers from inside out, in its core being the effective gamification (Kappen & Nacke, 2013):

- Motivated Behaviour Layer - intrinsic (e.g., competence) and extrinsic motivation (e.g., badges),
- Game Experience Layer (e.g., challenges),
- Game Design Process Layer (e.g., interface design elements), and
- Percieved Layer of Fun.

Hamari and associates have studied effects of gamification through three elements: motivational affordances (dependent variables), psychological outcomes (dependent and independent variables) and behavioral outcomes (dependent variables) (Hamari, Koivisto, & Sarsa, 2014). Their review of 24 empirical studies showed that (Hamari, Koivisto, & Sarsa, 2014):

- there is a large variety of used motivational affordances (badges, leaderboards, challenges…),
- half of studies researched psychological effects, such as enjoyment, and almost all studies (22) researched effects on behaviour (participation, learning, content contribution…),
- most of research showed positive effects for specific motivational affordances, but those effects depend on number of factors (motivational affordances, system used) and their longevity is questionable, and
- gamification in the context of education, work and organizational system is most researched, and there were no cases of gamification in marketing research.

Another research that comprised 30 scientific papers concentrated on various elements in relation to gamification: the concept of gamification itself and its use, critics, connected concepts, frameworks, theoretical background and terminology, as well as various application domains (education, health, marketing…) (Seaborn & Fels, 2015). The effect of gamification on participants is also investigated, and results showed that it is mostly positive, but that it depends on context, such as application domain. Also, from several expected different effects, in some cases one part of them was positive and another part negative or one part positive and another neutral. Some research results even varied from one participant to another depending on, for example, age or gender. Authors also emphasize that efficacy could be increased if extrinsic motivators would follow intrinsic.

Of course, individual elements or mechanics were also investigated. For example, Hamari has conducted a research to find out how badges (extrinsic motivation) influence user activity (Hamari, 2017). His research was carried out on a platform for personal sales and purchases among individuals and lasted two years – in first year there was no gamification and in second gamification was implemented with badges as rewards. Results showed that gamification increased user activities in system usage, sales and commenting. Lieberoth has researched intrinsic motivation in a setting where students thought that they are using a tool for grading their satisfiability which university considers buying, and they did not receive any reward (Lieberoth, 2015). Each group was filling in different questionnaire: (1) regular, (2) with game artefacts (board, cards, figures) that didn't have special purpose and (3) with artefacts and mechanics of game (moving figures under certain conditions). It was found out that only framing a certain activity in gamification increases feelings of fun and interest, and that mechanics and real gamification additionally increase interest very little.

Effect of various gamification elements can also be observed only from design aspect. Recent research studied points, badges, leaderboards, performance graphs, meaningful stories, avatars and teammates exclusively as design elements and their influence (specific element for specific purpose) on satisfaction of psychological needs connected to intrinsic motivation elements: competence, autonomy (linked to decision making and task meaningfulness) and social relatedness (Sailer, Hense, Mayr, & Mandl, 2017). Authors have concluded that gamification alone does not has a significant effect, but that specific elements that are used influence the individual motivational aspects.

Companies dealing with the development of the contact centre systems always emphasize the importance of gamification, stating that it can increase productivity and with fun at work stimulate the employee's positive behaviour (Calabrio, n.d.-a), and they obligatory include it in their systems (Calabrio, n.d.-b). One of 20 ways to increase employee engagement that the company Puzzel (Puzzel, n.d.) describes is also a gamification, where they give an example of a company whose employees have even completed non-mandatory courses through it, resulting in increased customer satisfaction and reduced call time, but the Puzzel has also warned of possible negative effects in the case of wrong design and application. Various examples show that gamification can increase employee engagement and efficiency in targeted activities (CallMiner, 2016).

Therefore, gamification can have various positive effects on organization, but from the aforementioned research and information it can be concluded that it needs to be accessed individually. Attention should be payed to both intrinsic and extrinsic motivation as well as to individual elements, mechanics and the profile of the participants.

5 Ethical Considerations

For now, there is no accepted/agreed ethical framework of gamification, code of ethics for gamification or ethical guidelines which should be considered and which are generally accepted. Nevertheless, there are authors who are trying to highlight the importance of ethics.

Kim and Werbach (2016) have tried to make framework for gamification ethics. They were encouraged because of next two reasons. First, the gamification ethics is partly underrated and poorly theorized because gamification is a technological novelty. Adoption of gamification in practice is often much faster than detailed consideration and theoretical research. Other reason is that proponents and critics of gamification tend to generalize according to specific examples. Kim and Werbach (2016) have proposed conceptual mapping of gamification ethics which consists of four categories of moral concerns – exploitation, manipulation, harm, character. The map is an outline of approach that can help gamification provider to take ethical issues in consideration but does not seem to be a complete framework for normative evolution of gamified systems (Kim & Werbach, 2016).

Raftopoulos has proposed Sustainable Gamification Design (SGD) model as a conceptual framework for gamification design and in midst of it she put values and ethics frame (Raftopoulos, 2014). Proposed design phases of SGD model are *Discover* (context and actors of the system), *Reframe* (discovered information as opportunities and potential solutions), *Envision* (a preferred solution), and *Create* (the gamified application) and they in total consist of seven steps of design process. Values and ethics frame are established in the first phase and its purpose is to deal with so called "value-destroying" gamification elements that author also identified, such as "coercive participation" or "loss of human agency". Values and ethical principles are inspected at each of seven steps.

Within the design of gamification and gamified systems, Marczewski defines ethics as set of principles which should facilitate the solution design process which should be in balance with the promotion of desirable outcome for users. The emphasis should be placed on designer's intention to create systems that help, and not those that cause harm to others. However, it should be kept in mind that defining the harm can be potentially subjective. Thus, it is useful to have frameworks or ethical guidelines that would prevent potential danger when designers become focused on implementation and forget the potential issues and dangers. Therefore, it is very important to know that all cases of ethical concern related to gamification are not fault of gamification as a concept, but designers whom should use available techniques to make gamification ethical (Marczewski, 2017b).

Also, there is Open Gamification Code of Ethics and it can be found on the website http://ethics.gamified.uk/. It includes next five aspects: honesty, integrity, transparency, quality and respect. This code of ethics is voluntary and has no legal obligation. Since the last update (July 2017.), it has been signed by 72 people and list of all those who agree with it is located on the same website ("Open Gamification Code of Ethics," n.d.)

6 Conclusion and Future Work

In this work, we have studied the literature on gamification in a business context. First, we have described game studies related to gamification and a transition from game design to gamification design. Gamification has its foundations in games, game theory and research of human-computer interaction. MDA framework is usually recommended for design process of gamification by many authors in reviewed literature. This framework is taken from the game design theory and it was developed because of need for a formal and recognized proposal in the context of the game design. In the analysed literature, it is mentioned that term Emotions is more suitable then term Aesthetics for the non-gaming environment. So, the used framework in the gamification context is MDE framework (Mechanics, Dynamics, and Emotions). Marczewski's framework consists of set of questions and theoretical concepts which should be considered during the gamification design, but that doesn't mean that intrinsic motivation framework RAMP should not be neglected. DMC Pyramid show the structure of the game elements even more precisely. Also, in order to create satisfying user/player experience, characteristics of the users should be considered during gamification design step.

Before starting implementation of gamification in a particular environment, gamification designers should consult general principles for gamification. The adopted principles present the foundations for applying strategic decisions related to implementation of gamification in working environment. Effects of gamifications are considered mostly positive but are also very dependent on various factors. They therefore have to be approached individually for each case with both intrinsic and extrinsic motivation in mind.

Since the gamification is considered as technological novelty, the gamification ethics is still not sufficiently explored and theorized. There are some conceptual maps, guidelines etc., but complete framework for gamification ethics has yet to be developed. In our future work, we plan to study application of gamification in contact centres. One of the key elements of creating a successful gaming experience is certainly the understanding of all kinds of players and the understanding of what they are different to each other. This theme is also promising topic for future research. We also hope our overview presented in this paper will help researchers and practitioners to explore the state of the art of a

gamification in a business context, and to identify other interesting future research areas.

Acknowledgment

This paper is based on research findings from a project „Customer Experience of the Future - Smart Specialization and Contemporary Communication and Collaboration Technologies“ - HYPER, funded by European Union funds (ESI). Authors want to express gratitude to the HYPER project partners.

References

Bartle, R. (1996). Hearts, clubs, diamonds, spades: Players who suit MUDs. *The Journal of Virtual Environments*, *1*.

Burke, B. (2014, April 4). Gartner Redefines Gamification. Retrieved June 13, 2018, from https://blogs.gartner.com/brian_burke/2014/04/04/gartner-redefines-gamification/

Calabrio. (n.d.-a). 4 ways to rock contact center gamification. Retrieved from https://www.calabrio.com/contact-center-gamification-infographic/

Calabrio. (n.d.-b). The Definitive Primer on Contact Center Employee Engagement. Retrieved from https://www.calabrio.com/definitive-primer-employee-engagement/dl-dpee/

CallMiner. (2016). Using Gamification to Improve Contact Center Performance. Retrieved from https://callminer.com/wp-content/uploads/2016/08/Using-Gamification-to-Improve-Contact-Center-Performance.pdf

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments* (pp. 9–15). ACM.

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: a systematic mapping study. *Journal of Educational Technology & Society*, *18*(3), 75.

Gartner Inc. (n.d.). Hype Cycle Research Methodology. Retrieved March 26, 2018, from https://www.gartner.com/technology/research/methodologies/hype-cycle.jsp

Gartner Releases “Hype Cycle for the Digital Workplace, 2017.” (2017, August 17). Retrieved June 13, 2018, from https://www.gartner.com/newsroom/id/3785664

Gartner Says by 2014, 80 Percent of Current Gamified Applications Will Fail to Meet Business Objectives Primarily Due to Poor Design. (2012, November 27). Retrieved March 21, 2018, from https://www.gartner.com/newsroom/id/2251015

Gopaladesikan, S. (2012, December 11). Following Gamification Through Gartner’s Hype Cycle. Retrieved June 13, 2018, from http://www.gamification.co/2012/12/11/following-gamification-through-gartners-hype-cycle/

Groh, F. (2012). Gamification: State of the art definition and utilization. *Institute of Media Informatics Ulm University*, *39*, 31.

Hamari, J. (2017). Do badges increase user activity? A field experiment on the effects of gamification. *Computers in Human Behavior*, *71*, 469–478. https://doi.org/10.1016/j.chb.2015.03.036

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does Gamification Work? -- A Literature Review of Empirical Studies on Gamification (pp. 3025–3034). Presented at the 2014 47th Hawaii International Conference on System Sciences, IEEE. https://doi.org/10.1109/HICSS.2014.377

Hunicke, R., LeBlanc, M., & Zubek, R. (2004). MDA: A formal approach to game design and game research. In *Proceedings of the AAAI Workshop on Challenges in Game AI* (Vol. 4, pp. 1–5). AAAI Press San Jose, CA.

Huotari, K., & Hamari, J. (2012). Defining Gamification - A Service Marketing Perspective. In *Proceeding of the 16th International Academic MindTrek Conference*. https://doi.org/10.1145/2393132.2393137

Kappen, D. L., & Nacke, L. E. (2013). The kaleidoscope of effective gamification: deconstructing gamification in business applications (pp. 119–122). ACM Press. https://doi.org/10.1145/2583008.2583029

Kim, T. W., & Werbach, K. (2016). More than just a game: ethical issues in gamification. *Ethics and Information Technology*, *18*(2), 157–173. https://doi.org/10.1007/s10676-016-9401-5

Kumar, J., & Herger, M. (2013). *Gamification at Work: Designing Engaging Business Software*. Aarhus, Denmark: The Interaction Design Foundation. https://doi.org/10.1007/978-3-642-39241-2_58

Lieberoth, A. (2015). Shallow Gamification: Testing Psychological Effects of Framing an Activity as a Game. *Games and Culture*, *10*(3), 229–248. https://doi.org/10.1177/1555412014559978

Marczewski, A. (2012a). A Simple Gamification Framework / Cheat Sheet. *Gamified UK - #Gamification Expert*. Retrieved from https://www.gamified.uk/gamification-framework/

Marczewski, A. (2012b, March 13). Gamification for your company. Retrieved June 13, 2018, from https://www.capgemini.com/2012/03/gamification-for-your-company/

Marczewski, A. (2014, January 28). Marczewski's Gamification User Types. Retrieved March 12, 2018, from https://elearningindustry.com/marczewski-gamification-user-types

Marczewski, A. (2015). User Types. In *Even Ninja Monkeys Like to Play: Gamification, Game Thinking and Motivational Design* (1st ed., pp. 65–80). CreateSpace Independent Publishing Platform. Retrieved from https://www.gamified.uk/user-types/

Marczewski, A. (2017a). A Revised Gamification Design Framework. Retrieved March 20, 2018, from https://www.gamified.uk/2017/04/06/revised-gamification-design-framework/

Marczewski, A. (2017b). The ethics of gamification. *XRDS: Crossroads, The ACM Magazine for Students*, *24*(1), 56–59. https://doi.org/10.1145/3123756

McCarthy, I. P., & Gordon, B. R. (2011). Achieving contextual ambidexterity in R&D organizations: a management control system approach. *R&D Management*, 240258. https://doi.org/10.1111/j.1467-9310.2011.00642.x

Mora, A., Riera, D., Gonzalez, C., & Arnedo-Moreno, J. (2015). A literature review of gamification design frameworks. In *Games and virtual worlds for serious applications (VS-Games), 2015 7th international conference on* (pp. 1–8). IEEE.

Open Gamification Code of Ethics. (n.d.). Retrieved March 20, 2018, from http://ethics.gamified.uk/

Oprescu, F., Jones, C., & Katsikitis, M. (2014). I PLAY AT WORK—ten principles for transforming work processes through gamification. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00014

Pelling, N. (2011, August 9). The (short) prehistory of "gamification"…. Retrieved May 21, 2018, from https://nanodome.wordpress.com/2011/08/09/the-short-prehistory-of-gamification/

Puzzel. (n.d.). 20 Ways to Improve Employee Engagement in Contact Centres. Retrieved from https://www.callcentrehelper.com/images/reports/20170912-Puzzel_Employee-Engagemen-in-Contact-Centres.pdf

Raftopoulos, M. (2014). Towards gamification transparency: A conceptual framework for the development of responsible gamified enterprise. *Journal of Gaming & Virtual Worlds*, *6*(2), 159–178. https://doi.org/10.1386/jgvw.6.2.159_1

Robson, K., Plangger, K., Kietzmann, J. H., McCarthy, I., & Pitt, L. (2015). Is it all a game? Understanding the principles of gamification. *Business Horizons*, *58*(4), 411–420. https://doi.org/10.1016/j.bushor.2015.03.006

Robson, K., Plangger, K., Kietzmann, J. H., McCarthy, I., & Pitt, L. (2016). Game on: Engaging customers and employees through gamification. *Business Horizons*, *59*(1), 29–36. https://doi.org/10.1016/j.bushor.2015.08.002

Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior*, *69*, 371–380. https://doi.org/10.1016/j.chb.2016.12.033

Salen, K., & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. MIT Press.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006

Swacha, J. (2016). Gamification in Enterprise Information Systems: What, Why and How (pp. 1229–1233). https://doi.org/10.15439/2016F460

Werbach, K. (n.d.). *History of Gamification*. Retrieved from https://www.coursera.org/lecture/gamification/1-5-history-of-gamification-7Wp4p

Werbach, K., & Hunter, D. (2012). *For the win: how game thinking can revolutionize your business*. Philadelphia: Wharton Digital Press.

The Use of Social Media in Political Campaigns: The Case of Croatian Local Elections 2017

Alen Kišić
Competence Center for Renewable Energy Sources
Zagrebačka 89, Varaždin
kisic.alen@gmail.com

Abstract. *Social media is a tool that allows political parties and candidates to interact with citizens. This research examines usage of social network Facebook by Croatian political candidates during the 2017 local elections. The aim of the research is twofold: to identify extent in which Facebook usage affects election results and to explore activities with the largest effect on election results. To address this questions, data from the Facebook pages of the candidates for the capital city major were collected and analysed by correlation analysis. Results indicate differences between candidates` activity which affect election results.*

Keywords. social media, campaign strategy, political campaign, election prediction

1 Introduction

Internet has enabled new approaches to communication and distribution of information. Participants in political campaigns have always used new and innovative ways of communication with the aim to reach the voters. Nowadays, internet has changed political communication and attracted the attention of researchers in the domain of political communication. The new dimension of research in this domain is achieved by the emergence of a social media platform (Strömbäck and Esser, 2009). Many authors point out that social media has changed the way of creating, distributing, and measuring political communication. Dynamic interactions and complex interdependencies at different levels brought by social media represent a challenge to traditional understanding of political communication (Strömbäck and Esser, 2009). Rapid expansion and application of social media in political campaigns around the world urged scientists to explore how does usage of social media affects political orientation, participation and attitudes of voters. Boyd states that a profound understanding of the characteristics and dynamics of social media communication provides a valuable framework for understanding the logic of social processes (Boyd, 2010). Buettner (2016) defines social media as a computer-mediated tools that allow you to create, share, or exchange information, ideas, images, or videos in virtual communities and networks. There are various forms of social media: Blogspot, LiveJournal, Yahoo! answers, Epinions, Flickr, YouTube, Digg, Reddit, microblogs (Twitter, Foursquare) and social networks (Facebook, Myspace, LinkedIn, Twitter, Tribe) as the most popular form of social media (Wright and Hinson, 2009). Social media play a significant role in politics. Recent research has been conducted with the aim of identifying the influence of social media on political campaigns from several perspectives: political participation, political knowledge and political efficiency (Rahmawati, 2014). Common characteristic for each of them is persuasion, which is crucial to political campaigns. Every speech, every phone call, or announcement through social media has the purpose of propaganda created to influence the voters. Diffusion of social networks has allowed new approaches for measuring public opinion. Election prediction of election by using large social networking data is a new form of political prediction and is mainly used to generate predictions of election results by collecting relevant social network data. Some authors suggest that analysing data collected from social networks during the pre-election campaign could be a useful addition to traditional methods (e.g. Sang and Bos, 2012; Schober et.al., 2016). The main advantages of measuring public opinion through social media are accessibility and speed (Schober et.al., 2016).

The aim of this research is to determine the frequency and means of using social media in political campaigns and associate it to the outcomes of the election. To be more specific, Facebook usage by Croatian political candidates during the 2017 local elections is examined. In order to achieve the objective, this paper analyses the data collected in the context of Croatian local elections by applying descriptive statistics, correlation analysis and ANOVA.

This paper is organized as follows. Literature review of this topic is explained in chapter 2. Chapter 3 presents research questions along with the data and methodology description. Chapter 4 presents results. Chapter 5 presents conclusion and directions for a further research.

2 Literature review

Literature review has been conducted and results shown that there is a need to monitor the content being published and to analyse the impact of that content on the targeted audience. Today, in the era of big data this is possible more than ever. Access to data provided a breakthrough from a mere assessment to data-based decision making. Which data and for what purposes have been used in the previous researches?

Housholder and LaMarre (2015) investigated relationship between expectations of the campaign on social media and information on participation in the elections that occurred as a result of activity on social media. The results indicate that social media engagement has a positive impact on the campaign. Authors point out that engagement on social networks can also help in prediction the outcome of the election. Housholder and LaMarre believe that future research should investigate to what extent certain types of engagement on social media result with desired outcomes (Housholder and LaMarre, 2015). Guleria et al. (2016) discussed the impact of an electronic campaign on the behaviour of voters, their awareness and understanding of political parties. Authors emphasizes the benefits of such campaigns in terms of wider coverage compared to traditional approaches. Their research is qualitative. The guidelines for future research indicate the need for quantitative research to prove the relationship between electronic media and political participation (Guleria et. al., 2016). Chen and Chang investigated the link between Facebook and blog usage with motivation for information and political discussion (Chen and Chang, 2017). The results of the regression analysis have shown that the desire for political discussions is significant predictor of blogging, while the motivation for information is related to the Facebook usage. Bond et al. (2012) conducted an investigation of political messages to Facebook users during of the US election to Congress in 2010. The results show that messages sent by politicians through social media have a direct impact on the elections result, not just on the person who reads the message, but also to the people interacting with them. The interpretation of the results emphasizes the importance of a large number of social network followers as factors of election success (Bond et al., 2012). Hong and Nadler have been researching whether and to what extent Twitter's use for political purposes has the potential to influence public opinion (Hong and Nadler, 2012). Their results are based on the activities of American presidential candidates on Twitter. Their results have shown that Twitter's activity is not statistically significantly associated with the number of mentioning on Twitter (Hong and Nadler, 2012).

A review of recent papers revealed that very little study so far was conducted about the impact of social media on the election results themselves. Most of these studies focus on describing social media as a marketing tool for politicians (e.g. Ahmad and Popa, 2014 or Christmann et. al., 2010). To understand the impact of social media and realize effectiveness of this tool, it is necessary to measure the use of social media for the purpose of political activities (Rahmawati, 2015). Borah (2016) highlights increasing number of research about application of social media sites for political purposes, but also points out that a small number of these surveys examine the content of the candidates' pages. On the same path are directions for further research from Praude and Skulme (2015). They highlighted the need for measuring the effectiveness of messages that are sent via social media.

3 Research methodology

An implementation of Facebook campaign has become a norm in the elections of most modern democratic societies, however it is yet to discover an impact of Facebook pages on the election results in the Croatian context. Accordingly, this study will investigate the Facebook usage during the 2017 local elections.

Following research questions are set up:

RQ1: To what extent usage of Facebook affects election results?

RQ2: Which activities have the largest effect on election results?

RQ3: Are there statistically significant differences between candidates in Facebook usage?

To address research questions, primary data were collected. Data set consists of posts on the Facebook pages of the candidates for the capital city major. The researchers limited the data to the one month activity: their page activity is monitored during the official campaign: from 21st April till 21st May. Table 1 depicts variables measured in the research.

Table 1. Variables description

Variable name	Variable description
Photos	Overall number of photos posted during the campaign
Videos	*Number of videos posted during the campaign*
Links	*Number of links posted during the campaign*
Statuses	*Number of statuses posted during the campaign*
Events	*Number of events created during the campaign*
Authors	*Number of different authors who created the page content*
Reactions	Average number of reactions on posts
Shares	Average number of posts shares
Comments	Average number of posts comments
Commenters	Average number of different commenters of posts

Reactors	Average number of different reactors
Top post	Type of most with most reactions
Top post reaction	Number of reactions on top post
Top post shares	Number of shares of top post
Top post comments	Number of comments on top post
Page likes	Overall number of page likes
Score	Percentage of votes on elections

Sociograph tool (2017) was used in order to extract data from the Facebook pages of candidates. The unit of analysis for this study was limited to the candidates for capital city major. There were eight candidates. Out of those, two candidates did not have active pages. Thus, six candidates were included in the research. Furthermore, election results in terms of candidates' percentage gained on the elections were also included. Data analysis includes use of a methodical approach to describe the findings from data, extract reasoning, and answer the research questions. This study relies on the use of statistical methods to analyse collected data. The findings were analysed by using descriptive statistics. Descriptive statistics is used to summarize and display the quantitative data. The descriptive statistical methods present data in a summarised way that the underlying information contained in the data can be easily identified (Collis and Hussey, 2013). The following section presents descriptive statistics, correlation analysis and differences testing contained in the data about social media usage for political activities.

4 Research results

This section presents the results of the empirical study: descriptive analysis of the data and results´ interpretation. This section also provides answers to the research questions.

By means of Sociograph (2017) data about six political candidates for capital city major were extracted. Out of eight candidates for capital city major, six of them (75%) had active Facebook pages during the official campaign. Candidate 1 did not have active page for whole period of campaign, just the half of period. Croatian candidates use Facebook in large extent to interact with citizens since three quarters of the candidates had active pages.

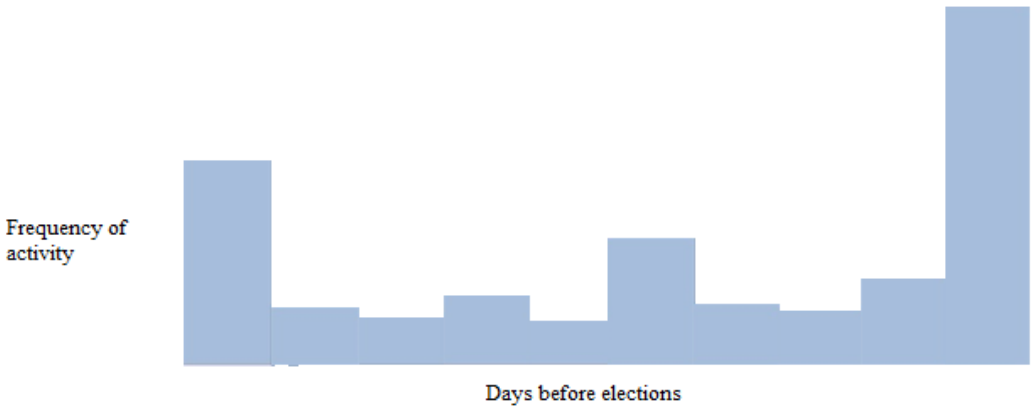


Figure 1. Frequency of candidate 1 activities

Closer look at the dynamics of each candidate activities during the observed period gave insight into frequency of the interaction with the audience. Fig 1 shows distributions of candidate 1 activities days before elections. Multimodal distribution is seen for Candidate 1. This shown to be common feature of all candidates, so only one candidate is emphasized here. This data provide a good indication that all candidates are more engaged at the beginning of campaign and on the end of the campaign. The frequency of the activities is not the same during the campaign. There are only slight differences in the dynamics of candidates' activity. In general, candidates are assigning higher relevance to the use of social media for political communication closer to the elections.

Data contained in table 2 in show social media functions used by candidates. The type of activities that the politicians engage the most are photos, following by videos and statuses. One could infer that multimedia elements are dominant here. Statuses without multimedia elements such as photos or videos are rarely used by most candidates. Exception is candidate 2.

Table 2. Social media functions

	Photos	*Videos*	*Links*	*Statuses*	*Events*	*SUM*
C1	*77*	*6*	*3*	*6*	*0*	*92*
C2	*53*	*18*	*33*	*61*	*1*	*166*
C3	40	28	17	39	4	128
C4	88	33	28	0	2	151
C5	24	20	4	5	4	57
C6	20	18	12	11	0	61
SUM	**302**	**123**	**97**	**122**	**11**	**655**

Whereas all candidates prefer photo as main medium for communication with voters, usage of other activities is different among the candidates. For instance, candidates 2 and 3 prefer statuses over videos and links, whereas candidate did not post single status. There is small number of created events for each of the candidates: candidate 3 created the highest number of events – four events. Candidate 2 posted most of the content during the observed period, following candidate 3 (see Table 2). Political candidate which were most frequent in using Facebook are second and third ranked at the elections. It is to be noted that first ranked candidate had active page only for a half of official campaign.

Candidates used almost all social media functions, but not with the same frequency. Candidates are engaged in various kinds of function with the aim to encourage citizens to participate in discussion. The harnessing of these functions can aid an effective communication between the politicians and citizens.

Table 3. Candidates' reach

	Authors	**Reactions**	**Shares**	**Comments**	**Commenters**	**Reactors**
C1	1	42	2	1	72	1127
C2	54	161	6	11	865	4508
C3	47	190	12	17	1082	3993
C4	1	115	10	7	471	2920
C5	1	331	55	39	1181	3384
C6	10	124	6	6	217	2177

In order to investigate are there statistically significant differences between candidates` activity, ANOVA was performed. Statistical test demonstrates if observed difference between the means of a parameters, is statistically significant. Analysis of variance indicated that there was a significant difference between the 8candidates` activity (Wilks' lambda = 0.5749, $p < .005$). Thus, we can reject the null hypothesis of no relationship at the 0.05 level test of independence. Variables which were responsible primarily for the overall significant difference in the activities are number of reactions and number of authors.

In order to identify reach of posts we have examined reactions and posts comments. The data presented in table 3 provides a good indication that the candidates provoke reactions on the social media. Citizens are engaged by reacting, sharing and commenting posts. Candidate 2 had the largest number of individual reactors on the posts (4508), following by candidate 3 and candidate 5. Candidate 5 stands out as candidate with highest number of average reactions, shares and comments.

This amplifies the opportunity for candidate to take advantage of the social connectedness on social media and influence the decision of followers and their friends politically. In line with the results from previous chapter that shows that a good percentage of the functions are actively accepted from citizen's. While this category did not indicate that they agree with the presented content, they chose to react on it. Under the right circumstances and with the right content, this category of people can arguably be targeted as potential voters.

A further analysis was carried out on the level of post to drill down categories of post with highest number of reactions. This is to discover the function on which the respondents are more active politically and the owners of which function are likely to cause reactions on social media politically.

Table 4. Top post statistics

	Top post type	*Top post reaction*	*Top post shares*	*Top post comments*
C1	Photo	283	8	16
C2	Link	4070	67	237
C3	Video	3682	286	309
C4	Photo	765	26	98
C5	Video	2134	2001	174
C6	Status	1295	60	48

The above result shows that the link is post with highest number of reactions, whereas video is post with highest number of shares. Video and photo are most frequent top post among candidates. The most commented post is video. It can therefore be deduced that the popularity of multimedia elements application across different spectrum makes social media phenomenon that has come to stay. This makes it necessary for politicians to explore the possibility of making use of social media.

Table 5. Elections score

	Page likes	*Score*
C1	11 708	30,87
C2	22 280	24,48
C3	11 008	19,14
C4	12089	5,6
C5	14092	2
C6	3144	1,88

Social media provides the platform for the politicians to interact, and develop connections online that can generate election success. This could provide increased, participation and drive election victory. Table 5 depicts number of pages likes and election score measured by percentage votes gained by each candidate. Raw data gives potential to explore interdependencies among activity and election results, so we have performed correlation analysis. Results are presented in table 6.

Table 6. Correlation analysis

	Authors	***Reactions***	***Shares***	***Comments***	***Commenters***	***Reactors***	***Top post reactions***	***Top post shares***	***Top post comments***	***Score***
Authors	-	0.09	-0.29	-0.02	0.48	0.76	0.92	-0.28	0.81	-0.35
Reactions	0.09	-	0.91	0.98	0.87	0.59	0.47	0.90	0.56	0.47
Shares	-0.29	0.91	-	0.96	0.66	0.25	0.11	0.99	0.23	0.45
Comments	-0.02	0.98	0.96	-	0.83	0.47	0.37	0.96	0.48	0.39
Commenters	0.48	0.87	0.66	0.83	-	0.85	0.77	0.64	0.88	0.12
Reactors	0.76	0.59	0.25	0.47	0.85	-	0.91	0.22	0.91	0.04
Top post reactions	0.92	0.47	0.11	0.37	0.77	0.91	-	0,12	0.93	-0,13
Top post shares	-0.28	0.90	0.99	0.96	0.64	0.22	0.12	-	0.22	0.40
Top post comments	0.81	0.56	0.23	0.48	0.88	0.91	0.93	0.22	-	-0.11
Score	-0.35	0.47	0.45	0.39	0.12	0.04	-0,.3	0.40	-0.11	-

Table 6 examines correlations between candidates score measured as rank and social media activity seen as: number of authors on candidates' page, number of reactions, shares, comments, commenters and reactors of posts. The results indicate a statistically significant correlation among candidates' score at the elections and number of reactions, number of comments and number of top post shares. Those correlations are positive indicating that candidates with higher number of followers` reactions and comments achieved better result. There is also statistically significant, but negative correlation between candidates result and number of authors of content at the Facebook page of the candidate. Negative correlation coefficient indicates that candidates with higher number of posts achieved lower score at the elections. Top post reactions, shares and comments are highly correlated with number of overall reactions. High number of statistically significant correlations between observed variables points out the need for further investigation of relationship and indicates possible need to group these indicators in index of candidates' activity.

5 Conclusion

This research aims to provide insight into the use of social media in political campaigns. Politicians are attempting to connect with their target market via social media, but not all of them are gaining competitive edge from it. Literature review suggested that the concept of social media is associated with elections results and those links should be explored. Therefore, the social network activities of candidates provided a basis for the study that was carried out on the usefulness of social media in political campaigns. Analysis presented here tried to identify how politicians are taking advantage of social media.

The contributions of this paper are as follows. First, this paper provides an extensive literature review on the social media for political communication. Second, based on the empirical results, high level of interest for social media based communication of political parties is identified. Results also indicated high level of interest of citizens' to engage with reactions, comments and shares. Furthermore, this study reveals that politicians are willing to engage in social media and citizens have the need to stay updated about current political situation and about politicians' reputation in social media. Social media are excellent tool to detect new trends and identify influential politicians. In Croatian context, there are yet untapped opportunities for the use of social media in politics. This could be an opportunity for the politicians to engage more via social media. As a limitation, our study focuses on Croatian politicians and only one city. In the future research should be repeated in different legislative periods and with politicians from other cities. Also, further investigation of interdependencies should be investigated by means of advanced statistical methods and machine learning approaches. Research results must be interpreted with caution because social networks are only one part of campaigns and they are focused on specific segment of people. Deep examination of election prediction should include other groups, e.g. elder people without social media profiles.

References

Ahmad, N., & Popa, I. L. (2014). The social media usage and the transformation of political marketing and campaigning of the emerging democracy in Indonesia. In Social Media in Politics, Springer International Publishing, 97-125.

Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. Nature, 489(7415), 295-298.

Borah, P. (2016). Political Facebook use: Campaign strategies used in 2008 and 2012 presidential elections. Journal of Information Technology & Politics, 13(4), 326-338.

Boyd, D. (2010). Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications. In Networked Self: Identity, Community, and Culture on Social Network Sites (ed. Zizi Papacharissi), 39-58

Buettner, R. (2016, June). Personality as a predictor of Business Social Media Usage: an Empirical Investigation of Xing Usage Patterns. In PACIS (p. 163).

Chen, C. Y., & Chang, S. L. (2017). User-orientated perspective of social media used by campaigns. Telematics and Informatics, 34(3), 811-820.

Christmann, S., Melcher, J., Hagenhoff, S., Stock Gissendanner, S., and W. Krumbein (2010). In Web 2.0-Technologien in politischen Meinungsbildungsprozessen von Parteien: Ein Beispiel aus der Praxis (Fähnrich, K.-P. and Franczyk, B. Eds.), 687-698, INFORMATIK 2010 - Service Science -Neue Perspektiven für die Informatik, Band 1, Leizpig, Germany.

Collis, J., & Hussey, R. (2013). Business research: A practical guide for undergraduate and postgraduate students. Palgrave Macmillan.

Guleria, A., Sharma, A., Bansal, D., & Sharma, G. (2016). The Impact of the Electronic Campaign by the Political Parties on the Voters, Review of Applied Socio- Economic Research, 7(1), 5-15.

Hong, S., & Nadler, D. (2012). Which candidates do the public discuss online in an election campaign? The use of social media by 2012 presidential

candidates and its impact on candidate salience. Government Information Quarterly, 29(4), 455-461.

Housholder, E., & LaMarre, H. L. (2015). Political social media engagement: Comparing campaign goals with voter behavior. Public Relations Review, 41(1), 138-140.

Praude, V., & Skulme, R. (2015). Social Media Campaign Metrics in Latvia. Procedia-Social and Behavioral Sciences, 213, 628-634.

Rahmawati, I. (2014). Social media, politics, and young adults: the impact of social media use on young adults' political efficacy, political knowledge, and political participation towards 2014 Indonesia general election (Master's thesis, University of Twente).

Sang, E. T. K., & Bos, J. (2012). Predicting the 2011 Dutch senate election results with twitter. In: Proceedings of SASN 2012, the EACL 2012 Workshop on Semantic Analysis in Social Networks, Avignon, France, 53-60.

Schober, M. F., Pasek, J., Guggenheim, L., Lampe, C., & Conrad, F. G. (2016). Research synthesis social media analyses for social measurement. Public Opinion Quarterly, 80(1), 180-211.

Sociograph, available at: https://sociograph.io/report.html, accessed: 25.08.2017.

Strömbäck, J., & Esser, F. (2009). Shaping politics: Mediatization and media interventionism. In: Mediatization: Concept, Changes, Consequences, New York: Peter Lang Publishing Group, 2009, 205-223.

Wright, D. K., & Hinson, M. D. (2009). An analysis of the increasing impact of social and other new media on public relations practice. In 12th annual International Public Relations Research Conference, Miami, Florida, 1-22.

Digital Transformation Insights and Trends

Igor Pihir, Katarina Tomičić-Pupek, Martina Tomičić Furjan
University of Zagreb, Faculty of Organization and Informatics
Department of Information Systems Development, Department of Organization
Pavlinska 2, Varaždin, Croatia
{igor.pihir, katarina.tomicic, martina.tomicic}@foi.hr

Abstract. *Digital transformation (DT) is an emerging paradigm, which introduces strategy-oriented and customer-centric changes in infrastructure and processes in modern organizations relying on contemporary information and communication technologies (ICT). Transforming in DT is not a one-time process, it is a holistic approach of shifting organizations towards implementation of new methods for raising organizational performances by boosting the organizational capabilities and competitiveness, creating thereby new models of doing business. This paper provides an insight into digital transformation in the form of an overview for the purpose of defining digital transformation and identifies its drivers. A brief bibliographic analysis of available literature is presented. Key determinants of digital transformation are defined and an overview of the methodologies used to assess the digital maturity of enterprises is made. Contextual influence factors and examples of enablers of digital transformation were discussed and existing trends and technologies in this challenging and promising field are positioned.*

Keywords. digital transformation, literature insights, key determinants, digital maturity, future trends.

1 Introduction

Digital transformation of enterprises is a new paradigm in the context of implementing contemporary technologies to set new products and services and change the mind-set of delivering them to the global market. It is defined as "the profound transformation of business and organizational activities, processes, competencies and models to fully leverage the changes and opportunities of a mix of digital technologies and their accelerating impact across society in a strategic and prioritized way, with present and future shifts in mind." [i-SCOOP.eu, 2016].

All digital changes in organizations can be grouped into one of the three categories [Westerman et al, 2014, pp 108]:

1. **Substitution**, where digital technologies are used to replace a function/process that is already performed in the organization;
2. **Extension**, where digital technologies are used to improve the functionality of a process/product; and
3. **Transformation**, where digital technologies are used to fundamentally redefine a process/product.

Digitization has not penetrated yet equally in all industries, but it has begun to transform many of them, and it has a significant impact on the economic performance of companies within those industries. As the growth continues, the implications for revenues, profits, and opportunities will be dramatic [Bughin et al, 2018, pp 2].

While identifying the main drivers of digital transformation, we came across several statements that might sound superfluous but, in our experience, depict well what is being put in front businesses as demands from the market or as requests from customers:

- Customer expectations continue to rise [Tiersky, 2017]
- SPEED is more important than ever [Tiersky, 2017]
- Digital transformation means business transformation [Tiersky, 2017]
- Unlocking Data Silos [Candito, 2017]
- Information "On the Go" [Candito, 2017]
- Intelligent Work Processes [Candito, 2017]
- Innovate or Die [Innovate or die, 2018]
- Everyone is doing it [Roche, 2018]
- With enormous change comes opportunity [Roche, 2018]
- At the heart of the change is reducing costs and driving greater efficiencies [Roche, 2018]

With aim to explore the developments in this area, researches were conducted and insights provided into the field of digital transformation as part of the research project (IRI) *Development of innovative*

platform for digital transformation of enterprises. Following analyses were made:

- **literature analysis on digital transformation**, including the analysis of existing scientific databases, journals, papers and areas or activities pertinent to digital transformation (as explained in chapter 2);
- **key determinants of digital transformation**, focusing on industrial areas and contextual factors which contribute to technological improvements with the purpose of digital transformation of organizations (as explained in chapter 3);
- **methodologies of assessing digital maturity of enterprises**, which influences and enables digital transformation of businesses (as explained in chapter 4);
- **contextual influence factors and examples of enablers** in the context of digital transformation (as explained in chapter 5);
- **new trends and technologies** in the field of digital transformation (chapter 6).

The following chapters present insights into and trends in digital transformation.

2 Literature analysis on digital transformation

With the aim to obtain some insights into this research field, the literature was collected from Scopus as one of the most relevant and high-quality scientific databases of scientific papers in this field. Regardless of the limitations, as this literature review cannot be considered as exhaustive, it does give a valuable insight into digital transformation. The search of data was conducted on 15th March 2018, based on keyword "digital transformation" in the contribution title, and it resulted in 154 hits. (Databases like Scopus index papers all the time, from 1-3 years back so they are not stochastic!).

The papers were reviewed by the year of their publication. It was established that the term „digital transformation" in the present-day organizational and ICT sense was first defined and used between 2000 and 2003. The following years saw an increase in the amount of works dedicated to this field, which culminated in 2016, remaining high in 2017 as well as in first months of 2018. Fig. 1 shows the number of works in Scopus database published from 2000 to date.

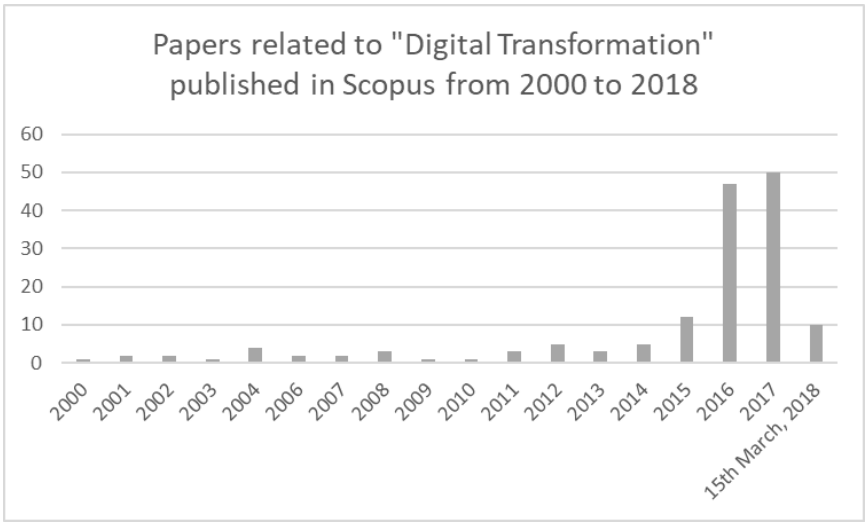


Figure 1. Works in Scopus database using term "digital transformation" from 2000 to 2018

The chart presented above shows that the number of works has significantly increased over the past three years with a tendency of further growth in 2018, given that only those published in the first quarter of the year have been indexed so far. The number of indexed works is typically higher in the second half of the year due to slower pace of indexation and publication of scientific papers. The works published before 2000 have not been taken into account because "digital transformation" mentioned there refers to the digitalization of pictures and texts from analogue media and not to the meaning of the term in the modern sense.

As regards the publication fields, the titles listed in Scopus database are found in a variety of fields of human activities. As for scientific fields, we find them, quite understandably in information/computer science, but also in all other fields of human activities, especially in business, management, accounting, engineering, social sciences, decision science, economics and finance, medicine, protection of environment, etc. Fig. 2 provides an overview of the number of works by industry, where a single paper may be listed in several fields.

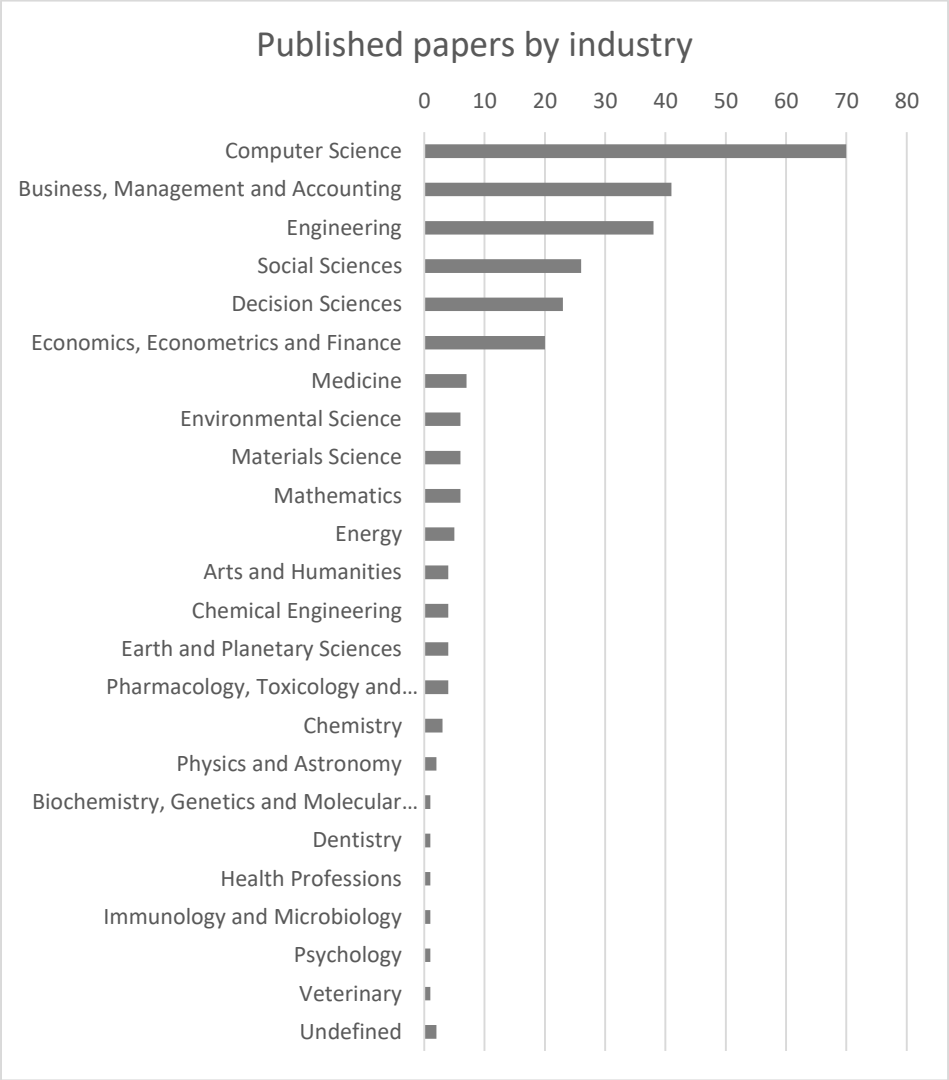


Figure 2. Published papers by industry

Judging by the number of citations, the titles listed below can be considered as the most important in the field of DT:

1. The digital transformation of traditional businesses, authored by Andal-Ancion, A., Cartwright, P.A., Yip, G.S. from 2003, which defines in detail this field from the scratch; it was published in the MIT Sloan Management Review 44(4), pp. 34-41 and was cited 55 times [Andal-Anction, Cartwright & Yip, 2003].
2. Innovation diffusion in global contexts: Determinants of post-adoption digital transformation of European companies, authored by Zhu, K., Dong, S., Xu, S.X. and Kraemer, K.L. from 2006. It was published in the European Journal of Information Systems. 15(6), pp. 601-616, and cited 182 times [Zhu et al., 2006].
3. Digital transformation: Opportunities to create new business models, authored by Berman, S.J. from 2012, published in the Strategy and Leadership 40(2), pp. 16-24, cited 35 times [Berman, 2012].

These papers constituted the cornerstone of all relevant researches in the field over the past three years and, at the same time, provide initial framework of what we today call "digital transformation".

3 Key determinants of digital transformation

The analysis of identified scopes, pillars and areas of methodologies used to assess DT made it possible to identify the convergence by key determinants of digital maturity:

a) **Strategy orientation –** vision, management, leadership
b) **Customer centricity** – monitoring of customers' experiences, prediction of their needs

c) **ICT and process infrastructure** – ICT resources, management of business processes
d) **Talent, capability and capacity strengthening** – culture of permanent investment in new skills, knowledge and capacities
e) **Innovation culture and organizational commitment** – commitment to organizational culture, innovation culture and organizational factors.

A more detailed analysis of expert reports and scientific literature and a closer insight into the features of each methodology made it possible to classify both those methodologies and the models created therefrom, by several criteria:

1. **According to the stage of the assessment methodology implementation,** to those that are implemented:
   a) *Only initially, at the beginning of the DT project* for the purpose of establishing initial state of affairs;
   b) *Both before and after digital* transformation in order to establish the initial state and measure effects, i.e. improvement;
   c) *Only at the end of the DT* project for the purpose of assessing new positions.
2. **With regard to the inclusion of technological** support, some methodologies entail:
   a) *Technological approach*, which refers to IT solutions that facilitate self-assessment by means of web applications, online questionnaires and similar platforms, or
   b) *Personalized approach* with a high level of consultant engagement in detailed assessment.
3. **According to applicability of maturity assessment methodology** in the domain of related industry:
   a) *Specialized methodologies* for maturity assessment applied to specific industries
   b) *General methodologies* for maturity assessment.

Given the key determinants of digital transformation, the conclusion is that organizations are expected to be agile. Agility refers to the set of properties required in every modern organization to enable them not only to respond to the challenges of the market but also to create new products and services as well as to create demands for them. These properties include changing environment adaptability, organizational flexibility oriented towards team work and project-type approach, high collaboration skills, evolutivity in the realization of business processes and incremental approach to the delivery of products and services.

Agility is today hard to master without proper support based on contemporary and emerging technologies which, however, should not be the starting point in creating development plans, but rather the means of its realization.

Thus, taking into consideration basic categories of digital transformation associated with the strategy orientation and focused on the customer/user, monitoring of customers' experiences, prediction of customers' needs, different technologies may have different impacts on organizations in different industries.

Some technologies involve minimum standards of survival and competitiveness (mainstream technologies), whereas some others enable industries to make a strategic step forward in relation to their competition in the market and contribute to achieving strategic advantages over their competition.

4 Methodologies of assessing digital maturity of enterprises

The paradigm of "digital transformation" gave rise to the need for organized approaches to its implementation. In order to analyse possible effects of investments into digital transformation, which is a reasonable thing to do, various methodologies of assessing digital maturity of organizations have been developed. They are supposed to be implemented in the digital transformation projects with the purpose to measure the effects of transformation.

Some examples that illustrate the types of digital maturity assessment were selected from the wide range of global and regional methodologies:

1. **Digital Maturity Model** - TM Forum (three key elements) [TM Forum, 2018].
2. **Digital Maturity Model 5.0** – Forrester (four dimensions) [Forrester, 2018].
3. **Digital Maturity Assessment Tool** - Government of South Australia [Government of South Australia, 2015]
4. **Key pillars of digital transformation** - Chief information Officer (CIO) Report (key pillars). [Evans, 2017]
5. **Framework for digital maturity of schools** – CARNet (five areas) - regional model intended for the eduational sector and developed as part of the project "e-Škole: Uspostava sustava razvoja digitalno zrelih škola (pilot projekt)" /e-Schools: Creation of the system of digitally mature schools (pilot project/ [e-Škole, 2018]

Key pillars, categories, i.e. dimensions of the selected methodologies for assessing digital maturity are given in Table 1.

Table 1: Pillars, categories and dimensions of the maturity assessment technologies

Digital Maturity Model – TM Forum (three key elements)	Digital Maturity model 5.0. - Forrester (four dimensions)	Digital Maturity Assessment Tool Government of South Australia – (five pillars)	Key pillars of digital transformation – CIO (key pillars)	Framework for digital maturity of schools – CARNet (5 areas)
Customer centricity	Culture	Governance and leadership	Strategy and vision	Planning, management and leadership
Organization and Culture	Organization	People and culture	People and culture	ICT in learning and teaching
IT and software skills	Technology	Capacity and capability	Process and governance	Development of digital competences
	Insights	Innovation	Technology and capabilities	ICT culture
		Technology		ICT infrastructure

5 Contextual influence factors and examples of enablers of digital transformation

In the course of activities conducted in previous researches, an analysis of the implementation of social platforms and media as technology-based concepts of the process enhancement was done. Relevant scientific sources were examined, scientific papers collected, and an extensive analysis of their contents was conducted. The existing models of similar researches were used as a setting for our research [Niehaves & Plattfaut, 2011]. The incidence of classification and identification of particular properties is described in the research work [Suša Vugec, Tomičić-Pupek & Bosilj Vukšić, 2018]. The authors show that, in terms of structural dimensions, the advanced technologies of the social BPM are equally applied both to build and to use the new systems. They also show that, in terms of the process dimension, the most dominant are the initiatives aimed at improving the collaboration inside the organization and the collaboration with users/customers. As for the dimension of the contents of initiatives, new systems are most often implemented with the purpose of restructuring or enhancing business processes as part of continuous management of organizational performance.

As reported, the properties incidence in the examined literature covering case studies shows that organizations pursue initiatives to introduce agility based on the following indications:

a) The initiative is focused on building a system that will be intensely exploited;
b) The initiative is focused mainly on the management and the employees in the company and, to a lesser extent, on the need to communicate with customers.
c) The initiative is focused on improvement of the collaboration within the organization and with customers, and on better distribution of tasks and roles within the process, and
d) On enhancement and restructuring, permanent management of processes and knowledge.

The results of the conducted research, published in [Suša Vugec, Tomičić-Pupek & Bosilj Vukšić, 2018], confirm previously obtained results [Niehaves & Plattfaut, 2011]. This indicates consistency of the research framework which can therefore be used to research the implementation of other technologies as well.

These initiatives for introducing agility in doing business are influencing organization's choice of suitable technologies for improving their processes and further development of an agile business architecture [Meffert, Breuer, Evers, 2018]. Improvements can be related to increasement of efficiency or effectiveness, development of ecosystems of businesses, increase of agility in order to react on disruptions of regular business models before the competitors do, creating new markets, products or services, managing customer journey and experience and other challenging issues [Pejić Bach, Spremić & Suša Vugec, 2017, Dhawan et al, 2018]. The improvement initiative is the context of designing and introducing often a mix of suitable technologies for supporting processes engaged in the digital transformation. Various technologies are being related to digital transformation like mobile and cloud technologies, social platforms, big data and Data analytics, Internet of Things, Virtual and Augmented Reality, Drones, Robotics and autonomous systems, 3D printing, blockchain, Artificial Intelligence, Reference models, knowledge management and other all oriented towards making the organization "future-ready" [Ross, 2018; Weill & Woerner, 2018]. All these new digital technologies have been developed independently of one another and independently from the trend of digital transformation, but they all can be

used simultaneously to digitally transform organization's business processes. Technologies are not all suitable for all organizations, therefore improvement initiatives as well as legacy technologies are the contextual influence factors and enablers of DT. While some of the mentioned technologies have an emerging significance in some industries, the same technology can already be mainstream for some other industry, like VR or autonomous systems that are already mainstream in the automotive industry can be seen as emerging technologies in agriculture.

6 New trends and technologies in the field of digital transformation

New development trends involve implementation of new digital technologies under the common name of Industry 4.0 or Fourth Industrial Revolution.

Thereby "the rate of the technological development in Industry 4.0 is exponential and, therefore, anticipating the challenges and even the benefits is much more difficult than what the world experienced in the previous industrial revolutions. This increased difficulty is due to the high convergence of technologies that could complement or compete with different possible diffusion scenarios that may result in more frequent breakthroughs that are difficult to forecast." [Morrar, Arman & Mousa, 2017]

IoT (Internet of Things), virtual and expanded reality, robotics, autonomous systems, mobile technology, blockchain and technologies in the field of social media and platforms have *de facto* become standards in some industries (e.g. automobile manufacturing industry, software industry, creative industries, and industries linked with financial institutions).

Schwab [2017] made a list of emerging technologies (presented as results of a survey, conducted in September 2015 by the World Economic Forum's Global Agenda Council), with their positive and negative impacts. A selection of those technologies is shown in Table 2:

Table 2: Emerging trends in digital transformation

Technology	Description	Positive impacts	Negative impacts
Implantable technologies	Devices implemented into bodies, from pacemakers, and smart tattoos to „built-in" smartphones	Beneficial to health monitoring or locating missing children	Threat to privacy and data security
Wearable Internet	Technologies in mobile phones designed to fit in clothes and accessories	Self-sufficiency and better decision making	Threat to privacy and data security, addiction
Internet of things	Connecting to „everything" on the Internet via sensors and appropriate applications	Rise in productivity, improved quality of life, safety (of food, planes…), creation of new businesses, connection with environment	Privacy concerns, loss of traditional jobs, security threats
Smart cities	Management of energy, material flows, logistics and traffic through sensors and data platforms	Rise in productivity, improved quality of life, lower rate of crime, increased mobility, better access to education	Privacy concerns, risk of system collapse, cyber-attacks
Big data	Management and use of huge amounts of data in automated decision making and real-time services customization	Better and faster decision making, cost savings, new job categories	Job losses, privacy concerns, questionable trust in data, questionable ownership of data
Driverless cars	Cars started and driven by means of built-in applications	Improved safety, lesser impacts on environment, improved mobility for the old and disabled	Job losses, cyber-attacks, lower revenue for public transportation
Robotics	Design, construction, operation, and application of robots	Substitution to people's hard work	Job losses, liability and accountability
Blockchain	Distributed trust mechanism designed to keep track of transactions	Disintermediation of financial institutions explosion of tradable assets, increased transparency	Trust of people, fear of losing financial reality

Sharing economy	Exchange of physical goods, assets or services	Increased access to resources, better asset utilization	More contract labour, decrease of grey economy, abuse of trust
3D printing	Creation of a physical object by printing it layer by layer from a drawing or model	Accelerated product development, rising demand for product designers, more personalized products	Job losses, piracy, uncontrolled production of body parts, opportunity for printing objects with high level of abuse like guns

All listed technologies have a potential to increase productivity and bring digital transformation into an organization, **either introduced or used separately**, or as a combination of both, the only challenge being the right choice.

As stated by Westerman et al.: "New digital technologies can fuel innovation and improve company performance, but only if applied in the right places" [Westerman et al, 2014, pp 36].

Digital transformation methodological guidelines such as The digital transformation compass [Westerman, et al, 2014, pp 174], Business model canvas [Osterwalder et al., 2014, pp 16] or the Digital transformation framework [Matt et al, 2015, pp 342] should help businesses to digitally change their work and/or the results thereof by leading them through the transformation. The start of the transformation depends on defined strategic goals: if the goal is driving new digital revenue, then the start should be with strengthening the digital content – the what is consumed; if the goal is cross-selling and driving more revenue per customer, the focus should be on improving customer experience – the how it is packaged; if the goal is efficiency and flexibility, then focus is to be made on building and exploiting shared digital platforms – the how it is delivered [Weill & Woerner, 2013, 73].

8 Conclusion

This paper provides an insight into the field of digital transformation and an explicit description of its key determinants and trends. A bibliographic analysis of the relevant works was done in this field through investigation of 154 papers from the Scopus database. The purpose was to show the trends and relevance of digital transformation and prove it to be a novelty in the ICT implementation approaches. Having reviewed the literature, the frequency of published works per industries was presented so as to show that digital transformation is not just a matter of technology and information science, but also that it is happening in all organizations, regardless of their form or type of business processes.

Defining key determinants of digital transformation and reviewing of the methods of assessing digital maturity of organizations made it possible to identify the existing methods. The paper describes methods and techniques available to any and all organizations which are required to identify, describe and classify their current state or key factors and, based on the findings, take a set of steps forward and move to a higher level of maturity and carry out digital transformation of business processes.

Set of contextual influence factors were defined and examples of enablers impacting the management of business processes in order to try to achieve organizational restructuring and eventual digital transformation are described. At closing, a presentation is provided of the present-day trends and technologies related to this field, the impacts of which and their interconnection allow for unique combinations of organizational and technological achievements, which in turn lead to the creation of new models of business practice by creating added value in, sometimes, most unexpected places.

In conclusion, we can only say that digital transformation through strategy-oriented and customer-centric changes in infrastructure and processes in modern organizations relying on contemporary information and communication technologies, is likely to ignite innovations and enhance performance only if "applied in the right places, as stated by [Westerman et al, 2014, pp 3]. This, however, requires a shift in mind-sets on the part of the people in whose organizations digital transformation is to be made.

Acknowledgments

This research has been conducted as part of the wider research in the project Development of innovative platform for digital transformation of enterprises (RDI) which is funded by European Union through the European Regional Development Fund (ERDF).

References

Andal-Ancion, A., Cartwright, P.A. & Yip, G.S. (2003). The digital transformation of traditional businesses. *MIT Sloan Management Review* 44(4), pp. 34-41.

Berman, S.J. (2012). Digital transformation: Opportunities to create new business models. Strategy and Leadership. Vol 40(2), pp. 16-24.

Bughin, J., LaBerge, L., & Mellbye, A. (2018). The case for digital reinvention. *Digital reinvention: Unlocking the 'how' – Digital McKinsey.* Retrieved from https://www.mckinsey.com/ business-functions/digital-mckinsey/our-insights/digital-reinvention-unlocking-the-how

Candito, N. (2017). Unmasking Digital Transformation: 6 B2B Drivers for 2018. Retrieved from http://www.industryweek.com /technology-and-iiot/unmasking-digital-transformation-6-b2b-drivers-2018

Dhawan, R., Heid, B., Küderli, P. & Laczkowski, K. (2018). How industrial companies can respond to disruptive forces. Retrieved from https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/how-industrial-companies-can-respond-to-disruptive-forces

e-Škole. (2018). Okvir za digitalnu zrelost škola. Retrieved from https://www.e-skole.hr/ hr/rezultati/digitalna-zrelost-skola/okvir-digitalne-zrelosti-skola/

Evans, N. (2017). Assessing your organization's digital transformation maturity. Retrieved from https://www.cio.com/article/3213194/digital-transformation/assessing-your-organization-s-digital-transformation-maturity.html

Forrester. (2018). The Digital Maturity Model 5.0. Retrieved from https://hootsuite.com/resources/ forrester-the-digital-maturity-model-5-0#

Government of South Australia. (2015). Digital Maturity Assessment Tool – Governance and leadership. Retrieved from https://digital.sa.gov.au/sites/default/files/content_files/toolkits/Digital_Maturity_Assessment.pdf

Innovate or die. (2018). A journey towards innovations. Retrieved from https://www.innovateordie.eu/

i-SCOOP.eu (2016). Digital transformation: online guide to digital business transformation. Retrieved from: https://www.i-scoop.eu/digital-transformation/

Matt, C., Hess, T., Benlian, A. (2015): Digital Transformation Strategies. *Business and Information Systems Engineering*. 57(5), 339–343

Meffert, J., Breuer P. & Evers M. (2018). How companies are reinventing themselves. Retrieved from https://www.mckinsey.de/publikationen /leading-in-a-disruptive-world/how-companies-are-reinventing-themselves

Morrar, R., Arman, H. & Mousa, S. (2017). The Fourth Industrial Revolution (Industry 4.0): A Social Innovation Perspective, Technology Innovation Management Review, 7(11);

Niehaves B. & Plattfaut R. (2011). *Collaborative business process management: status quo and quo vadis*. Business Process Management Journal. 17(3): 384-402.

Osterwalder, A., Pigneur, Y., Bernarda, G. & Smith, A. (2014). *Value proposition design*, Hoboken, New Yersey: John Wiley &Sons

Pejić Bach, M., Spremić, M. & Suša Vugec, D. (2017). *Integrating Digital Transformation Strategies into Firms: Values, Routes and Best Practice Examples*. Management and Technological Challenges in the Digital Age / Novo Melo, P.; Machado, C.(eds.). Boca Raton, Florida: Taylor & Francis Group: CRC press, pp. 107-128

Roche, B. (2017). The 4 Key Drivers for Digital Transformation in the Enterprise (an excerpt of a chapter from my forthcoming ebook on Digital Transformation). Retrieved from https://www.linkedin.com/pulse/4-key-drivers-digital-transformation-enterprise-excerpt-brian-roche/

Ross, J. (2018). Digital Is About Speed — But It Takes a Long Time, MITSloan Management Review, Retrieved from https://sloanreview.mit.edu/article/digital-is-about-speed-but-it-takes-a-long-time/

Schwab, K. (2017). The fourth industrial revolution. Portfolio penguin: UK.

Suša Vugec, D., Tomičić-Pupek, K. & Bosilj Vukšić, V. (2018). *Social business process management in practice: Overcoming the limitations of the traditional business process management* // International Journal of Engineering Business Management, Vol 10, pp 1-10.

Tiersky, H. (2017). The 5 key drivers of digital transformation today. CIO from IDG. Retrieved from https://www.cio.com/article/3198121/it-industry/whats-now-in-digital-transformation.html

TM forum. (2018). Digital Maturity Model Retrieved from https://www.tmforum.org/digital-maturity-model-metrics/

Weill, P., Woerner, S.L. (2018) Is Your Company Ready for a Digital Future? *MIT Sloan Management Review* 59(2), pp. 21-25.

Weill, P., Woerner, S.L. (2013) Optimizing your digital business model. *MIT Sloan Management Review* 54(3), pp. 71-78.

Westerman, G., Bonnet, D. & McAfee, A. (2014). *Leading Digital – turning technology into business transformation*. USA: Harvard business review press.

Zhu, K., Dong, S., Xu, S.X. & Kraemer, K.L. (2006). Innovation diffusion in global contexts: Determinants of post-adoption digital transformation of European companies. European Journal of Information Systems. 15(6), pp. 601-616.

ICT Entrepreneurship and Innovation

Forecasting Cryptocurrency Investment Return Using Time Series and Monte Carlo Simulation

Nikola Zornić, Aleksandar Marković
University of Belgrade, Faculty of Organizational Sciences
Department of Business Systems Organization
Jove Ilica 154, 11000 Belgrade, Serbia
{nikola.zornic, markovic.aleksandar}@fon.bg.ac.rs

Sava Čavoški
FEFA, Metropolitan University
Bulevar Zorana Đinđića 44,
11000 Belgrade, Serbia
scavoski@fefa.edu.rs

Abstract. *Cryptocurrencies are attracting significant amount of attention. Everything started with Bitcoin and built up to the situation where we have over 1500 cryptocurrencies. One can say that cryptocurrency market is the new stock market. This market is still highly volatile, but decentralized, open, and widely accessible. In this paper we will use time series analysis and Monte Carlo simulation for forecasting cryptocurrencies' return for selected time period. With huge price oscillations present it is hard to provide precise return predictions, but any step towards analysing cryptocurrencies adds to understanding the market.*

Keywords. cryptocurrency, return on investment, time series, simulation, model, Monte Carlo simulation

1 Introduction

Cryptocurrencies can be defined as digital, computer currencies whose implementation stands on the principles of cryptography, both to validate the realised transactions and to enlarge the currency in circulation (Cocco, Concas, & Marchesi, 2017).

Bitcoin is seen as the first decentralized digital currency platform, a currency without central authority to regulate it's usage, validate and settle transactions (Gandal & Halaburda, 2016). It was introduced in 2009, but only recently it's value and popularity has significantly grown. Although newer cryptocurrencies gained popularity much faster, Bitcoin (BTC) is still the one with highest market valuation, usage, merchant acceptance and popularity (Hayes, 2015).

Following Bitcoin's footsteps, other cryptocurrencies were launched (Iwamura, Kitamura, & Matsumoto, 2014). It should be noted that anyone can create its own cryptocurrency in minutes (Long, 2018). Popular name for all cryptocurrencies released after Bitcoin is *altcoins*. Some of the most popular altcoins are Ethereum (ETH), Ripple (XRP), EOS, Litecoin (LTC), Zcash (ZEC), and Monero (XMR).

In scientific literature Bitcoin is the most studied cryptocurrency. Barber, Boyen, Shi, and Uzun (2012) pointed to several problems with Bitcoin, such as technical vulnerability, potential deflationary spiral, accidental loss of bitcoins, and malware attacks. Urquhart (2016) showed that Bitcoin market returns are significantly inefficient if observed at once, on the whole sample, but when sample is split into two subsample periods, tests indicate that Bitcoin is efficient in the latter period. Yermack (2013) analysed Bitcoin price against fiat currencies and showed that its volatility undermines its usefulness as a currency. Baek and Elbeck (2015) presented strong evidence to suggest that Bitcoin volatility is internally (buyer and seller) driven – leading to the conclusion that the Bitcoin market is highly speculative. Cheah and Fry concluded that Bitcoin price is prone to speculative bubbles and that the market is highly volatile (2015b; 2016). Bitcoin showed vast success and popularity since its creation (more in the recent years), thanks to its added value (Marian, 2013). Namely, some of the most important pros of Bitcoin are anonymity, decentralised nature, enhanced revenue, and use of proof-of-work mechanisms (Moore & Christin, 2013).

Cryptocurrencies have recently become a topic of interest for scientific papers. Only 193 articles have been published by the end of 2017 in journals indexed on Clarivate Analytics Web of Science Social Sciences Citation Index (SSCI) and Science Citation Index Expanded (SCIE) with the topic "cryptocurrency" OR "bitcoin" (Clarivate Analytics, 2018; Zornić & Marković, 2018).

Most of the authors tried to determine factors influencing cryptocurrency price, trade volume, and volatility. For example, Sovbetov (2018) examined factors that influence prices of cryptocurrencies such as Bitcoin, Ethereum, Dash, Litecoin, and Monero. Those factors include cryptomarket-related factors and attractiveness of cryptocurrencies in long-run.

Glaser, Zimmermann, Haferkorn, Weber, and Siering (2014) showed that most of the interest in cryptocurrencies on Wikipedia is due to the cryptocurrencies' investment asset aspect and not due to the usage as currency itself. When cryptocurrencies

become widely usable for their primary purpose – as medium of exchange for goods and services, they will be even more interesting to analyse (Zornić & Marković, 2018). Concept of purely digital currency will highly increase convenience of payments in digital world, but at the moment their usage involves high risk. Regardless digital wallet security risk, currently the highest is risk related to cryptocurrency value due to its instability.

Aim of this paper is to provide a model for cryptocurrency return analysis using a combination of two well-known methods, time series analysis and Monte Carlo simulation, using widely accessible data source. This paper is organized as follows. Firstly, the cryptocurrency is defined and short literature review is presented. Afterwards, descriptive statistics of collected cryptocurrencies' price data is displayed, followed by time series fit functions. Those time series fit functions are used as input for Monte Carlo simulation, whose results are analysed afterwards. Finally, in the last section, concluding remarks and future directions of the research are provided.

2 Model for Forecasting Cryptocurrency Investment Return

Five cryptocurrencies with highest market capitalization on 29th of June 2018 have been chosen for building the model for forecasting cryptocurrency investment return: Bitcoin ($108.88 B), Ethereum ($45.64 B), Ripple ($17.67 B), EOS ($8.06 B), and Litecoin ($ 4.58 B). Data has been collected for the daily closing price in USD ($) for each of them. Data for all cryptocurrencies except for EOS covers the period from 01.01.2016 to 29.06.2018 (CryptoCompare, 2018a, 2018c, 2018e, 2018d), and the period from 29.06.2017. to 29.06.2018. for EOS (CryptoCompare, 2018b). There are a lot of online cryptocurrency historical price databases, but CryptoCompare is selected as the one that has prices for all the mentioned cryptocurrencies. Simple descriptive overview of the prices and natural logarithm of daily returns is presented in Table 1.

We can see that Bitcoin price has grown from an average of $567.00 in 2016 to $9131.34 in 2018. The rate has been even higher for some periods of time, reaching the maximum value of $19,345.49 on 16.12.2017. Other cryptocurrencies saw excessive growth in value, too. Ripple had the highest mean daily return of 1.57% in 2017. Other cryptocurrencies had their highest mean daily return in the same year, too.

Table 1. Cryptocurrency prices and daily return

		2016			2017			2018		
		Days	**Mean**	Std. Dev.	Days	**Mean**	Std. Dev.	Days	**Mean**	Std. Dev.
Bitcoin	**Price [$]**	366	**567.00**	138.35	365	**3981.07**	3987.18	180	**9131.34**	2288.07
	Daily return [%]		**0.22**	2.58		**0.73**	4.91		**-0.47**	5.11
Ethereum	**Price [$]**	366	**9.76**	3.67	365	**221.68**	183.80	180	**717.69**	232.40
	Daily return [%]		**0.59**	7.25		**1.24**	6.88		**-0.32**	6.09
Ripple	**Price [$]**	366	**0.01**	0.00	365	**0.20**	0.25	180	**0.90**	0.48
	Daily return [%]		**0.05**	5.76		**1.57**	11.52		**-0.84**	7.42
EOS	**Price [$]**				186	**2.22**	2.35	180	**10.41**	3.59
	Daily return [%]					**1.19**	13.35		**-0.03**	9.33
Litecoin	**Price [$]**	366	**3.76**	0.47	365	**49.85**	64.13	180	**157.09**	44.86
	Daily return [%]		**0.06**	2.78		**1.08**	8.12		**-0.62**	6.35

Time series analysis and Monte Carlo simulation is conducted using Palisade @RISK 7.5.2 plugin software for Microsoft Office Excel. Total of 11 time series algorithms have been employed for each cryptocurrencies' return. @RISK features automatic detection of required transformations to achieve time series stationarity. Akaike Fit (Akaike information criterion – AIC) and Bayesian Fit (Bayesian information criterion – BIC) are used as quality measures for time series fit. Results for the best ranked algorithms for each cryptocurrency are presented in Table 2. In addition to fitting time series data with appropriate model, forecasting for seven days is done. These fitting results are used as input for Monte Carlo simulation

Table 2. Time series fit details

Cryptocurrency [type]	Bitcoin return [MA1]	Ethereum return [MA1]	Ripple return [GARCH]	EOS return [MA1]	Litecoin return [MA1]
Data Transform	Auto Detect	Auto Detect	Auto Detect	Auto Detect	Auto Detect
Function	None	None	None	None	None
Detrend	None	None	None	First Order	None
Deseasonalize	None	None	None	None	None
Seasonal Period	N/A	N/A	N/A	N/A	N/A
Akaike (AIC) Rank	#1	#1	#1	#1	#1
Akaike (AIC) Fit	-3187.55	-2283.47	-2157.93	-527.64	-2493.80
Bayesian (BIC) Rank	#1	#1	#1	#1	#1
Bayesian (BIC) Fit	-3.19E+03	-2.28E+03	-2.14E+03	-5.28E+02	-2.49E+03
Parameters	3	3	4	3	3
Parameter #1	Mu	Mu	Mu	Mu	Mu
Value	0.002869743	0.006715739	0.004815913	-0.00022336	0.00335444
Parameter #2	Sigma	Sigma	Omega	Sigma	Sigma
Value	0.042167778	0.068930561	0.003604131	0.116822606	0.061950577
Parameter #3	B1	B1	A	B1	B1
Value	0.003730151	0.009485065	0.988957199	-0.934305919	0.022356184
Parameter #4			B		
Value			0.043549253		

Graphs with best time series fit functions, including historical data and 7-days prediction period are presented on Figures 1-5. The *x-axis* represents time, where 0 is data of data collection (29.06.2018) and the *y-axis* represents daily return. These functions and parameters will be used for creating financial model for cryptocurrency return analyses.

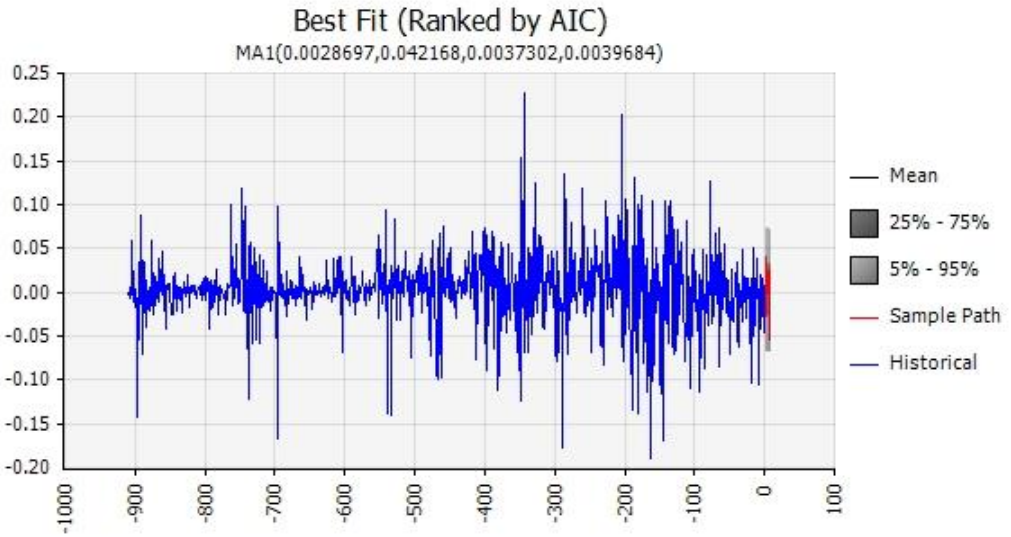


Figure 1. MA1 time series for Bitcoin daily returns

Figure 1 presents time series for Bitcoin daily returns. As we can see from historical data, oscillations are high, from -18.92% (16.01.2018 - t_{-164}) to 22.76% (20.07.2017 - t_{-344}). In the last 30 days, there were two peaks with return less than -10%. The 90% confidence interval for daily return in the future 7 days will be between -6.65% and 7.23%.

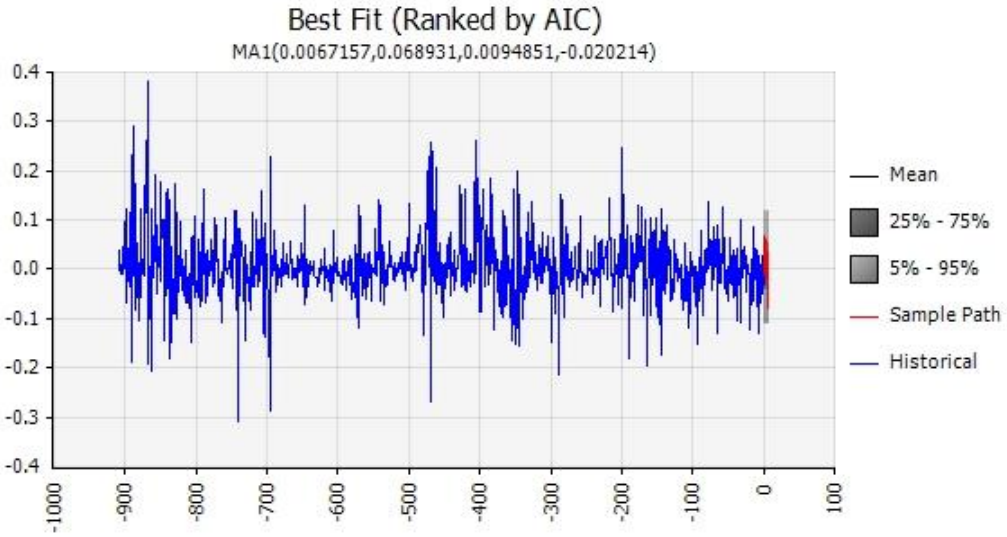


Figure 2. MA1 time series for Ethereum daily returns

Ethereum had even higher return peaks. As we can see from Figure 2, there are periods with high oscillations, from -31.01% (18.06.2016 - t_{-741}) to 38.30% (11.02.2016 - t_{-869}). The 90% confidence interval for daily return in the future 7 days will be between -10.7% and 12.00%.

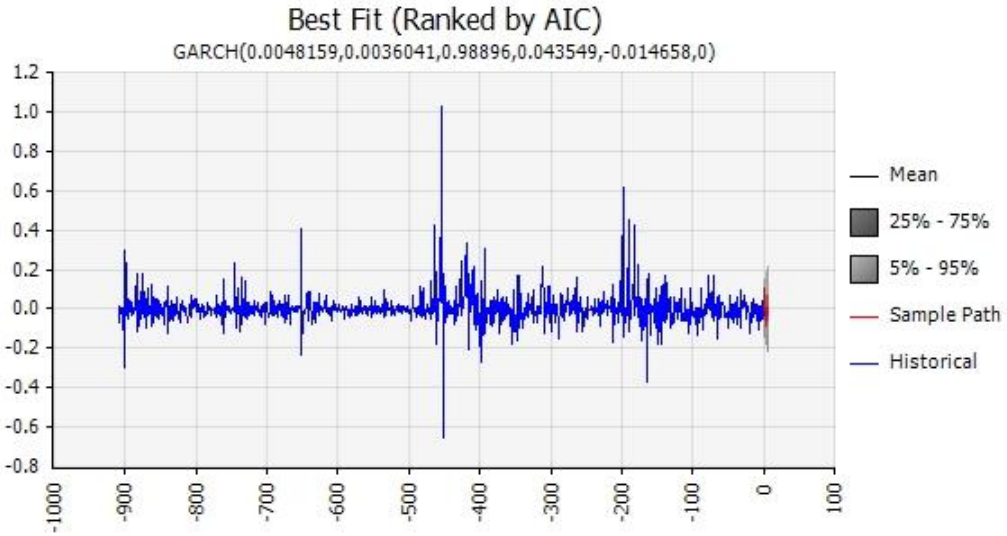


Figure 3. GARCH time series for Ripple daily returns

Ripple return had had one of the largest peaks of observed cryptocurrencies. Analysing Figure 3, we can see high peaks, from -65.30% (03.04.2017 - t_{-452}) to 102.80% (02.04.2017 - t_{-453}). The 90% confidence interval for daily return in the future 7 days will be between -20.00% and 21.00%.

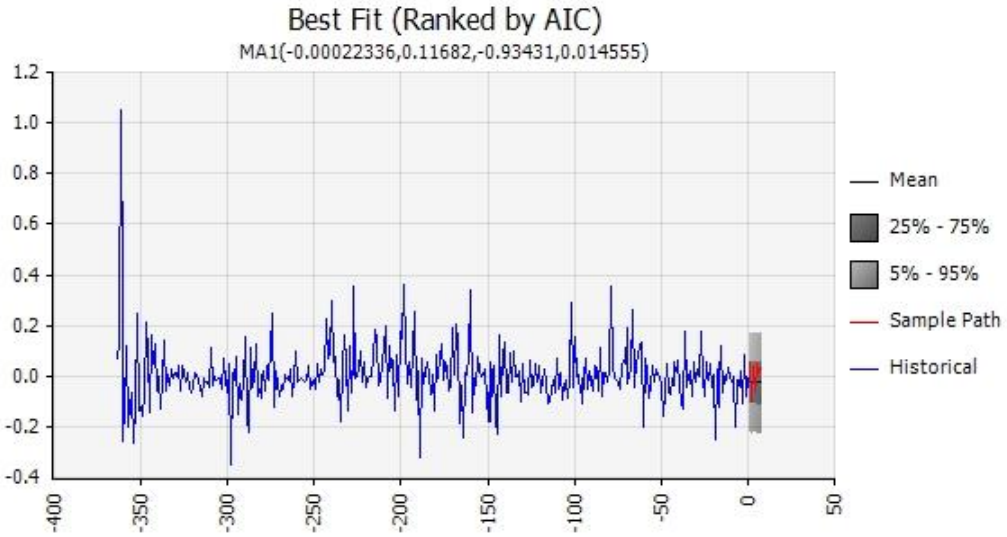


Figure 4. MA1 time series for EOS daily returns

Figure 4 presents time series for EOS daily return. Highest daily return of 104.98% was achieved 02.07.2017 - t_{-298}. On the other side, the highest decrease in value of -35.21% was achieved 04.09.2017 - t_{-362}. The 90% confidence interval for daily return in the future 7 days will be between -22.00% and 16.90%.

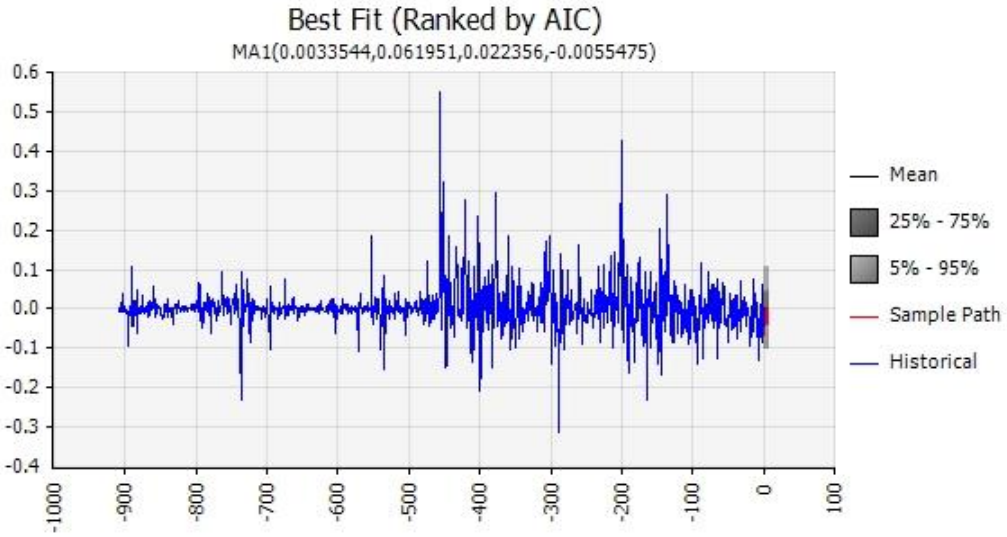


Figure 5. MA1 time series for Litecoin daily returns

Litecoin daily return time series is presented in Figure 5. Analysing time series, we can see that return is oscillating from -31.25% (14.09.2017 - t_{-288}) to 55.16% (30.03.2017 - t_{-456}). The 90% confidence interval for daily return in the future 7 days will be between -9.90% and 10.50%.

Figure 6 presents the Monte Carlo model in Microsoft Office Excel. Yellow fields are for input variables, blue for calculation, and green for output. Five different time series functions are used for calculating 7-days return and based on the results each cryptocurrencies' price is calculated.

Monte Carlo simulation was run using @RISK software with 100.000 iterations. Cryptocurrencies' return is generated using specific time series function in each iteration and afterwards other values are calculated and iteration results are saved in database. Those results are presented in form of probability distributions together with excessive statistical indicators.

	A	B	C	D
1	**Cryptocurrency**	**Price [29.06.2018]**	**Price [06.07.2018]**	**Ln(return)**
2	Bitcoin	$ 5,910.5100	$ 6,030.5132	**2.01%**
3	Ethereum	$ 414.8600	$ 434.7369	**4.68%**
4	Ripple	$ 0.4402	$ 0.4551	**3.32%**
5	EOS	$ 7.3000	$ 6.1115	**-17.77%**
6	Litecoin	$ 73.8800	$ 75.6292	**2.34%**
7				
8		=D4/B4	=EXP(D2)*B2	
9				
10				Time series functions

Figure 6. Predicted cryptocurrency 7-days returns

One can say that investment into cryptocurrencies carries high risk, as the market is highly volatile (Yermack, 2013). On the other side, potential profits are significant. Our results presented on Figures 7-11 confirm these claims. Based on our results, Ethereum showed the highest mean return and Bitcoin lowest standard deviation.

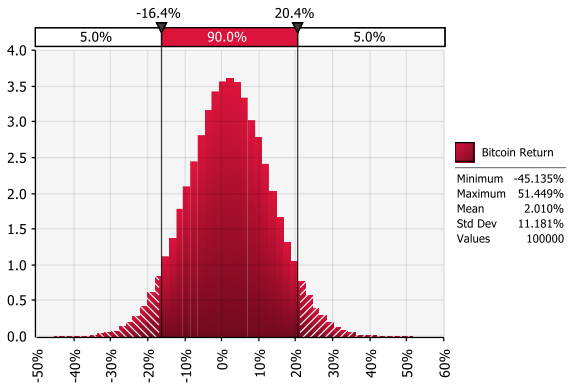


Figure 7. Bitcoin return

Average return on Bitcoin (Figure 7) for 7-days period is 2.01% with standard deviation 11.18%.

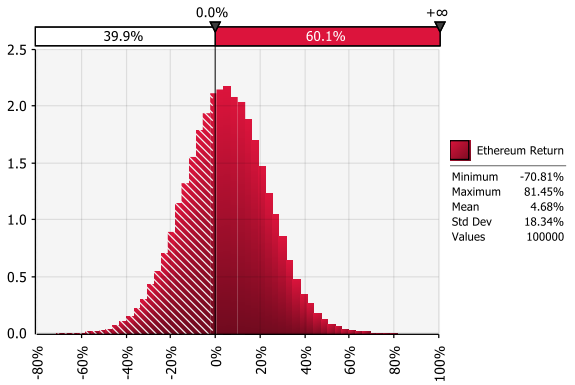


Figure 8. Ethereum return

Average return on Ethereum (Figure 8) for 7-days period is 4.68% with standard deviation 18.34%. Maximal predicted return in this period is 81.45%, and on the other side, minimal -70.81%. From the same figure we can also see the probability that profit will be achieved (60.1%).

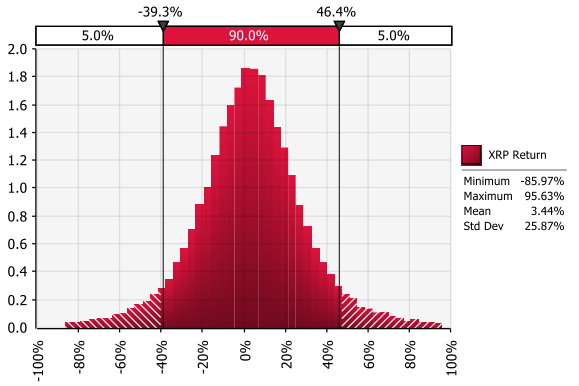


Figure 9. Ripple return

Ripple 7-days period mean return is 3.44% with standard deviation 25.87% (Figure 9).

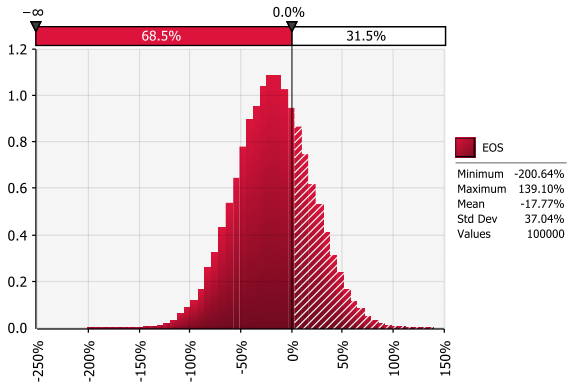


Figure 10. EOS return

EOS showed the worst result, when return on investment is observed (Figure 10). Mean return is negative (-17.77%) with standard deviation 37.04%. Probability of losing money when investing in EOS is 68.5%.

Litecoin has the second-best standard deviation – 16.76% and average return for 7-days period 2.34% (Figure 11).

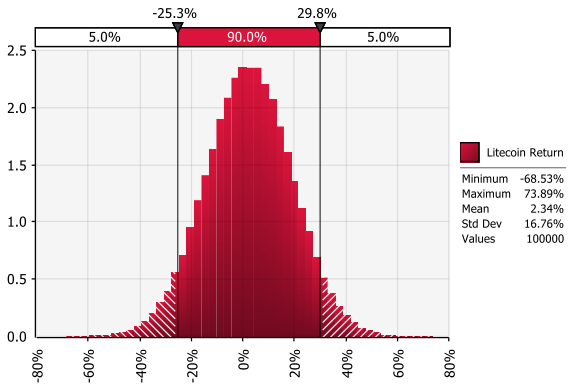


Figure 11. Litecoin return

3 Conclusion

Aim of this paper was to provide a model for cryptocurrency return analysis using a combination of two well-known methods, time series analysis and Monte Carlo simulation. Additionally, popular cryptocurrency price tracker, CryptoCompare, was used as data source.

Five cryptocurrencies were included in analysis: Bitcoin, Ethereum, Ripple, EOS, and Litecoin. Time series model was fitted for each of them and resulting @RISK functions are used as input for Monte Carlo simulation. Monte Carlo simulation results for observed cryptocurrencies show high investment risk, but also high potential profits. Based on results, all cryptocurrencies except EOS have positive mean return.

Unfortunately, we already concluded that Bitcoin price is prone to speculative bubbles and that the market is highly volatile. This means that it is currently not possible to give precise return and profit predictions based on scientific methods. Regardless, any kind of information gathering and analysis are giving higher insight into possible price trends. Provided model gives possibility to analyse results in terms of probabilities and statistical indicators.

During the research, several future directions of the study emerged. Firstly, more cryptocurrencies can be included in the analysis, especially the newly created and fast-rising ones. The other direction is including portfolio optimization features.

References

Baek, C., & Elbeck, M. (2015). Bitcoins as an investment or speculative vehicle? A first look. *Applied Economics Letters*, *22*(1), 30–34. https://doi.org/10.1080/13504851.2014.916379

Barber, S., Boyen, X., Shi, E., & Uzun, E. (2012). Bitter to Better — How to Make Bitcoin a Better Currency. In *Financial Cryptography and Data Security* (pp. 399–414). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-32946-3_29

Cheah, E.-T., & Fry, J. (2015a). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, *130*, 32–36. https://doi.org/10.1016/J.ECONLET.2015.02.029

Cheah, E.-T., & Fry, J. (2015b). Speculative bubbles in Bitcoin markets? An empirical investigation into the fundamental value of Bitcoin. *Economics Letters*, *130*, 32–36. https://doi.org/10.1016/j.econlet.2015.02.029

Clarivate Analytics. (2018). Web of Science. Retrieved March 25, 2018, from https://apps.webofknowledge.com

Cocco, L., Concas, G., & Marchesi, M. (2017). Using an artificial financial market for studying a cryptocurrency market. *Journal of Economic Interaction and Coordination*, *12*(2), 345–365. https://doi.org/10.1007/s11403-015-0168-2

CryptoCompare. (2018a). Bitcoin (BTC) - USD - Historical OHLC chart, social data chart and multiple chart indicators. Retrieved June 29, 2018, from https://www.cryptocompare.com/coins/btc/charts/USD?p=ALL&t=LC&e=CCCAGG

CryptoCompare. (2018b). EOS (EOS) - USD - Historical OHLC chart, social data chart and multiple chart indicators. Retrieved June 29, 2018, from https://www.cryptocompare.com/coins/eos/charts/USD?t=LC&p=ALL

CryptoCompare. (2018c). Ethereum (ETH) - USD - Historical OHLC chart, social data chart and multiple chart indicators. Retrieved June 29, 2018, from https://www.cryptocompare.com/coins/eth/charts/USD?t=LC&p=ALL

CryptoCompare. (2018d). Litecoin (LTC) - USD - Historical OHLC chart, social data chart and multiple chart indicators. Retrieved June 29, 2018, from https://www.cryptocompare.com/coins/ltc/charts/USD?p=ALL&t=LC&fTs=1382565600&tTs=1525125600

CryptoCompare. (2018e). XRP (XRP) - USD - Historical OHLC chart, social data chart and multiple chart indicators. Retrieved June 29, 2018, from https://www.cryptocompare.com/coins/xrp/charts/USD?t=LC&p=ALL

Fry, J., & Cheah, E.-T. (2016). Negative bubbles and shocks in cryptocurrency markets. *International Review of Financial Analysis*, *47*, 343–352. https://doi.org/10.1016/j.irfa.2016.02.008

Gandal, N., & Halaburda, H. (2016). Can We Predict the Winner in a Market with Network Effects? Competition in Cryptocurrency Market. *Games*, *7*(3), 16. https://doi.org/10.3390/g7030016

Glaser, F., Zimmermann, K., Haferkorn, M., Weber, M. C., & Siering, M. (2014, April 15). Bitcoin - Asset or Currency? Revealing Users' Hidden Intentions.

Hayes, A. (2015, March 16). What Factors Give Cryptocurrencies Their Value: An Empirical Analysis. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2579445

Iwamura, M., Kitamura, Y., & Matsumoto, T. (2014). Is Bitcoin the Only Cryptocurrency in the Town? Economics of Cryptocurrency And Friedrich A. Hayek. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2405790

Kristoufek, L. (2013). BitCoin meets Google Trends and Wikipedia: Quantifying the relationship between phenomena of the Internet era. *Scientific Reports*, *3*(1), 3415. https://doi.org/10.1038/srep03415

Long, E. (2018). How to Create Your Own Cryptocurrency. Retrieved April 29, 2018, from https://lifehacker.com/how-to-create-your-own-cryptocurrency-1825337462

Marian, O. Y. (2013, October 1). Are Cryptocurrencies "Super" Tax Havens?

Moore, T., & Christin, N. (2013). Beware the Middleman: Empirical Analysis of Bitcoin-Exchange Risk (pp. 25–33). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39884-1_3

Sovbetov, Y. (2018). Factors Influencing Cryptocurrency Prices: Evidence from Bitcoin, Ethereum, Dash, Litcoin, and Monero. *Journal of Economics and Financial Analysis*, *2*(2), 1–27. Retrieved from https://mpra.ub.uni-muenchen.de/85036/1/MPRA_paper_85036.pdf

Urquhart, A. (2016). The inefficiency of Bitcoin. *Economics Letters*, *148*, 80–82. https://doi.org/10.1016/j.econlet.2016.09.019

Yermack, D. (2013). Is Bitcoin a Real Currency? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2361599

Zornić, N., & Marković, A. (2018). Cryptocurrency Price Forecasting Using Time Series and Monte Carlo Modeling and Simulation. In *SymOrg 2018 - Doing Business in the Digital Age: Challenges, Approaches and Solutions*. Zlatibor, Serbia: University of Belgrade, Faculty of Organizational Sciences. In press.

Intelligent Information Systems

Designing Deep Neural Network Topologies with Population-Based Metaheuristics

Grega Vrbančič, Iztok Fister Jr., Vili Podgorelec
University of Maribor
Faculty of Electrical Engineering and Computer Science
Koroška 46, 2000 Maribor, Slovenia
{grega.vrbancic, iztok.fister1, vili.podgorelec}@um.si

Abstract. *Over last years the deep neural network has become one of the most popular classification methods with performance comparable and in some cases even superior to humans in the wide range of applications. However, there are still some major challenges regarding the deep neural networks. One of the biggest, with the huge impact on the classification performance, is the design of such deep neural network. In this paper, we propose a population-based metaheuristics approach for designing a deep neural network topology in a straightforward automatic manner, which performance we compare against the conventional classifiers across three different datasets. With the usage of our proposed method, unlike the conventional classifiers, we were able to achieve high classification performance with no major performance drops throughout all tested datasets.*

Keywords. machine learning, neural networks, swarm and evolutionary computation, optimization

1 Introduction

In recent years, deep neural networks (DNNs) have demonstrated performance comparable, in some cases even superior to humans in areas such as image recognition [8, 27, 6], speech recognition [7], natural language processing [5] and even in playing two-player games such as Go [21]. Such accomplishments in the application of DNNs can be largely credited to increasing computer power and a growing abundance of data. As more and more computation power becomes available, more data - bigger datasets can be used to train DNNs with more layers and more neurons at each layer, which should eventually translate to higher accuracy of such DNN models [31].

Regardless of all the major successes of utilizing the DNNs to various problems, the researchers are still facing the two major problems: the design of DNN and parameter setting for the training of the DNN. The design of DNN including the number of hidden layers, the number of neurons at each layer, the type of activation function for each layer is, generally speaking, done manually, as well as picking out the right parameters for training such DNN. However, the choice of design of DNN, as well as the choice of training parameters, has an important impact on how a DNN is going to perform a specific task. The major problem in designing the DNN topology and in picking the right training parameters is the lack of some general rules or recipes to follow, which would guarantee a good DNN performance. Basically, it depends more or less on our previous experience and trying out different DNN topologies and/or training parameters.

There are not many studies on using the population-based nature-inspired algorithms [4] for tackling the mentioned problems of using the DNNs since such approach has high computational costs. Most of the studies are focusing on solving just one part of the problem in order to keep the search space as small as possible. For example in [24, 19, 12, 28] the authors are trying to optimize the architecture of feed-forward DNNs and recurrent neural networks using evolutionary approach and in [9, 30] authors are attempting to optimize the training parameters of neural networks. Based on those encouraging results from mentioned studies, we developed a method for designing DNNs with the use of population-based metaheuristics in a straightforward automatic manner titled as GWODNN or DEDNN (Grey Wolf Optimizer for Deep Neural Network/Differential Evolution for Deep Neural Network).

The main goal of our research is to study whether a model based on DNN topology designed with our proposed population-based metaheuristics approach for designing a DNN topology in a straightforward manner, will give us generally better classification accuracy across different datasets over the conventional classifiers (e.g. k-nearest neighbor, decision tree, multilayer perceptron). The main advantages of the proposed method are the very straightforward usage and the adaptability to different datasets while achieving high classification accuracy.

The remaining of this paper is organized as follows. Section 2 briefly describes methods we used. In Section 3 we present the proposed GWODNN/DEDNN

method, whose performance of classification on various datasets are presented in Section 4. Conclusion and final remarks are gathered in Section 5.

2 Methods

2.1 Population-based metaheuristics

This subsection familiarizes readers with population-based metaheuristics, i.e. differential evolution and grey wolf optimizer.

2.1.1 Differential evolution

Differential Evolution (DE) [25] is a population-based metaheuristic algorithm that belongs to the family of evolutionary algorithms [3]. DE was introduced in 1995 by Storn and Price [25]. Because of many wins at international competitions, DE is considered as one of the most powerful algorithms appropriate for continuous optimization. DE consists of Np real-coded vectors representing the candidate solutions, as follows:

$$\mathbf{x}_i^{(t)} = (x_{i,1}^{(t)}, \ldots, x_{i,n}^{(t)}), \quad \text{for } i = 1, \ldots, Np, \tag{1}$$

where each element of the solution is in the interval $x_{i,1}^{(t)} \in [x_i^{(L)}, x_i^{(U)}]$, and $x_i^{(L)}$ and $x_i^{(U)}$ denotes the lower and upper bounds of the i-th variable, respectively. There are three different operators in DE, i.e.: mutation, crossover, and selection.

DE mutation is expressed as follows:

$$\mathbf{u}_i^{(t)} = \mathbf{x}_{r1}^{(t)} + F \cdot (\mathbf{x}_{r2}^{(t)} - \mathbf{x}_{r3}^{(t)}), \quad \text{for } i = 1, \ldots, Np, \tag{2}$$

where F denotes the scaling factor as a positive real number that scales the rate of modification while $r1,\ r2,\ r3$ are randomly selected values in the interval $1 \ldots Np$.

Crossover in DE is expressed as follows:

$$w_{i,j}^{(t+1)} = \begin{cases} u_{i,j}^{(t)} & \text{rand}_j(0,1) \leq CR \vee j = j_{rand}, \\ x_{i,j}^{(t)} & \text{otherwise}, \end{cases} \tag{3}$$

where $CR \in [0.0, 1.0]$ controls the fraction of parameters that are copied to the trial solution.

Selection is expressed mathematically as follows:

$$\mathbf{x}_i^{(t+1)} = \begin{cases} \mathbf{w}_i^{(t)} & \text{if } f(\mathbf{w}_i^{(t)}) \leq f(\mathbf{x}_i^{(t)}), \\ \mathbf{x}_i^{(t)} & \text{otherwise}. \end{cases} \tag{4}$$

2.1.2 Grey wolf optimizer

Grey wolf optimizer or simply GWO [13] is a swarm intelligence based algorithm [4], which mimics the leadership hierarchy and hunting mechanism of grey wolves. For simulating the leadership hierarchy, four types of grey wolves are employed, i.e. alpha, beta, delta, and omega [13]. In line with this, algorithm consists of three main steps:

- searching for prey,
- encircling prey, and
- attacking prey.

Basic GWO algorithm is presented in Alg. 1. For deeper outline of GWO algorithm, readers are referred to the paper [13].

Algorithm 1 Grey wolf optimizer

```
Initialize the grey wolf population X_i (i = 1, 2, ..., n)
Initialize a, A, and C
Calculate the fitness of each search agent
X_α = the best search agent
X_β = the second best search agent
X_δ = the third best search agent
while termination_condition_not_meet do
    for each search agent do
        Update the position of the current search agent
    end for
    Update a, A, and C
    Calculate the fitness of all search agents
    Update X_α, X_β, and X_δ
end while
return X_α
```

2.2 Deep Neural Network

A standard neural network (NN) [11] consists of many simple and connected processors called neurons, each producing a sequence of real-valued activations. The input neurons get activated through sensors perceiving the environment, on the other side, other neurons get activated through weight connections from a previously active neuron. Learning of such NNs is about finding weights that make the NN exhibit the desired behavior (e.g. recognize a person from the picture). Such behavior may require long causal chains of computational stages, where each stage transforms (most often in a non-linear way) the aggregate activation of the network [20].

For decades have been around shallow NN-like models with few such stages. Models with several successive nonlinear layers of neurons date back to 50s and 60s years in a previous century. In last few years, deeper NN-like models also known as deep neural networks or DNN are gaining on the popularity.

In formal terms, a feed-forward (acyclic) and deep neural network is a tuple $N = (L, T, \Phi)$, where each of its elements is defined as follows [26]:

- $L = \{L_k | k \in \{1, ..., K\}\}$ is a set of layers, where L_1 is the input layer, L_K is the output layer, and layers in between are known as hidden layers. Each layer L_k consists of s_k nodes known as neurons.

- $T \subseteq L \times L$ is a set of connections between layers in such way that except input and output layers, each of hidden layer has an incoming connection and an outgoing connection.
- $\Phi = \{\Phi_k | \in \{2, ..., K\}\}$ is a set of activation functions, one for each non-input layer.

3 Proposed method

Our proposed method for designing DNN topology is presented in Alg. 2. The method is based on the grey wolf optimizer and differential evolution algorithm variants named GWODNN and DEDNN. GWO and DE are used to find the optimal DNN topology - in our case to find the number of hidden layers, number of neurons, dropout probability and activation function for each layer and the optimizer function for DNN. The output layer of each DNN topology is fixed, using Sigmoid activation function and Uniform function for the initialization of kernel.

Algorithm 2 Proposed method

Output: The $best$ DNN topology based on best solution

```
gwo_alg.init_bat();
while termination_condition_not_meet do
    solution = gwo_alg.get_best_solution();
    layer_num = map_batch(solution[75]);
    optimizer = map_optimizer(solution[76]);
    layers = map_layers_triples(solution[0 − (3 ∗ layer_num)]);
    fitness = train_and_eval(layers, optimizer, 100, 32);
    gwo_alg.generate_new_solution(fitness);
end while
best = create_model(gwo_alg.get_best_solution());
```

As shown in Alg. 2[1] used algorithms are producing a solution with the dimension of 77. The dimension of the problem relates to our predefined limitation with regard to the maximum number of hidden layers of DNN. In our case, the maximum number of hidden layers is set to 25 but it could be easily changed to any value. Each of those layers is defined with 3 values: number of layers, number of neurons and dropout probability. Besides definitions of layers, the last two values of a solution are defining the number of hidden layers used to create DNN topology and an optimizer function used to build the DNN model.

Therefore, the individuals in GWODNN and DEDNN are presented as real-valued vectors:

$$\mathbf{x}_i^{(t)} = (x_{i,0}^{(t)}, \ldots, x_{i,n}^{(t)}), \quad \text{for } i = 0, \ldots, Np - 1, \quad (5)$$

where each element of the solution is in the interval $x_{i,1}^{(t)} \in [0, 1]$.

[1]DEDNN is basically the same algorithm, only GWO steps are replaced by DE steps.

The real values of real-valued solution vectors are then mapped according to the equations 6, 7, 8, 9 and 10 where y_1 presents the number of hidden layers, y_2 optimization function, y_3 number of neurons in hidden layer, y_4 dropout probability of hidden layer and y_5 activation function for hidden layer. Each y_2 is member of population $O = \{sgd, rmsprop, adagrad, adadelta, adam, adamax, nadam\}$ which represents a group of available optimization functions, while y_5 is member of population $A = \{softmax, elu, selu, softplus, softsign, relu, tanh, sigmoid, hard\_sigmoid, linear\}$ which represents a group of available activation functions.

$$y_1 = \lfloor x[i] * 20 + 5 \rfloor; y_1 \in [5, 25] \quad (6)$$

$$y_2 = \begin{cases} \lfloor x[i] * 7 + 1 \rfloor; y_2 \in [1, 7] & x[i] < 1 \\ 7 & \text{otherwise,} \end{cases} \quad (7)$$

$$y_3 = \begin{cases} \lfloor x[i] * 100 + 1 \rfloor; y_3 \in [1, 100] & x[i] < 1 \\ 100 & \text{otherwise,} \end{cases} \quad (8)$$

$$y_4 = \lfloor x[i] * 30 + 20 \rfloor; y_4 \in [20, 50] \quad (9)$$

$$y_5 = \begin{cases} \lfloor x[i] * 10 + 1 \rfloor; y_5 \in [1, 10] & x[i] < 1 \\ 10 & \text{otherwise,} \end{cases} \quad (10)$$

The fitness function was defined in two variants: one using the test accuracy calculated on validation set and one using train accuracy calculated on the train set. In equations (11) and (12) are presented the formal definitions of mentioned fitness functions, where $test\_acc$ stands for accuracy calculated on validation set and $train\_acc$ stands for accuracy calculated on train set. Used implementations of GWO and DE are optimized to search for the global minimum, so our fitness functions are defined in a way, which is converting the problem of searching maximal accuracy to the problem of searching minimum as presented in formal definitions.

$$f(test\_acc) = 1 - test\_acc \quad (11)$$

$$f(train\_acc) = 1 - train\_acc \quad (12)$$

4 Experiments and Results

4.1 Datasets

For the purpose of better overall evaluation of our proposed method, we chose 3 datasets with a different number of instances, features, and distribution of classes. In the following chosen datasets are presented in more details.

4.1.1 Phishing Websites Data Set

Phishing Websites Data Set [14] from UCI Machine Learning repository [10], prepared by Mohammad et al. [15] was used for purpose of predicting phishing websites. The basic information of this dataset is presented in Table 1.

Parameter	Value
Number of features	31
Number of instances	11,055
Number of classes	2 classes
Distribution of classes	3,793 phishing websites
	7,262 legitimate websites

Table 1: The basic information about the Phishing Websites Data Set.

4.1.2 Pima Indians Diabetes Database

Pima Indians Diabetes Database [22] from Kaggle, prepared by Smith et al. [23] is dataset used for diagnostically predict whether or not a patient has diabetes. The basic information of this dataset is presented in Table 2.

Parameter	Value
Number of features	8
Number of instances	768
Number of classes	2 classes
Distribution of classes	268 with diabetes
	500 without diabetes

Table 2: The basic information about the Pima Indians Diabetes Database.

4.1.3 Bank Marketing Data Set

Bank Marketing Data Set [16] from UCI Machine Learning repository [10], prepared by Moro et al. [17] is data set related to direct marketing campaigns of a Portuguese banking institution. Dataset is used for predicting if a client will subscribe a term deposit or not. The basic information of this dataset is presented in Table 3.

Parameter	Value
Number of features	17
Number of instances	45211
Number of classes	2 classes
Distribution of classes	5289 subjects subscribed
	40922 subjects not subscribed

Table 3: The basic information about the Bank Marketing Data Set

4.2 Experimental settings

4.2.1 GWO and DE parameters

GWO and DE algorithms were initialized with parameters presented in Table 4. Algorithms were used for searching for an optimal number of hidden layers, a number of neurons, dropout probability and activation function for each hidden layer as well as for searching for optimization function used for building the model.

In our experiments, 20% of the initial dataset was used when calculating fitness using equation (12) presented in Section 3.2. When calculating fitness using equation (11) those 20% of the initial instances are further divided to train and test split in ratio 70:30 where 70% is used for training and calculating the $train\_accuracy$, while 30% is used to calculate the $test\_accuracy$ in previously mentioned equations.

Parameter	GWO	DE
Dimension of the problem	77	77
Population size	40	40
Number of function evaluations	500	500
Lower bound	0.0	0.0
Upper bound	1.0	1.0
F (Scaling factor)		0.5
CR (Crossover probability)		0.9

Table 4: Used parameter values for GWO and DE algorithms.

4.2.2 Deep Neural Network

The base of our proposed method is feed-forward NN with predefined input layer using uniform kernel initializing function and output layer with one neuron also using uniform kernel initializing function and activation function set to sigmoid. The number of the hidden layers, the configuration of each of the hid-

den layers and optimization function used for compiling the model is defined with the solution found using GWODNN and DEDNN.

4.2.3 Learning parameters

Beside the DNN topology and its layer parameters which were obtained from the solution of our GWODNN / DEDNN the other learning parameters (e.g. batch size, epochs) were picked based on our previous experience in machine learning. The training of DNN was performed with a batch size of 32 and 100 epochs. We used the same learning parameters for each of the used experimental datasets.

4.3 Results

The proposed method and experiments were implemented in Python programming language using the following frameworks and/or libraries: NiaPy [29], Keras [2] with Tensorflow [1], NumPy, Pandas and scikit-learn [18]. For all of the existing classification algorithms the implementations from scikit-learn with their default settings were used.

In order to objectively evaluate the performance of our proposed method, we followed an established methodology. All of the used datasets were initially divided into two subsets in a ratio 80:20. The smaller subset was used for finding the best DNN topology using the proposed metaheuristic methods. After obtaining the optimal DNN topology, only the remaining larger subset was used to perform the ten-fold cross-validation procedure. Namely, in order to keep the evaluation procedure as fair as possible, the 20% of the original instances, which have been already used to search for DNN topology, were not included in the final results. In the case of conventional classifiers, the standard ten-fold cross-validation approach was used on the original datasets.

Results presented in following sections are minimum, maximum, average and median of the ten-folds achieved by a specific method on test sets. With the labels *GWODNN/DEDNN - train* we refer to our GWODNN/DEDNN methods with fitness calculated based on train accuracy, while with the labels *GWODNN/DEDNN - test* we refer to our GWODNN/DEDNN methods with fitness calculated based on test accuracy. Performance of our proposed method variations were compared against multilayer perceptron classifier (MLP), with 5 hidden layers each with 100 neurons, k-nearest neighbor (KNN) with number of neighbors set to 5 classifier and decision tree classifier (DT) with minimum number of samples required to split an internal node set to 2. Beside the mentioned conventional classifiers, we also compared the performance of our proposed method variants against baseline DNN with an input layer, 2 hidden layers and an output layer. The first hidden layer is having 12 neurons while the second is having 8 neurons, both of them are utilizing ReLU activation function. On the output layer, the Sigmoid activation function is used. The same baseline DNN is used over all three datasets, which provides us with the solid baseline performance for further performance comparison against our GWODNN/DEDNN variants.

4.3.1 Performance on Phishing Websites Data Set

The results of the performance of the classifiers on the Phishing Websites Data Set is presented in Figure 1. The best performing classifier with the average accuracy of 96.9% is MLP, closely followed by the DEDNN - train and DT with the average accuracy of 96.6%. In general, all classifiers performed comparable, except the Base DNN and the KNN classifier which are lacking behind 2.8% and 2.4%. Looking closer at the performance of our proposed method variations, we can see that none of them is performing noticeably better than the others, with all the accuracies in the range of 0.6%.

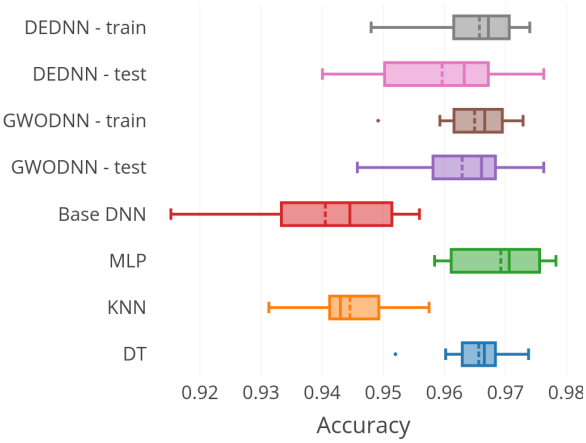


Figure 1: Comparison of accuracy between our proposed methods and other conventional classifiers using 10-fold cross validation on Phishing Websites Data Set.

4.3.2 Performance on Pima Indians Diabetes Database

The performance of classifiers on the Pima Indians Diabetes Database is presented in Figure 2. With the average accuracy of 72.6%, GWODNN - test proves to be the best performing classifier by the noticeable margin of 2.1% against the second best the KNN classifier. The worst performance achieved the DT classifier with 67.1% of average accuracy. Focusing on our methods, we can see, that the GWODNN variations did outperform the DEDNN variations by a noticeable margin of 3.9% on the case of the *test* variations and by 1.0% on the case of *train* variations. The best performing classifier on the previous dataset - MLP classifier is lagging behind the best performing classifier significantly by the margin of 3.8%.

Table 5: Average accuracy performance for all classifiers against all of the datasets

	Phishing Websites	Pima Indians Diabetes	Bank Marketing
DEDNN - train	0.966	0.692	0.895
DEDNN - test	0.960	0.687	0.889
GWODNN - train	0.965	0.702	0.889
GWODNN - test	0.963	**0.726**	**0.896**
Base DNN	0.941	0.692	0.895
MLP	**0.969**	0.688	0.889
KNN	0.945	0.705	0.882
DT	0.966	0.671	0.874

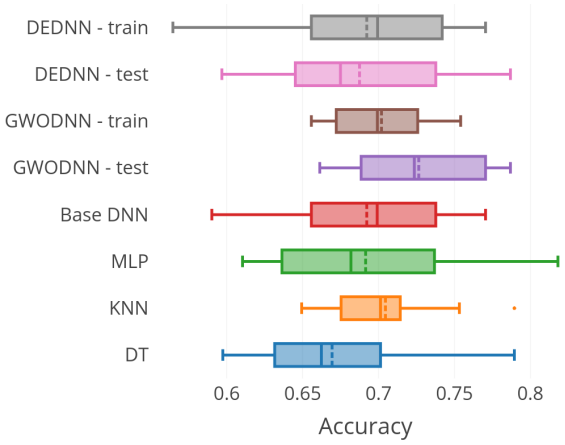


Figure 2: Comparison of accuracy between our proposed methods and other conventional classifiers using 10-fold cross validation on Pima Indians Diabetes Database

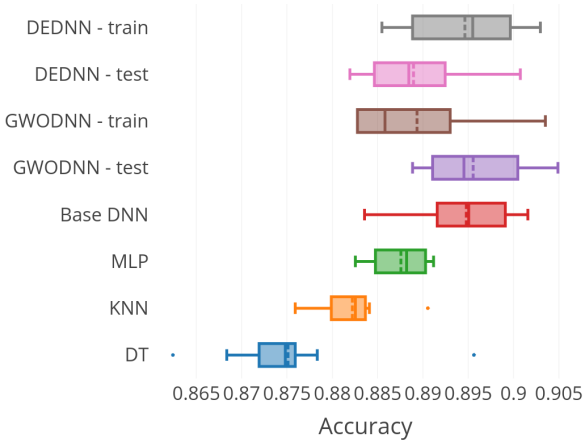


Figure 3: Comparison of accuracy between our proposed methods and other conventional classifiers using 10-fold cross validation on Bank Marketing Data Set

4.3.3 Performance on Bank Marketing Data Set

The experimental results of the performance of the classifiers on the Bank Marketing Data Set is presented in Figure 3. Looking at the boxplots, we can see that all of tested classifiers, except the DT, performed somewhat similar. The best performing classifier is GWODNN - test with the 89.6% of average accuracy. The second best classifier, leaving out the results of our method variations, is Base DNN with 89.5% lacking behind the best performing classifier by 0.1%. Comparing the performance of our method variations, we can see that results are ranging from 0.1% to 0.7%.

4.3.4 Performance comparison across all datasets

Shown in Table 5 are results of all average accuracy performances for all of the tested classifiers and on all datasets. In general, the GWODNN - test classifier performs the best in two out of three datasets, achieving the best performance on the Pima Diabetes Database and Bank Marketing Data Set, while not much (0.6%) lagging behind the MLP on the Phishing Websites Data Set. Generally speaking, all of our proposed method variants, were able to achieve high classification performance with no major performance drops throughout all tested datasets, unlike the conventional classifiers, which in some cases are lagging behind by a noticeable margin. The same performance drops throughout all tested datasets, can also be observed when comparing the performance of base DNN against out GWODNN/DEDNN variants performance.

5 Conclusions

In this paper, we presented a new approach utilizing a population-based metaheuristics algorithms to design a DNN topology in a straightforward automatic manner. The results, obtained from conducted experiments, have proven our method to be very promising, giving us high classification performance on all of the tested datasets, with no major performance drops, in comparison with conventional classifiers.

In the future, based on those encouraging results, we would like to expand our work with the use of different algorithms (e.g. Cuckoo search algorithm and Particle swarm optimization), as well as tackle the problem of time complexity using such population-based metaheuristics approach for designing a DNN topology, with parallelization methods and techniques.

Acknowledgments

The authors acknowledge the financial support from the Slovenian Research Agency (Research Core Funding No. P2-0057).

References

[1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[2] François Chollet et al. Keras. `https://keras.io`, 2015.

[3] Swagatam Das and Ponnuthurai Nagaratnam Suganthan. Differential evolution: A survey of the state-of-the-art. *IEEE transactions on evolutionary computation*, 15(1):4–31, 2011.

[4] Iztok Fister Jr, Xin-She Yang, Iztok Fister, Janez Brest, and Dušan Fister. A brief review of nature-inspired algorithms for optimization. *Elektrotehniski vestnik*, 80(3):116–122, 2013.

[5] Yoav Goldberg. A primer on neural network models for natural language processing. *Journal of Artificial Intelligence Research*, 57:345–420, 2016.

[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. In *European Conference on Computer Vision*, pages 630–645. Springer, 2016.

[7] Geoffrey Hinton, Li Deng, Dong Yu, George E Dahl, Abdel-rahman Mohamed, Navdeep Jaitly, Andrew Senior, Vincent Vanhoucke, Patrick Nguyen, Tara N Sainath, et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 29(6):82–97, 2012.

[8] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[9] Frank Hung-Fat Leung, Hak-Keung Lam, Sai-Ho Ling, and Peter Kwong-Shun Tam. Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *IEEE Transactions on Neural networks*, 14(1):79–88, 2003.

[10] M. Lichman. UCI Machine Learning Repository, 2013. Available at http://archive.ics.uci.edu/ml, Accessed: 2018-01-15.

[11] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.

[12] Risto Miikkulainen, Jason Liang, Elliot Meyerson, Aditya Rawal, Dan Fink, Olivier Francon, Bala Raju, Hormoz Shahrzad, Arshak Navruzyan, Nigel Duffy, et al. Evolving deep neural networks. *arXiv preprint arXiv:1703.00548*, 2017.

[13] Seyedali Mirjalili, Seyed Mohammad Mirjalili, and Andrew Lewis. Grey wolf optimizer. *Advances in engineering software*, 69:46–61, 2014.

[14] Rami M Mohammad, Fadi Thabtah, and Lee McCluskey. Phishing Websites at UCI Machine Learning Repository. Available at http://archive.ics.uci.edu/ml/datasets/Phishing+Websites, Accessed: 2018-05-17.

[15] Rami M Mohammad, Fadi Thabtah, and Lee McCluskey. An assessment of features related to phishing websites using an automated technique. In *Internet Technology And Secured Transactions, 2012 International Conference for*, pages 492–497. IEEE, 2012.

[16] Sérgio Moro, Paulo Cortez, and Paulo Rita. Bank Marketing Data Set. Available at https://archive.ics.uci.edu/ml/datasets/Bank+Marketing, Accessed: 2018-05-17.

[17] Sérgio Moro, Paulo Cortez, and Paulo Rita. A data-driven approach to predict the success of bank telemarketing. *Decision Support Systems*, 62:22–31, 2014.

[18] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[19] Tim Salimans, Jonathan Ho, Xi Chen, Szymon Sidor, and Ilya Sutskever. Evolution strategies as a scalable alternative to reinforcement learning. *arXiv preprint arXiv:1703.03864*, 2017.

[20] Jürgen Schmidhuber. Deep learning in neural networks: An overview. *CoRR*, abs/1404.7828, 2014.

[21] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[22] Jack W Smith, JE Everhart, WC Dickson, WC Knowler, and RS Johannes. Pima Indians Diabetes Database. Available at https://www.kaggle.com/uciml/pima-indians-diabetes-database, Accessed: 2018-05-17.

[23] Jack W Smith, JE Everhart, WC Dickson, WC Knowler, and RS Johannes. Using the adap learning algorithm to forecast the onset of diabetes mellitus. In *Proceedings of the Annual Symposium on Computer Application in Medical Care*, page 261. American Medical Informatics Association, 1988.

[24] Kenneth O Stanley, David B D'Ambrosio, and Jason Gauci. A hypercube-based encoding for evolving large-scale neural networks. *Artificial life*, 15(2):185–212, 2009.

[25] Rainer Storn and Kenneth Price. Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4):341–359, 1997.

[26] Youcheng Sun, Xiaowei Huang, and Daniel Kroening. Testing deep neural networks. *CoRR*, abs/1803.04792, 2018.

[27] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, et al. Going deeper with convolutions. Cvpr, 2015.

[28] Petra Vidnerová and Roman Neruda. Evolution strategies for deep neural network models design.

[29] Grega Vrbančič, Lucija Brezočnik, Uroš Mlakar, Dušan Fister, and Iztok Fister Jr. NiaPy: Python microframework for building nature-inspired algorithms. *Journal of Open Source Software*, 3, 2018.

[30] Grega Vrbančič, Iztok Fister Jr., and Vili Podgorelec. Swarm intelligence approaches for parameter setting of deep learning neural network: Case study on phishing websites classification. In *WIMS '18: Proceedings of the 8th International Conference on Web Intelligence, Mining and Semantics*. ACM, accepted for publishing.

[31] Xiaowei Xu, Yukun Ding, Sharon Xiaobo Hu, Michael Niemier, Jason Cong, Yu Hu, and Yiyu Shi. Scaling for edge inference of deep neural networks. *Nature Electronics*, 1(4):216, 2018.

SportyDataGen: An Online Generator of Endurance Sports Activity Collections

Iztok Fister Jr., Grega Vrbančič, Lucija Brezočnik, Vili Podgorelec and Iztok Fister
University of Maribor
Faculty of Electrical Engineering and Computer Science
Koroška cesta 46, 2000 Maribor, Slovenia
{iztok.fister1, grega.vrbancic, lucija.brezocnik, vili.podgorelec, iztok.fister}@um.si

Abstract. *Analyzing sport data becomes, every year, more interesting for a wide spectrum of researchers in the sports domain. Recently, more and more data relating to sports have become available to researchers due to the huge progress of information technologies. New wearable devices enable athletes to track performance data that are saved into sport activity datasets and later analyzed by sport trainers using data mining methods. These datasets are also very useful to researchers in the sports domain who usually employ a collection of them in their researches. Typically, these researchers are confronted with two problems: Firstly, how to gain the real data, and secondly, how to process them. In this paper, we propose a new tool for generating a collection of sport activity datasets online (named SportyDataGen), where data from various endurance sports disciplines, obtained from different mobile devices worn by amateur, as well as professional athletes are stored. The data are generated in CSV format according to user demands, and do not require any preprocessing. Here, the proposed tool is described systematically, while one example is also presented of a generated collection.*

Keywords. sport data analysis, data mining, triathlon, online generator

1 Introduction

Analyzing sports activity datasets has become an unavoidable tool for monitoring advances of athletes involved in the process of sport training. This statement is supported by the rising power of Information Technology (IT) on the one hand, and development of more sophisticated algorithms in the domain of Artificial Intelligence (AI) on the other. Simultaneously, a lot of people have been involving into sports, into which they are forced due to their, usually sedentary, modern life-style. Mainly, the endurance sports disciplines are preferred by the athletes, because those are individual sports and do not need any sports infrastructure (e.g., sports hall, special sports tools, etc.), where athletes train, when they have time, and therefore, there is no need for special coordination between people as in the case of a team sports disciplines. Nowadays, these athletes use smart watches or smart phones during the training session in order to analyze tracked data about realized their sport activities after training. Typically, these devices are the origin of producing the huge amount of data. Indeed, data in this domain have been rising exponentially.

Swimming, biking, and running are typical endurance sports disciplines that are more eligible for amateur athletes. However, combining these three sports caused the formation of the more complex sports disciplines, such as, for example, triathlon. The word **triathlon** refers to the modern sport that is one of the more attractive sports disciplines recently, where the final time of the competition is the sum of all three sports disciplines, together with two transitions. Thus, the first transition is devoted to preparing an athlete from swimming to cycling, and the second from cycling to running. It is worth mentioning that there are various triathlon distances that spread from the short distance toward the medium and long distances, until the ultra-triathlons.

Although the triathlon is a very young sport, every year it attracts more and more competitors from all over the world. This trend is already evidenced by the following facts: (1) The majority of international races (e.g. IRONMAN triathlons) are filled with competitors from the whole world shortly after the Internet registrations are opened, (2) Many personal trainers have become more busy due to providing the process of sports training for an increased number of triathlon enthusiasts during the whole year, and (3) Triathlon has also become an interesting subject for intensive scientific researches [11, 16, 13, 10].

In the past years, most of the researches related to performance studies, tactics, and health in cycling, running, and triathlon. These sports also became interesting for researchers in Data Science and other AI domains by rising the volume of data produced during training sessions. For example, Matabuena et al. [14] showed that the performance of a group of athletes can be predicted without forcing them to fatigue by exer-

cises of low intensity. Not long ago, Mnadla et al. [15] outlined data concerning the related activities and interest for the Ironman competition on the Internet, while Knechtle et al. [12] analyzed participation and performance trends in ultra-triathlons. Moreover, analyzing cycling data [2, 17] also produced valuable research for a wide audience. On the other hand, Cintia et al. [3] shows how to exploit the massive data exported from Strava to construct personalized training programmes for cyclists.

Interestingly, public access to endurance sports data for research is still a very big bottleneck for researchers in this research domain, because a lot of data remain unreachable due to the policy of athletes, trainers, or even sport clubs. For that reason, studies cannot be replicated, while some studies cannot even be conducted because of the lack of data. In the past, we have tried to supply data produced by wearable devices (smart phones and smart watches) to researchers in some individual sports disciplines [18, 8]. In line with this, we were focused on collecting data obtained from real athletes. However, this task was not so easy, because some athletes did not want to share their activities publicly. Therefore, we have assembled sports activity data obtained from a group of athletes who were willing to share their data. Every year, we published a big collection of these data online. Fascinatingly, we received a very positive feedback over the years.

Unfortunately, these datasets were collected in their native format (i.e., TCX and GPX file formats), and, therefore, preprocessing (i.e., parsing) is needed before using them. Obviously, this task is far to be simple, especially for researchers without deep knowledge of computer programming. In order to help them, we decided to develop an online sports activity generator of sports activity collection, where users would be able to select what and how much of the generated data they want. This generator, named **SportyDataGen**, can be found at the Internet page `www.sport-slo.net`.

The aim of the paper is to present the features of the online generator of an endurance sports activity collection that are available to a potential user on the Internet. The structure of the remainder of the paper is as follows. Section 2 discusses problems arising in using existing collections of an endurance sports activity datasets. Generation of the endurance sports activity collection with SportyDataGen is described in Section 3. Examples of the generated endurance sports activity datasets are presented in Section 4. The paper is concluded with Section 5, which summarizes the performed work and outlines directions for the future.

2 Problems using existent collections of endurance sports activity datasets

SportyDataGen relies on real data that were obtained from mobile devices worn by amateur and professional athletes during their endurance sports training sessions. Most of the data were exported directly from Internet applications, like Garmin Connect [9], Strava Connecting the world's athletes [20], or Suunto MovesCount [21]. Athletes gave us their credentials in order to access their profiles on such applications. At the same time, they agreed that data can be used in this research. The exported datasets can either be in TCX or GPX file format. Actually, both formats basically have their roots in eXtensible Markup Language (XML).

Although these activity datasets obtained from real athletes were collected and made available publicly for the purpose of research, they were rarely used in practice, especially because they must be preprocessed before their use. However, the preprocessing is not a simple task, due to the fact that there is no sophisticated parser of these file formats available in particular programming languages. As a result, the parser needs to be developed from scratch, but this demands a sufficiently high level of knowledge in computer programming. Unfortunately, a lot of users suffer from a lack of this knowledge and, therefore, are unable to carry out such preprocessing.

Obviously, the authors of these collections are aware of these problems. Consequently, some previous versions of the collections allow downloading the activity datasets in a raw data format suitable for exploring without preprocessing [8]. Although these collections are very useful, they are too static, and do not consider any variability (e.g., using similar activity datasets from two different athletes). Indeed, this situation calls for using a generation tool for generation of the endurance sports activity collections that are suitable for generating data on user demand.

3 Generation of endurance sports activity collection

The SportyDataGen is a generator for endurance sports activity collection consisting of more components working together sequentially. Actually, the architecture of the generator that is presented in Fig. 1 is very complex. There are the following six components of the generator:

- Collection: This component enables automatic downloading of the endurance sports activities from online profiles to be at the disposal by producers of specific wearable devices, like Garmin, Strava, Suunto, etc.

- Processing: Downloaded raw endurance sports activities saved in TCX or GPX file formats are parsed by the mentioned components. This means that the most important features are extracted from raw files. Thus, features, like the average Heart Rate (HR), the total sports training Time Duration (TD), the total distance and consumed calories, characterize a typical endurance sports training session.
- Database: The purpose of the component is to store extracted features into a database for further analysis. Typically, each training session in the database is identified by the training load indicator - TRaining IMPulse (TRIMP), proposed by Banister [1]. This load indicator is expressed simply as:

$$\text{TRIMP} = TD \cdot HR, \tag{1}$$

  where TD denotes the time duration of the endurance sports training session in minutes, and HR is an average Heart Rate in beats per minutes.
- Generation: The corresponding endurance sports activities are selected from a database according to the user's preferences, like the sports discipline, the number of generated endurance sports activities in the collection, and the features that should be generated. In line with this, these activities can also be clustered regarding the TRIMP measure.
- Output: The component is devoted for exporting the selected collection of measured sports activities in CSV format. Additionally, the table with the same features are also made available to the user.
- Web application: It allows control of the database and the generation process of the measured sports activity collection. On the other hand, the application enables users to interact with a User Application Interface (API) for tailoring the collection according to their needs.

The first three components are already part of the Artificial Sports Trainer (AST) proposed by Fister et al. [7], dedicated to acquire the sports training activities into a database. Indeed, the database serves as a basis for decision-making about realization of the sports training process by the AST. In line with this, the AST uses it for planning the sports training sessions [5], adapting the training sessions [6], proposing the dietary plan necessary for covering energy consumption as demanded by the specific training plan [4].

In this study, the database was exploited for generation of endurance sports activity collection, where the prescribed number of endurance sports activities are selected according to the user's preferences in a CSV file format that does not need any preprocessing and is, therefore, also available for use by users with less knowledge of computer programming. The following preferences are available to users by generating the collection of sports activities:

- the endurance sports discipline
- the number of generated activities
- features
- the number of clusters
- intensity of the generated endurance activity collection in TRIMP
- the generation mode

At this moment, two endurance sports disciplines are supported, i.e., cycling and running. However, more endurance sports disciplines could be supported in the future (e.g., triathlon). The number of generated endurance sports activities determines the endurance sports activities collection. Users can select between the following features: heart rate, duration, distance, calories, maximum and average altitude, and pace. The number of clusters prescribes those numbers of different clusters used by the k-means clustering algorithm, from which the activities could be selected. The intensity of the generated activities is optional, and actually refers to the generation mode. If the generation mode is purely random, then the value of zero is expected, which means that the intensity of all training activities in the collection is ignored. In contrast, when the evolutionary approach is selected, a minimum TRIMP intensity value I_{C_i} of the appropriate participating cluster is specified by the evolutionary optimization process.

As already mentioned, the SportyDataGen supports two modes of generating the in the remained of the paper sports activity collection, i.e., purely random, and evolutionary. In the remained of the paper, both modes are described in detail.

3.1 Generation modes

Two modes of generating the endurance sports activity collection exist. Actually, both modes start with clustering the endurance sports activities into a definite number of clusters. The clusters are determined according to the TRIMP measure, while the k-means clustering algorithm is applied for clustering. However, the number of clusters is crucial for the performance of the generation. The more clusters are selected, the more diverse activities can be generated. At the moment, the SportDataGen supports from 3 to 10 numbers of clusters. Typically, the proper number of clusters must be determined experimentally.

The set of endurance training sessions are clustered according to the TRIMP training load indicator in order to obtain groups of training sessions of similar intensities. As can be seen from Eq. (1), time duration and average heart rate have influence on the TRIMP training load indicator. However, the main disadvantage of this indicator is that it is insensitive to the different levels of training. Users of SportyDataGen can choose the

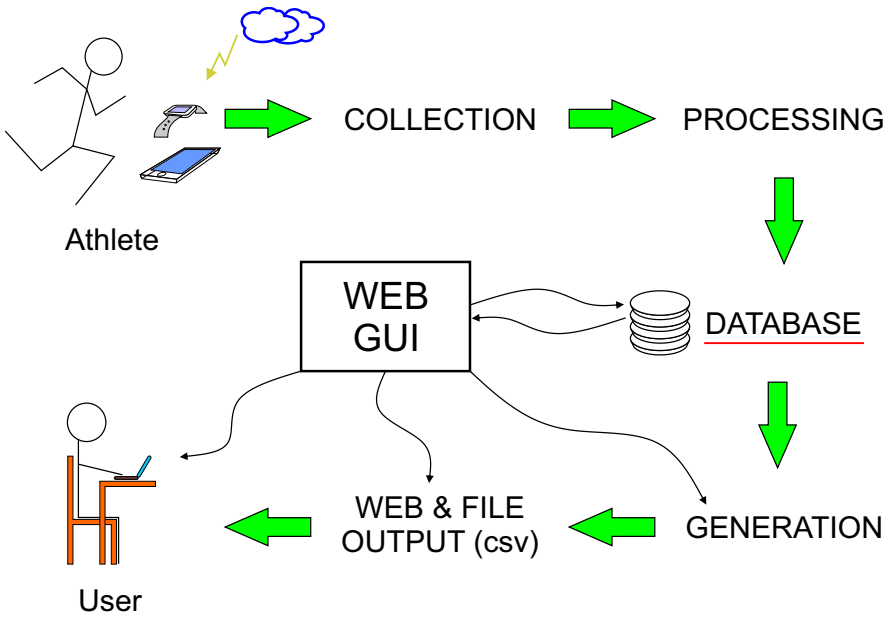


Figure 1: Architecture of SportyDataGen.

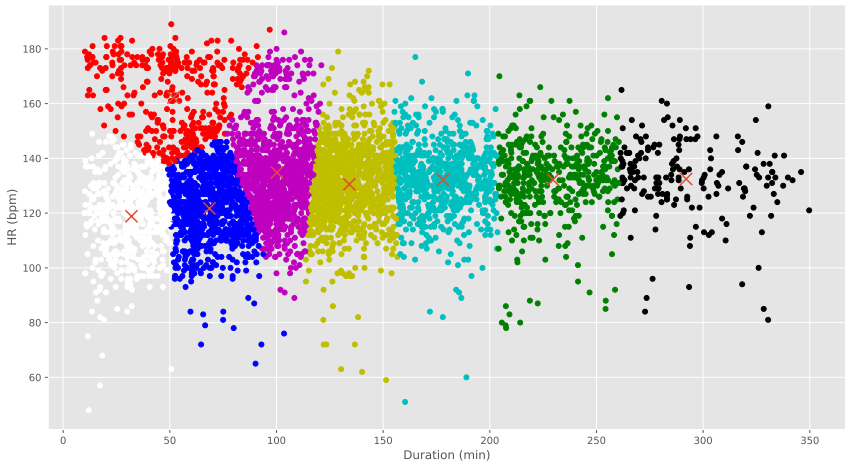


Figure 2: Example of clustering (C = 8).

desired cluster settings. Fig 2 presents an example of clustering, where the number of clusters is 8.

Generation of the endurance sports activity collection can be defined mathematically as follows. Let us assume a collection of endurance sports activities as a real-valued vector:

$$\mathbf{x}_i = [x_{i,1}, \ldots, x_{i,D}]^T, \quad \text{for } i = 1, \ldots, NP, \quad (2)$$

and a set of clusters are given:

$$C = \{C_1, C_2, ..., C_n\}, \quad (3)$$

where NP denotes the number of different collections, D is the number of endurance activities in the collection, n the number of clusters. Each cluster is defined as:

$$C_k = \{t_{k,1}, \ldots, t_{k,m_k}\}, \quad \text{for } k = 1, \ldots, n, \quad (4)$$

where $t_{k,l}$ denotes a specific endurance training session and m_k is the size of the k-th cluster. A vector denoting a sequence of clusters

$$\mathbf{c}_i = [c_{i,1}, \ldots, c_{i,D}]^T, \quad \text{for } i = 1, \ldots, NP, \quad (5)$$

is assigned to each of the vector $\mathbf{x}_i$ that determines a specific cluster from which the endurance sports activity can be selected. The sequence of clusters is then generated according to the following assumptions. Until the vector $\mathbf{c}_i$ is not full, a tournament is played between the specific cluster j and its corresponding threshold value is expressed as follows:

$$threshold(C_i, C_j) = \left| I_{C_i} - I_{C_j} \right| / 10. \quad (6)$$

If the value drawn from the uniform random distribution in interval $[0, 1]$ is less than the threshold value, the j-th cluster is placed on the observed position in the vector $\mathbf{c}_i$. The intensity cluster entered via user's API $I_{C_i} = 0$ favors those clusters C_j that gave the closest distances $|I_{C_i} - I_{C_j}|$ according to the following ordering:

$$I_{C_0} \succeq \ldots \succeq I_{C_{i-1}} \succeq I_{C_i} \succeq I_{C_{i+1}} \succeq \ldots \succeq I_{C_n}, \quad (7)$$

where $' \succeq'$ denotes the relation of 'is better than or equal to', and $I_{C_i} = 0$. Actually, a distance $|I_{C_i} - I_{C_{i-1}}| = 1$, or $|I_{C_i} - I_{C_{i+1}}| = 1$, a distance $|I_{C_i} - I_{C_{i-2}}| = 2$, or $|I_{C_i} - I_{C_{i+2}}| = 2$, while a distance $|I_{C_i} - I_{C_0}| = i - 1$, and $|I_{C_i} - I_{C_n}| = n - i - 1$.

Then, the corresponding sports activity $t_{c_{i,j},l}$ is determined, where l is calculated according to the following equation:

$$l = \lfloor x_{i,j} \cdot m_{c_{i,j}} \rfloor \cdot m_{c_{i,j}}, \quad (8)$$

If the purely random selection mode is drawn, each element $x_{i,j}$ is selected simply according to the following equation:

$$x_{i,j} = \text{rand}(0, 1), \quad (9)$$

where $\text{rand}(0, 1)$ generates the random number drawn from the interval $[0, 1]$.

The evolutionary approach is intended for producing more reliable and robust data. It is powered by Differential Evolution (DE) [19] that is considered as an Evolutionary Algorithm (EA). Due to the nature of EAs, DE is population-based, and consists of Np real-coded vectors representing the candidate solutions regarding Eq. 2. Each element of the solution is in the interval $x_{i,1}^{(t)} \in [x_i^{(L)}, x_i^{(U)}]$, where $x_i^{(L)}$ and $x_i^{(U)}$ denotes the lower and upper bounds of the i-th variable, respectively.

DE guided search is controlled by three operators:

- mutation,
- crossover and
- selection.

Mutation selects two solutions randomly, and adds the scaled difference between these to the third solution. This mutation is expressed as follows:

$$\mathbf{u}_i^{(t)} = \mathbf{x}_{r1}^{(t)} + F \cdot (\mathbf{x}_{r2}^{(t)} - \mathbf{x}_{r3}^{(t)}), \text{ for } i = 1, \ldots, Np, \quad (10)$$

where F denotes the scaling factor as a positive real number that scales the rate of modification, while $r1,\ r2,\ r3$ are randomly selected values in the interval $1 \ldots Np$. Note that, typically, the interval $F \in [0.1, 1.0]$ is used in the DE community.

Uniform crossover is employed as a crossover by the DE, where the trial vector is built from parameter values copied from two different solutions. It is expressed as follows:

$$w_{i,j}^{(t+1)} = \begin{cases} u_{i,j}^{(t)} & \text{rand}_j(0,1) \leq CR \vee j = j_{rand}, \\ x_{i,j}^{(t)} & \text{otherwise}, \end{cases} \quad (11)$$

where $CR \in [0.0, 1.0]$ controls the fraction of parameters that are copied to the trial solution. Note, the relation $j = j_{rand}$ ensures that the trial vector is different from the original solution $\mathbf{x}_i^{(t)}$.

Selection is, in fact, a generalized one-to-one selection that is expressed mathematically as follows:

$$\mathbf{x}_i^{(t+1)} = \begin{cases} \mathbf{w}_i^{(t)} & \text{if } f(\mathbf{w}_i^{(t)}) \leq f(\mathbf{x}_i^{(t)}), \\ \mathbf{x}_i^{(t)} & \text{otherwise}. \end{cases} \quad (12)$$

In fact, each element $t_{c_i,l}$ denotes the l-th training session in the c_i-th cluster. As a result, the fitness function of the DE algorithm is expressed as follows:

$$f(\mathbf{x}_i) = \sum_{j=1}^{D} \text{TRIMP}_{t_{c_i,j}}, \quad (13)$$

where $\text{TRIMP}_{t_{c_i,j}}$ denotes the TRIMP training load indicator of the corresponding training session $t_{c_i,j}$ calculated by Eq. (13).

Let us mention that parameter settings for evolutionary approach are: $NP = 20$, N_{iter} [1] = 200, F = 0.5, CR = 0.9.

[1] number of iterations

3.2 Advances and weaknesses of SportyDataGen

SportyDataGen is easy to use. The user may only select the required parameters and sports data is generated. To use SportyDataGen, no annoying Register/Login features are needed, and no irritating advertisements are present. It generates real data obtained from different athletes. Thus, amateur, as well as professional, athletes are taken into account. SportyDataGen relies totally on references and, therefore, does not offer randomly generated data. Generated datasets are downloaded from a server. This means that they are easy to share. Each of them obtains a unique dataset link for the purposes of citing in a publication or sharing to your friends. On the other hand, they are also easy to reuse and for replication. In any case, reviewers may be equipped by a full dataset and convince themselves by reproducing your research. The generator supports a multiple format output. Striving to help researchers with time consuming tasks is taken into account. Sports data can, therefore, be downloaded in multiple formats, including CSV. More formats will be supported in the future. Generated datasets are stored on a server permanently, and daily backups are conducted. Data are added into our database regularly. To have access for sport volunteers for donating sports data is in accordance with the mission of SportyDataGen. The current version of SportyDataGen supports biking and running sports disciplines. However, SportyDataGen could support more sports disciplines in the future, including swimming and multi-sports (e.g., triathlon, duathlon). In other words, the current version of SportyDataGen is not finished yet. Indeed, new ideas will be implemented in the coming versions of SportyDataGen.

On the contrary, some weaknesses of the proposed method also exist. Due to the nature of our volunteers, some clusters are not supported by enough data. For example, according to the clustering in 10 clusters, the cluster with high intensities is not supported by enough activities. The main problems are that athletes do not perform extremely long training sessions with high average heart rate. One of the weaknesses that may also be considered is that raw data of a particular endurance activity is not presented online. Raw endurance activity consists of all GPS track points that some researchers can use for deeper analysis. However, this part is considered for future releases of SportyDataGen. Additionally, the current version is not concentrated on deeper validation of measures produced by wearable devices. In fact, some athletes do not use the newest wearable devices. For that reason, some measures may not be totally consistent due to the problems with devices. Anyway, we consider all original measures that are available in data. In the future, we intend to validate each activity deeply before saving it to the database.

4 Examples of generated endurance sports activity collection

Table 1 presents an example of a generated endurance sports activity collection in cycling. In this collection, the first row presents the ID of an activity and all the features that were selected by the user. It is worth mentioning that the value of duration is presented in minutes, while distance is presented in kilometers. Average HR is presented as beats per minute, altitude measures are presented in meters, while calories are based on kCal. Here, endurance activities are very similar, due to the selected parameter of intensity. These data can now be imported easily as test data for many tasks, e.g. planning sport training sessions with AI tools.

5 Conclusion

The objective of this paper was to present a very complex generator of endurance sport activity collection for research that is accessible online. In the past years, analyzing sport activities that were created by sport trackers has become a very interesting research area. However, data for analysis still represent a big bottleneck for many researchers. Although many datasets have been released in the past years, they were actually collections of raw activity data downloaded from various applications that producers of various wearable tracking devices offer their customers (i.e., athletes).

In line with this, raw data limited the use of such data due to the sophisticated processing. Some researchers are non-programmers and therefore can hardly deal with programming tasks. In this case, the proposed tool SportyDataGen is unavoidable. On the other hand, it really speeds up the preprocessing process, while it's web application is easy to use. Moreover, its architecture allows users to share data easily. Users can simply select parameters and thus they are capable of generating the various outputs.

In the future, we intend to improve this generator according to the user's feedback. More generation modes are going to be supported in a future release, along with more endurance sports disciplines. The next big challenge is to support generation of triathlon datasets.

Acknowledgments

The authors acknowledge the financial support from the Slovenian Research Agency (Research Core Funding No. P2-0057).

References

[1] EW Banister. Modeling elite athletic performance. *Physiological testing of elite athletes*, pages 403–424, 1991.

ID	Duration	Distance	Average HR	Average alt.	Max alt.	Calories	Ascent	Descent
1	153.72	61.45	152.00	214.83	267.00	1401.00	785.80	785.80
2	147.07	28.23	138.00	808.41	938.60	782.00	1294.40	1293.80
3	160.42	66.54	135.00	207.46	285.20	1109.00	690.80	695.40
4	164.38	22.92	132.00	207.09	274.80	1230.00	626.40	617.60
5	169.02	61.38	141.00	131.24	203.80	1558.00	228.80	226.80
6	144.60	41.06	134.00	251.24	334.60	1115.00	919.20	933.00
7	146.43	40.03	141.00	29.21	105.60	780.00	803.80	801.20
8	155.92	46.42	130.00	481.50	687.00	1520.00	1051.60	1059.20
9	143.75	33.02	144.00	795.67	919.80	766.00	1262.40	1262.80
10	148.17	42.22	119.00	224.80	288.80	851.00	627.40	636.60
11	146.35	35.55	130.00	304.37	535.60	1133.00	709.00	720.80
12	146.65	70.07	140.00	378.05	447.60	980.00	549.20	518.80
13	165.53	78.82	128.00	281.12	378.00	1069.00	985.00	988.00
14	173.92	19.50	115.00	275.66	432.40	765.00	572.40	650.40
15	151.07	67.79	120.00	406.36	628.60	1845.00	866.80	870.00
Avg.	154.47	47.67	133.27	333.13	448.49	1126.93	798.20	804.01

Table 1: Example of generated dataset using pure random approach

[2] Hana Charvátová, Aleš Procházka, Saeed Vaseghi, Oldřich Vyšata, and Martin Vališ. Gps-based analysis of physical activities using positioning and heart rate cycling data. *Signal, Image and Video Processing*, 11(2):251–258, 2017.

[3] Paolo Cintia, Luca Pappalardo, and Dino Pedreschi. " engine matters": A first large scale data driven study on cyclists' performance. In *Data Mining Workshops (ICDMW), 2013 IEEE 13th International Conference on*, pages 147–153. IEEE, 2013.

[4] Dušan Fister, Samo Rauter, Iztok Fister, and Iztok Fister Jr. Generating eating plans for athletes using the particle swarm optimization. In *17th International Symposium on Computational Intelligence and Informatics (CINTI)*, pages 193–198, 2016.

[5] Iztok Fister, Samo Rauter, Xin-She Yang, Karin Ljubič, and Iztok Fister Jr. Planning the sports training sessions with the bat algorithm. *Neurocomputing*, 149:993–1002, 2015.

[6] Iztok Fister Jr. and Iztok Fister. *Generating the Training Plans Based on Existing Sports Activities Using Swarm Intelligence*, pages 79–94. Springer International Publishing, Cham, 2017.

[7] Iztok Fister Jr., Karin Ljubič, Ponnuthurai Nagaratnam Suganthan, Matjaž Perc, and Iztok Fister. Computational intelligence in sports: Challenges and opportunities within a new research domain. *Applied Mathematics and Computation*, 262:178–186, 2015.

[8] Iztok Fister Jr., Samo Rauter, Dušan Fister, and Iztok Fister. A collection of sport activity datasets for data analysis and data mining 2017a.

[9] Inc. Garmin. Connect. https://connect.garmin.com/en-US/, 1996–2017. [Accessed on: December 2017].

[10] Mark J Horne. The relationship of race discipline with overall performance in sprint and standard distance triathlon age-group world championships. *International Journal of Sports Science & Coaching*, page 1747954117738878, 2017.

[11] Nicholas E Kimber, Jenny J Ross, Sue L Mason, and Dale B Speedy. Energy balance during an ironman triathlon in male and female triathletes. *International journal of sport nutrition and exercise metabolism*, 12(1):47–62, 2002.

[12] Beat Knechtle, Patrizia Knechtle, and Romuald Lepers. Participation and performance trends in ultra-triathlons from 1985 to 2009. *Scandinavian journal of medicine & science in sports*, 21(6), 2011.

[13] Romuald Lepers, Christoph A Rüst, Paul J Stapley, and Beat Knechtle. Relative improvements in endurance performance with age: evidence from 25 years of hawaii ironman racing. *Age*, 35(3):953–962, 2013.

[14] Marcos Matabuena, Mario Francisco-Fernández, and Ricardo Cao. Predicting the physiological limits of sport stress tests with functional data. In *Functional Statistics and Related Fields*, pages 179–187. Springer, 2017.

[15] Sofiane Mnadla, Nicola Luigi Bragazzi, Mehdi Rouissi, Anis Chaalali, Anna Siri, Johnny Padulo, Luca Paolo Ardigò, Francesco Brigo, Karim Chamari, and Beat Knechtle. Infodemiological data of ironman triathlon in the study period 2004–2013. *Data in brief*, 9:123–127, 2016.

[16] Oliver Neubauer, Daniel König, and Karl-Heinz Wagner. Recovery after an ironman triathlon: sustained inflammatory responses and muscular stress. *European journal of applied physiology*, 104(3):417–426, 2008.

[17] Aleš Procházka, Saeed Vaseghi, Hana Charvátová, Ondřej Ťupa, and Oldřich Vyšata. Cycling segments multimodal analysis and classification using neural networks. *Applied Sciences*, 7(6):581, 2017.

[18] Samo Rauter, Iztok Fister Jr., and Iztok Fister. A collection of sport activity files for data analysis and data mining 2016a.

[19] Rainer Storn and Kenneth Price. Differential evolution - a simple and efficient heuristic for global optimization over continuous spaces. *J. of Global Optimization*, 11(4):341–359, December 1997.

[20] Inc. Strava. Connecting the world's athletes. https://www.strava.com/, 2017. [Accessed on: December 2017].

[21] Inc. Suunto. Suunto MovesCount App. http://www.suunto.com/movescountapp, 2017. [Accessed on: December 2017].

Recursive Partitioning in Predicting Energy Consumption of Public Buildings

Marijana Zekić-Sušac
University of Josip Juraj Strossmayer in Osijek
Faculty of Economics in Osijek
Trg Ljudevita Gaja 7, 31000 Osijek, Croatia
marijana@efos.hr

Adela Has
University of Josip Juraj Strossmayer in Osijek
Faculty of Economics in Osijek
Trg Ljudevita Gaja 7, 31000 Osijek, Croatia
adela.has@efos.hr

Saša Mitrović
University of Zagreb
Faculty of Organization and Informatics Varaždin, PhD student
Pavlinska 2, 42000 Varaždin, Croatia
smitrovic@foi.hr

Abstract. *Recursive partitioning includes a number of algorithms that create a classification or a regression decision tree by splitting the values of independent variables. The aim of this paper is to compare the accuracy of four different recursive partitioning methods in predicting the electrical energy consumption of public buildings. The input space included 141 attributes of public buildings in Croatia describing their geospatial, construction, heating, cooling, meteorological and energy characteristics. Four methods that produce regression tree partitioning were trained and tested. The results show that the random forest (RF) has outperformed CART, conditional inference tree (CTREE), and gradient boosted tree (GBT). The selection of important predictors was also compared and discussed.*

Keywords. Recursive partitioning, energy consumption, public buildings

1 Introduction

Previous research has shown that buildings are the largest individual energy consumers. More precisely, the building sector itself contains 40% of total primary energy consumption (Tommerup et al., 2007). Efficient models for predicting energy consumption could be particularly useful in public sector, where the state institutions need to recognize large consumers and allocate resources for improving its energy efficiency. Most of public buildings in Croatia still uses non-renewable energy resources and greenhouse gases, and reduction of such energy consumption is in accordance with EU directives and national strategic and action plans. Zekić-Sušac (2017) has shown that several approaches were used by researchers in modelling energy consumption: (1) individual statistical methods such as linear regression, time series analysis, probability density functions, or similar methods, (2) comparison of statistical methods with machine learning methods, and (3) simulation modelling.

This paper focuses on machine learning approach, more precisely on recursive partitioning methods and investigates their potential in predicting energy consumption of public buildings. Four different methods were used: classification and regression trees (CART), conditional inference trees (CTREE), random forest (RF), and gradient boosted trees (GBT) on a real dataset of Croatian public buildings. The aim was to investigate which of the recursive partitioning methods best fits the data and has a potential to be used as a modelling approach in reducing the cost of energy in public sector.

2 Previous research

Energy consumption and management is a frequent topic in recent research, due to the global need of reducing pollution, usage of non-renewable natural resources and green gas emission. Zekić-Sušac (2016) brings and overview of methods used in modelling energy efficiency and consumption. There are efforts in building prediction models of energy consumption in different countries by using statistical methods, machine learning, simulation (Chou and Bui, 2014). Some authors build prediction models for households, such as Farzana et al. (2014) who predicted the energy demand in the urban residential buildings of Chongqing in south west China. They have compared the accuracy of artificial neural networks (ANN),

Grey models, regression models, a polynomial model and a polynomial regression model. According to their research, the artificial neural networks outperformed other methods. The largest number of machine learning methods in this domain was used by Chou and Bui (2014) but they have used experimental datasets from the literature instead of real data. Their intensive methodological tests have shown that the ensemble of support vector regression (SVR) and ANN has outperformed CART method, chi-squared automatic interaction detector, general linear regression, and ensemble inference model. Regarding the input space, they have used only 8 input attributes and two output variables (cooling load (CL) and heating load (HL)). Mangold et al. (2015) used Swedish energy performance certificate data to describe energy usage in buildings. Chung and Park (2015) investigated energy consumption in buildings in South Korea.

The dataset from the public sector was used by Son et al. (2015) who predicted energy consumption of government-owned buildings based on an RreliefF variable selection algorithm and support vector machines method. Has and Zekić-Sušac (2017) has investigated the potential of artificial neural networks in predicting the energy efficiency level.

Yu et al. (2010) utilized decision tree method (C4.5 algorithm) with annual average air temperature, house type, construction type, floor area, heat loss coefficient, equivalent leakage area, number of occupants, space heating, hot water supply and kitchen as input variables on 80 residential buildings in Japan to estimate residential building energy performance. Research results of the aforementioned authors demonstrated that decision tree method can classify and predict building energy demand levels with a high accuracy (93% for training dataset and 92% for test dataset.

Tsanas & Xifara (2012) compared a classical linear regression approach, Reweighted Least Squares (IRLS), against non-parametric random forests (RF) method for predicting heating load and cooling load of 768 diverse residential buildings by using relative compactness, surface area, wall area, roof area, overall height, orientation, glazing area, glazing area distribution as input variables. Their research has shown that RF outperformed IRLS. Wang et al. (2018) used random forest (RF) for predicting energy consumption of two educational buildings in Florida state in the United States. Their dataset consisted of 11 input variables (meteorological, occupancy and time related data) while the methodology included RF, and regression tree (RT), and Support Vector Regression (SVR). The comparison has revealed that RF was more accurate then RT and SVR.

Papadopoulos et al. (2018) evaluated performances of random forests, extremely randomized trees (extratrees), and gradient boosted regression trees on Tsanas & Xifara (2012) dataset of 768 diverse residential buildings with 8 input variables. The results showed that tested tree partitioning methods outperformed other methods in recently published works of other researchers. Following the experience and guidelines from previous research which did not exploit recursive partitioning enough in this area, it was our challenge to put more focus on this type of machine learning methods. In addition, previous authors emphasized the importance of using occupational data (Liang et al., 2016) in addition to building characteristics. Our dataset is among the most extensive ones including a large number of building attributes as well as geospatial, heating, cooling, occupational, and meteorological data.

3 Data and model evaluation

The dataset was extracted from the database maintained by the Agency for Legal Trade and Real Estate Brokerage (APN) in Croatia. The initial data consisted of 2048 public buildings from Croatia such as kindergartens, schools, medical buildings, sport objects, health institutions, military, and all other types of buildings that are owned or occupied by the public sector. They were described by 141 attributes that can be grouped into geospatial, construction, heating, cooling, meteorological and energy data. The variable names are given in Table 1. The output variable was the total electricity consumption and the total natural gas consumption of each building in 2016. In the pre-processing stage, the outliers (i.e. the cases above the upper quartile) were detected and removed from the dataset variable, thus the final sample consisted of 1858 cases. For the purpose of this research the variable reduction was not used in the pre-modeling stage since all the tested methods incorporate algorithms for retaining the most important variables in the training process. Therefore, the variable reduction was conducted in the post-modeling phase.

Table 1. Variables describing public buildings

No.	Group of variables	Variable name
1.	Geospatial data	county, object region, type of object, object geo type, cultural heritage building
2.	Construction data	share of use of total building area, year of completion of construction, year of last restoration, flat gross floor area of building, useful area surface of building, object dim cooled area, object dim cooled surface area, object dim cooled volume area, number of floors, internal project temperature, share of windows surface
3.	Heating data	heated surface of the building, heated volume area of the building,

		installed power el. motor for pumps heat, type of heat pump, energy generating product, heating pump, total heat capacity of heat pump, total body heat radiator, total power body heat radiator, total body heat function oil, total power body heat function oil, total body heat other, total power body other, thermal power of heaters, primary heat sys using electrical heaters, installed capacity of electrical heaters, primary heating sys using split sys, installed electrical power of split sys heat, installed heat power of split sys heat, total heating power, factor of building shape f0, h1max. allowed coefficient of transmission heat loss per surface, transmission coefficient of heat loss, annual thermal energy needed 4heat, number of interior light luminaries
4.	Cooling data	object dimension of cooled area object dimension of cooled surface area object dimension of cooled volume area
5.	Meteorological data	air temperature
6.	Occupational data	number of employees, number of users, number of working days per week, number of working days per year, no of working hours per workday
7.	Energy coefficients of 9 specific parts of buildings d1,…,d9	object construction coeff. Transmission, object construction iso. thickness, object construction surface, object construction thickness (d1=roof, d2=floor, d3=windows, d4=shades, d5=heated ceiling, d6=unheated ceiling, d7=external wall, d8=doors, d9=unheated wall)
8.	**Output variable**	**Yearly electricity consumption (kWh)**

For the purpose of training and testing recursive partitioning methods, the total sample was randomly divided into the train and the test data, such that 70% of data (1486 cases) was used for training and 30% (372 cases) for testing. Data were normalized before training. The mean square error (MSE) is used as a common measure of performance in regression trees to determine final splitting. However, for final comparison of accuracy, we have followed the suggestion of Tofallis (Tofallis, 2015) to use the symmetric mean average percentage error (SMAPE) since it more fairly treats positive and negative residuals. The measure was computed according to:

$$SMAPE = 100\frac{1}{n}\sum_{i=1}^{n}\frac{|y_t - y_c|}{|y_t| + |y_c|} \qquad (1)$$

where y_t the real is output value, y_c is the predicted value, and n is the number of cases in the test sample.

4 Recursive partitioning methods

The four recursive partitioning methods were used, namely CART, CTREE, random forest and gradient boosted tree. All the computations were conducted using R software.

4.1 Classification and regression tree (CART)

The classification and regression tree (CART) suggested by Breiman et al. (1984) is the basic and most commonly used recursive partitioning method. In this research the regression variant of the CART was used such that the output produces a real number instead of a class probability. In its standard form, it builds a binary tree by splitting the input vectors at each node according to a function of a single input. For each input variable, the parent node is divided into child nodes by separating the objects with values lower and higher than the split point with the highest reduction of impurity. After repeating the splitting process for all input variables using each node as a new parent node until the tree reaches its maximum size, the stage of pruning begins and the algorithm prunes the tree back using the cross-validation procedure to select the right-sized tree. The algorithm considers all possible tree splits in order to find the most successful one by Gini index defined as (Apté et al., 1997):

$$Gini(t) = 1 - \sum_i p_i^2 \qquad (2)$$

where t is a current node and p_i is the probability of class i in t. Prune of misclassification error was used as the stopping rule, with minimum n=5. In the regression tree, the response for any observation is computed by following the path from the root node down to the appropriate terminal node of the tree, where the values for the splitting variables are observed, and the predicted response value is calculated by averaging response in that terminal node [10]. The limitation of CART trees is in their biasness regarding the variable selection, since it does not treat fair the variables of different types, categories, or missing values (Grömping, 2009).

4.2 Conditional inference tree (CTREE)

The conditional inference tree (CTREE) was proposed by Hothorn et al. (2006) as a tree partitioning method which does not use pruning, and is therefore faster than the CART and also overcomes the variable selection bias which exists in CART. The CTREE uses multiplicity-adjusted conditional tests to determine the predictors of an output and to generate a tree structure. It conducts the test of the null hypothesis of no association between an input variable and the output, and it calculates it both globally for each node and separately for each individual variable in non-terminal nodes. The smallest p-value is used to determine the variable which becomes a split variable. The tree grows until there is no further statistically significant split. This method is also robust since it can work with different types of variables and missing data.

4.3 Random forest (RF)

The random forest (RF), as the name of the method says, is a tree partitioning method that generates a collection of decision trees based on a random subset of the data, and each split within each tree is created based on a random subset of candidate variables (Hartshorn, 2016). The final response is obtained by averaging responses of the individual trees. According to the ambiguity decomposition (Krogh et al., 1995, in Louppe, 2014), the generalization error of the ensemble is to be lower than the average generalization error of its constituents. That is also the main advantage of the random forest method, since ensemble overcomes instability of single-tree techniques and improves its performance. The main shortcomings are in its complexity and computing time (Fan and Gray, 2005), (Grömping, 2009).]. The RF-CART algorithm was used in our experiments a random forest based on CART. In order to save computing time in pruning trees, we have used the complexity parameter cp=0.01 with ANOVA, such that the overall R-squared must increase by cp at each step. Splits that do not improve the fit by cp will likely be pruned off by cross-validation. The maximum depth parameter of any node was set to 30.

4.4 Gradient boosted trees (GBT)

This method uses boosting process in generating trees, meaning that trees are grown sequentially such that each successive tree uses information from previously grown trees. The aim is to minimize the error of the previous trees (Garreth et al., 2014). In this research the gradient boosting implemented in the R package xgboost was used. It involves resampling of observations and columns in each round of 10 cross-validation steps. The fitting of a decision tree is obtained by using the model residual errors as the outcome variable of the new model. The new decision tree adjusted by a shrinkage parameter *lambda* is added into the fitted function and the residuals are updated. *Lambda* of 0.01 was used in our research. Since it builds trees sequentially, this method is time-consuming comparing to other recursive partitioning methods. However, previous research has shown that it often outperformed other methods (Touzani et al., 2018).

5 Results

The graphical presentation of the tree partitioning for the first two methods that produce a single final tree is shown in Figure 1 (for CART) and in Figure 2 (for CTREE). It can be seen from Figure 1 that the CART method has extracted a smaller tree, with only 6 final splits and using only four features. The selected features are V9 (heated surface of the building) in the root node of the tree, V40 (total building power of cooling in kW) in the left branch, and variables V17 (number of working hours per workday) and V5 (number of users) in the right branch. The accuracy of the CART decision tree model in the sense of MAPE was 33.7324%. Parameters for CART were: Gini index, cp = 0.01, minsplit = 20, cp = 0.01, maxcompete = 4, maxsurrogate = 5, usesurrogate = 2, no. of cross-validation = 10, max.depth = 30.

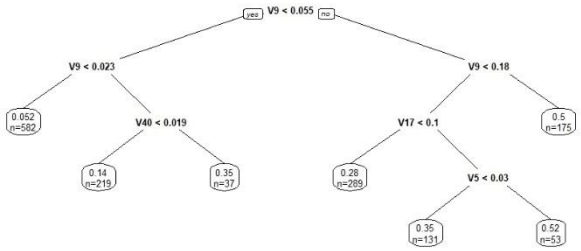


Figure 1. CART decision tree

The CTREE has produced a larger number of splits as shown in Figure 2. It has extracted 16 predictors as important, with XXX splits. Parameters for CTREE were: 19 terminal nodes, teststat = c("quad", "max"), testtype=c("Bonferroni", "MonteCarlo", "Univariate", "Teststatistic"), mincriterion=0.95, minsplit=20, minbucket=7, stump=FALSE, nresample=9999, maxsurrogate = 0, mtry = 10, savesplitstats = TRUE, maxdepth = 0. The accuracy of CTREE was slightly higher than the accuracy of CART, since it has produced MAPE of 31.21%.

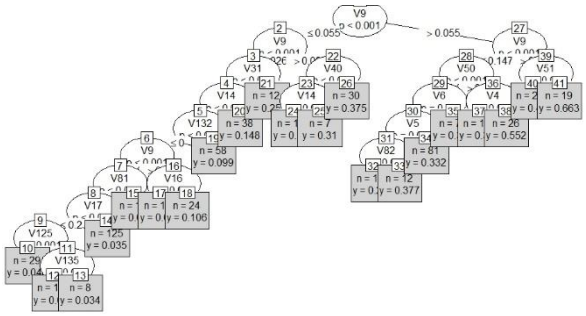


Figure 2. Conditional inference tree

The random forest method has created an ensemble of maximum 500 trees, while the number of variables tried at each split was 26. Since it does not produce a single tree as the output, it is possible to graphically observe the error conversion according to the number of generated trees, which is shown in Figure 3.

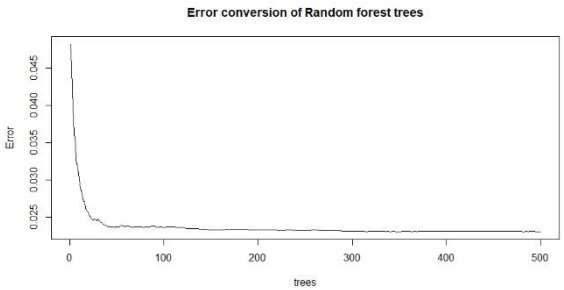


Figure 3. Error conversion of random forest

It can be seen in Figure 3 that the error converges as the number of trees increases and is relatively stable after 300 trees. The number of predictors extracted by random forest was 15, while its accuracy (MAPE) is 18.80%, which is higher than the previously tested CART and CTREE. The gradient boosted tree has produced the MAPE of 26.75%. The summary of the results of all four models is presented in Table 2.

Table 2. Accuracy of tested models

Recursive partitioning method	SMAPE (%)	No. of selected predictors
CART	33.73	4
CTREE	31.21	16
Random forest	18.80	15
Gradient boosted tree	26.75	20

It can be seen from Table 2 that the most accurate model was produced by the random forest method of recursive partitioning. In order to test the significance of the differences in results, the t-test of difference in proportions was conducted. The test has shown that the RF model is significantly different (p=0.000) from the CART model, as also significantly different from the CTREE model (p=0.000) and GBT model (p=0.049). Thus, it can be concluded that random forest performs significantly better than the other tested methods in modelling electricity consumption of public buildings.

Regarding the variable importance, it is interesting to analyse if the methods differ among themselves in selecting the important predictors. The selection of variables in all decision trees is obtained by using the information gain for regression trees calculated as (Hartshorn, 2016):

$$\boldsymbol{Information\ gain = n\left(MSE_{before} - MSE_{after}\right)}$$

where $\boldsymbol{MSE_{before}}$ is the MSE before the split, $\boldsymbol{MSE_{after}}$ is the MSE after the split, and *n* is the number of data points that split operated on. The variable importance is obtained by calculating information gain across all splits for a certain variable. Due to a lack of space, only first five highly ranked predictors extracted by each method were presented in Table 3.

Table 3. First five predictors extracted by each method

Rank	CART	CTREE	Random forest	Gradient boosted tree
1	V9 (heated surface of the building)	V9 (heated surface of the building)	V9 (heated surface of the building)	V9 (heated surface of the building)
2	V40 (total building power of cooling in kW)	V31 (installed electric power of split system for heating in kW)	V4 (number of employees)	V4 (number of employees)
3	V17 (number of working hours per workday)	V40 (total building power of cooling in kW)	V5 (number of users)	V5 (number of users)
4	V5 (number of users)	V14 (cooled volume area of the building in m^2)	V29 (total installed thermal power of heaters in kW	V6 (number of working days per week)
5	V4 (number of employees)	V132 (cool energy generating product code)	V70 (number of interior light luminaries)	V14 (cooled volume area of the building in m^2)

Table 3 reveals that the tested methods show certain similarities in selecting important predictors. The variable *V9 (heated surface of the building)* was ranked as the most important by all four methods. That variable belongs to the group of heating data (see Table 1) as well as the variable *V31 (installed electric power of split system for heating in kW)* extracted as the second most important by CTREE, *V29 (total installed thermal power of heaters in kW* and *V70 (number of interior light luminaries)* extracted as important variables by RF. The two variables from the group of occupational data: *V5 (number of users)* and *V4 (number of employees)* were selected by CART, RF, and GBT among the five most important ones but not by CTREE. The CART method has extracted an additional occupational variable: *V17 (number of working hours per workday)* while GBT extracted *V6 (number of working days per week).* The three variables from the group of cooling data were extracted in Table 3. The CART and CTREE have extracted *V40 (total building power of cooling in kW)*, while CTREE and GBT additionally selected *V14 (cooled volume area of the building in m2)* and *V132 (cool energy generating product code).*
The choice of important predictors generally shows that variables related to heating have the highest impact to electricity consumption, followed by the cooling-related variables, and occupational data.

6 Potentials for model implementation in reducing energy consumption in public sector

The experiments conducted in this research are a part of the research project "Methodological Framework for Efficient Energy Management by Intelligent Data Analytics" that aims to contribute the reduction of energy consumption of non-renewable natural resources by machine learning methods, such as artificial neural networks, recursive partitioning, support vector machines, and other methods. Due to the fact that buildings are the largest energy consumers, and that the state is in position to directly influence the energy consumption of public sector by allocating resources into measures to improve its energy efficiency, creating models that will support decisions on resource allocation is highly desirable.

Croatia has made significant steps by establishing the central Information system of energy management (ISGE) managed by the Agency for Legal Trade and Real Estate Brokerage (APN). However, the system still does not use machine learning to create prediction models or to extract important predictors of energy consumption. The models created in this research, especially the one based on random forest method can be implemented into ISGE as an intelligent module, a part of the web-based and mobile Internet of Things (IoT) applications that will automatically collect data, create models, and enable decision makers in determining actions that will lead to decreased energy consumption. Business analytics tools such as Alteryx, IBM Watson Analytics, Microsoft Azure Machine Learning, Amazon Web Services, and others enable to import data from ISGE system, R or Python scripts that define the algorithms, and create predictive models based on the methodological framework which selects the machine learning method that best fits the data and produce a prediction model.

7 Discussion and conclusion

The paper compares the accuracy and variable selection across four different recursive partitioning methods: classification and regression tree (CART), conditional inference tree (CTREE), random forest (RF), and gradient boosted tree (GBT) in modelling energy consumption of buildings in public sector. After data pre-processing, each method is trained and tested by using random subsampling procedure. The results have shown that the most accurate model was the one produced by the RF method which yielded the symmetric mean average percentage error (SMAPE) of 18.80%. The RF method significantly outperformed other tested recursive partitioning methods. The reason could be in the fact that RF uses ensemble of decision trees which improves the error in case of high-dimensional data that were present in this research. Although previous research showed that GBT usually produces more accurate results than other tree-partitioning methods, that was not the case in this research possibly due to a large number of input variables with a very similar effect on the error.

The most accurate RF model extracted 15 out of 141 predictors which belonged to the group of heating, cooling and occupational data. All four methods have extracted the variable *heated surface of the building* as the most important one.

The limitations of the recursive partitioning lie in the fact that in order to create efficient models for energy consumption of public buildings, other types of energy should be also considered, such as natural gas and water. In order to create a complete methodological framework, more machine learning methods should be compared and used in integrative manner. The created models have shown a potential of recursive partitioning methods in managing energy consumption in public buildings, and if implemented, could directly decrease energy consumption and expenditures in public sector, and significantly impact the state budget.

Acknowledgments

This work has been fully supported by Croatian Science Foundation under Grant No. IP-2016-06-8350 "Methodological Framework for Efficient Energy Management by Intelligent Data Analytics" (MERIDA).

References

Apté, C., & Weiss, S. (1997). Data mining with decision trees and decision rules. *Future generation computer systems*, 13(2-3), 197-210.

Breiman, L., Friedman, J., Ohlsen, R., & Stone, C. (1984). *Classification and regression trees*, Belmont, CA: Wadsworth International Group.

Chou, J. S., & Bui, D. K. (2014). Modeling heating and cooling loads by artificial intelligence for energy-efficient building design. *Energy and Buildings*, 82, 437-446.

Chung, M., & Park, H. C. (2015). Comparison of building energy demand for hotels, hospitals, and offices in Korea. *Energy*, 92, 383-393. DOI: 10.1016/j.energy.2015.04.016

Fan, G., & Gray, J. B. (2005). Regression tree analysis using TARGET. *Journal of Computational and Graphical Statistics*, 14(1), 206-218.

Farzana, S., Liu, M., Baldwin, A., & Hossain, M. U. (2014). Multi-model prediction and simulation of residential building energy in urban areas of Chongqing, South West China. *Energy and Buildings*, 81, 161-169.

Grömping, U. (2009). Variable importance assessment in regression: linear regression versus random forest. *The American Statistician*, *63*(4), 308-319.

Has, A., Zekić-Sušac, M., Modelling energy efficiency of public buildings by neural networks and its economic implications, in Zadnik-Stirn, L., Drobne, S. (Eds.), *Proceedings of the 14th International Symposium on Operations Research in Slovenia*, 27-29th September 2017, Bled, Slovenia. pp. 461-466.

Hartshorn, S. (2016). *Machine Learning With Random Forests And Decision Trees: A Visual Guide For Beginners*. Amazon Digital Services LLC: Seattle, Washington, USA.

Hothorn, T., Hornik, K., & Zeileis, A. (2006). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical statistics*, 15(3), 651-674.

Garreth, J., Witten, D, Hastie, T. & Tibshirani, R. (2014). *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company.

Liang, X., Hong, T., & Shen, G. Q. (2016). Improving the accuracy of energy baseline models for commercial buildings with occupancy data. *Applied energy*, 179, 247-260. doi: 10.1016/j.apenergy.2016.06.141.

Louppe, G. (2014). Understanding random forests, From theory to practice, PhD dissertation, University of Liège, Faculty of Applied Sciences, Department of Electrical Engineering & Computer Science. https://arxiv.org/pdf/1407.7502.pdf.

Mangold, M., Österbring, M., & Wallbaum, H. (2015). Handling data uncertainties when using Swedish energy performance certificate data to describe energy usage in the building stock. *Energy and Buildings*, 102, 328-336. doi: 10.1016/j.enbuild.2015.05.045.

Son, H., Kim, C., Kim, C., & Kang, Y. (2015). Prediction of government-owned building energy consumption based on an RReliefF and support vector machine model. *Journal of Civil Engineering and Management*, 21(6), 748-760.

Tofallis, C. (2015). A better measure of relative prediction accuracy for model selection and model estimation. *Journal of the Operational Research Society*, 66(8), 1352-1362.

Tommerup, H., Rose, J., Svendsen, S. (2007). Energy-efficient houses built according to the energy performance requirements introduced in Denmark in 2006. *Energy and Buildings*, 39(10), 1123-1130.

Touzani, S., Granderson, J., Fernandes, S. (2018). Gradient boosting machine for modelling the energy consumption of commercial buildings, *Energy and Buildings*, 158, pp. 1533-1543.

Zekić-Sušac, M. (2017). Overview of prediction models for buildings energy efficiency. In Mašek Tonković Anka (Ed.), *Proceedings of the 6th International Scientific Symposium Economy of Eastern Croatia – Vision and Growth*. (pp. 697 - 706). Faculty of Economics in Osijek, 25.05. - 27.05.2017.

Tsanas, A., & Xifara, A. (2012). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools.

Energy and Buildings, *49*, 560-567. doi:10.1016/j.enbuild.2012.03.003

Yu, Z., Haghighat, F., Fung, B. C. M., & Yoshino, H. (2010). A decision tree method for building energy demand modeling. Energy and Buildings, 42(10), 1637-1646. doi:10.1016/j.enbuild.2010.04.006

Wang, Z., Wang, Y., Zeng, R., Srinivasan, R. S., & Ahrentzen, S. (2018). Random Forest based hourly building energy prediction. Energy and Buildings, 171, 11-25. doi:10.1016/j.enbuild.2018.04.008

Papadopoulos, S., Azar, E., Woon, W. L., & Kontokosta, C. E. (2018). Evaluation of tree-based ensemble learning algorithms for building energy performance estimation. Journal of Building Performance Simulation, 11(3), 322-332. doi:10.1080/19401493.2017.1354919

Evaluating the Potential of a Data-Driven Approach in Digital Service (Re)Design

Tea Mijač, Mario Jadrić, Maja Ćukušić
Faculty of Economics, Business and Tourism, University of Split
Cvite Fiskovica 5, Split, Croatia
{tea.mijac, mario.jadric, maja.cukusic}@efst.hr

Abstract. *The amount of data is exponentially growing each day. With every interaction with digital services, users create their digital footprint. It is not unusual that usage data can demonstrate the need more objectively than users themselves leading to the fact that huge potential is hidden in data-driven development approaches. In order to investigate the reasoning behind using objective versus subjective data to changes in design of digital services, a pilot research study was conducted. The study was performed using A/B testing and a questionnaire. Original website was used as Design A, and with same functionalities kept, another webpage was launched (Design B). Among other findings, the analysis confirms that there is demonstrated need and rationale behind using objective users' data rather than subjective in testing the changes in design.*

Keywords. Digital service design, data-driven approach

1 Introduction

Number of digital artefacts is growing rapidly each day (Tomitsch, 2018). Since the amount of data produced by users nowadays exponentially increases, and by going through the more and more available usage data (generated through interaction with digital services), the development teams got the chance to understand what users are really doing and how they react better (King, Churchill, & Tan, 2017; Spiess, Joens, Dragnea, & Spencer, 2014). Since the user habits and interests are changing rapidly and new trends emerge daily, and having in mind that user experience (UX) is ultimately subjectively, dynamically and contextually dependent (Halvorsrud, Kvale, & Følstad, 2016), designers and developers have no choice but to take into consideration the data generated by different user actions and feedback collected from the overall user experience (Lee, Smith, Calvert, & Snajdr, 2016; Lemon & Verhoef, 2016).

Consequently, tracking the objective user data is extremely important and this data is a key component when evaluating the user experience (Sengers, Boehner, Mateas, & Gay, 2008). Although many studies confirm that direct contact with users is the key in user-oriented approaches and in fulfilling their expectations (Kujala, Kauppinen, & Rekola, 2001), the process of redesign is still often based on intuition rather than actual data (Havice, 2017) when, in effect, the designers could use e.g. the mouse movement, keyboard clicks, and so on, as the best input for the improvement (King et al., 2017). There is nothing more direct than the data that users produce themselves. With more and more data becoming available, the greater the chances are of understanding the users' needs (Anderson, 2015; Spiess et al., 2014).

The importance of using objective data in the context of improving the user experience has been a popular topic recently. For example, several authors studied data-driven development in telecommunications where they collected data to foresee potential customer complaints and tried to improve their user experience (Bao, Wu, & Liu, 2017). Similarly, Lee et al. (2016) investigated the approach in improving a library website. Despite the fact they highlighted the importance of using the objective data to improve the user experience, the main limitation of their research was using only the external source data – Google Analytics. Generally speaking, in the Human-Computer Interaction (HCI) field there is a lack of research about user experience metrics based on behavioural objective data produced by users themselves (Rodden, Hutchinson, & Fu, 2010).

In one of our recent studies (in the process of publication), we propose a methodological framework for user-oriented data-driven information systems modelling devised around the well-known IS development phases. The difference is that the proposed framework aims to emphasise the user experience and fosters the data-driven approach. The data can be used in order to either improve the user experience by way of eliminating critical errors, if these are detected, or to improve the whole users' journey while interacting with the system. In order to use the objective data in this whole process, metrics have to be devised and incorporated in the development phase. The proposed framework highlights these aspects as an important step/phase before any redesign.

Why we propose this methodological framework? Three major reasons why IT projects fail are the lack of information from end-users, unfinished specifications and frequent changes in specifications (Geogy & Dharani, 2016). Data-driven approach makes it easier to collect and understand the user needs and to provide much better quality of interaction – which takes us a step closer to collecting and analysing the actual needs of users as the basis for (re)design and development of user-oriented digital services. Data Driven Development (DDD) assumes that the development teams must base their decisions regarding new versions of digital services based on collected data (King et al, 2017; Maalej, Nayebi, Johann, & Ruhe, 2015). The development teams should be able to consider the requirements of the mass users when deciding on what needs to be developed (Spiess et al., 2014; Maalej et al., 2015). By adopting the user-oriented and data-driven approach - user experience should be improved. In general, the cycle of developing a digital service should never end.

To support the development of our framework, we need to inspect the data that is to be used in the process. It can come from different sources, external (tools used for passive tracking) or internal (server logs produced by users and by system itself). Despite the several divisions of objective data types, authors agree on two main types: passive (implicit) and active (explicit) data (Liikkanen, 2016; Maalej et al., 2015). Difference is that passive data comes from passive tracking such as session recordings of server logs and active data is collected via surveys, chats etc. (Liikkanen, 2016; Maalej et al., 2015). Authors (Rodden et al, 2010) also bring up types of metrics: (a) PULSE metrics (number of visits, activities per visitor etc.) and (b) HEART metrics: (1) Happiness – aesthetics, ease of use (2) Engagement – frequency of using, (3) Adoption – number of unique users (4) Retention – giving up using the service and (5) Task success – efficiency, effectiveness or error rate. Based on this, three types of (relevant) objective data (passive) are inspected further in the paper:

- Server logs – can contain all relevant information about former and current state of the system (digital service); there are of course many types of logs, depending on the defined settings.
- Visitors' metrics – most likely it is data collected via external sources such as Google Analytics.
- Visual metrics – heat maps and click maps.

The paper continues with a brief theoretical background addressing the methods relevant to data-driven approaches to support our study on the use of objective over subjective data in the digital services redesign projects. After the brief outline in section 2, the specific objectives and method is presented in section 3. The results of the survey and A/B experiment are presented in section 4 while the results are discussed and implications and limitations of the study are offered in section 5 and concluded in section 6.

2 Theoretical bases for the study

2.1 SUS and TAM as subjective measures

Standard ISO 9241-11 defines usability as an "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" (ISO, 1998). Usability is one of the main factors that influences the increase in the level of use of digital services (Huang & Benyoucef, 2014). The research has confirmed that end users prefer a website with a higher usability rating, although it is important to keep in mind that user's usability and design requirements depend on the type of digital service, users themselves, and the very purpose (Ilbahar & Cebi, 2017). As can be expected, some objective features of digital services such as loading speed, enhances user experience, and even speed perception itself has a positive impact on user experience (King et al, 2017). If for an example the task requires a large number of clicks, it is very likely that users will perceive that digital service as a complex and less usable (Venkatesh, Chan, & Thong, 2012).

Even though, there are a number of methods and theories in the literature for understanding, predicting, and assessing personal factors, behaviour, and the environment while interacting with software, the two most popular are System Usability Scale and Technology Acceptance Model (Harrati, Bouchrika, Tari, & Ladjailia, 2016). The System Usability Scale (SUS) is a well-researched and widely used questionnaire for assessing the usability of mostly web applications. It is considered that SUS is the simplest method which achieves the most reliable results according to sample size. With the number of 8 respondents used in SUS method, the expected accuracy of the results is over 75%, with the higher number of respondents the reliability increases, and the relevant conclusion can be deduced from the survey even if the sample size is 8 to 12 respondents (Brooke, 1996, 2013). The method was developed at DEC in 1986. It is a Likert scale where the respondent indicates his/her agreement or disagreement with the statement. By analysing each questionnaire one gets a result in the range from 0 to 100 – an indicator of the overall usability of the system being observed (ibid). Due to the lack of measurement for technology usage and acceptance, the TAM - Technology Acceptance Model has been developed in 1989 (Davis, 1989). To assess the user acceptance for technological products, it is one of the most well established models (Harrati et al, 2016). The first version of the model consisted of two variables that affect the acceptance or use of technology; the two variables are perceived usefulness and perceived ease of use. TAM has quickly become a dominant model for researching factors that affect user acceptance of technology (Marangunić & Granić, 2015). Variables of TAM model are also measured via a Likert scale.

2.2 Using server-side data metrics

Many authors explored the potential of server-side data-driven approaches in different contexts some of which are presented in table 1. As mentioned before, there are several types of server logs, which can be predefined and very useful. In addition to server logs, time metrics and error logs stand out as valuable data.

Table 1. Server-side data metrics

Server logs	Andrica & Candea, 2011; Garrido, Firmenich, Grigera, & Rossi, 2017; Gordillo, Barra, Aguirre, & Quemada, 2014; Grigera, Garrido, Rivero, & Rossi, 2017a; Harrati, Bouchrika, Tari, & Ladjailia, 2015; Inversini, Cantoni, & Bolchini, 2011; Maalej et al., 2015; Rodden et al., 2010a; Rodriguez, 2002
Time metrics	(Grigera et al, 2017)
Error rate	Au, Baker, Warren, & Dobbie, 2008; Rodden et al., 2010a

Regarding the external data sources, such as Google analytics, there are also several types of metrics, which can be used for this purpose. In one of our previous researches, we used only data from Google Analytics to develop a data-driven web persona (Mijač, Jadrić, & Ćukušić, 2018). Typical metrics are listed in table 2.

Table 2. Google Analytics metrics

High organic click-through rates for keyword(s)	(Bakaev, Bakaev, & Mamysheva, 2016; Lee et al, 2016; Rodden et al, 2010)
Bounce rate	(Bakaev et al, 2016; Lee et al, 2016)
Visitor traffic	(Bakaev et al, 2016; Lee et al, 2016; Rodden et al, 2010)
User activity	(Bakaev et al, 2016; Lee et al, 2016; Rodden et al, 2010)

2.3 Mouse metrics

Even though mouse metrics can also be obtained through server logs, these are usually separated from typical data server logs as it could be collected by using special software tools such as Mousotron. It is a mouse and keyboard activity monitor which enables tracking of different activities performed using a computer mouse, such as total number of left, right and double clicks, speed achieved and so on. (Blacksunsoftware, 2018). Table 3 below lists most common mouse metrics, used by different authors in their papers.

Table 3. Mouse metrics

Mouse clicks	(Andrica & Candea, 2011; Frantz, 2018; Garcia & Paiva, 2016; Harrati et al, 2015; Oertel & Hein, 2003)
Amount of scrolling and speed	(Au et al, 2008)

3 Research method

3.1 Research instrument and procedure

Drawing on the potential of generated usage data (presented in section 1) and based on the typical methods and metrics (presented in section 2), we devised a research study in order to substantiate the use of objective over subjective data in the digital services redesign projects. To be more precise, the purpose of the study was to investigate the sensitivity of objective and subjective variables to changes in design. An experiment using the A/B testing method was first performed. A/B testing is essentially an online experiment used when changes are made to a product/service to measure the effect (King et al, 2017; Lee et al, 2016; Rodden et al, 2010). In order to evaluate different versions of a digital product, in addition to the original website (design A), another one was made (design B). An experiment was then conducted with two different groups of participants (half of the classroom group A and the other half group B). With A/B testing it was possible to quantify the results and to compare the two versions of the design.

Besides automatically collecting usage data for two designs, data was also obtained using an online survey tool. The participants (second year undergraduate students of business studies) completed the questionnaires voluntarily and anonymously after the compulsory classes finished so as not to impact the results of the study. The exercise was conducted in a proctored environment, i.e. in computer labs and under supervision of a teaching fellow. The students were instructed to access the link to the online questionnaire, which was placed on the official e-learning website of the Faculty. They were instructed to close any other programs running on their computers. They were given enough time to complete the questionnaires finishing in approximately 30 minutes.

The questionnaire was divided in 5 parts. *Part 1* was designed to collect participants' demographic data and the information on whether users purchased online so far as well as an estimation of their own abilities in using web technologies. *Part 2* (Scenario) contained the link to the web page, (depending on a group the hyperlink led either to the design A or the design B) complemented with the instructions for executing four different tasks on the website. Participants received instructions to run the Mousotron software before starting each task and pause it after finishing it. After finishing each task, they also had to take the screenshot of Mousotron metrics and upload the picture to the foreseen place in part 2 of the survey. After performing each task, participants were required to evaluate the weight of the task as well as to evaluate time spent in comparison with their experience that they had in performing similar activities. *Part 3* contained the standard System Usability Scale questionnaire with 10 questions given to participants after briefly getting to know the system.

Part 4 contained Technology Acceptance Model Scale with nine Likert-type questions measuring: (1) perceived usefulness, (2) perceived ease of use as well as (3) intention for future usage. Three statements were allocated to each part of TAM. We used existing multi-item scales, adapted to suit the context of the study. Rating system was the same as in SUS, based on ratings 1 to 5 that correspond to the level of agreement with the two extremes: "Strongly agree" and "Strongly disagree". *Part 5* contained general questions regarding the overall satisfaction with the evaluated website.

Figure 1 illustrates the research procedure. Total number of the participants at the beginning was 161, but after allocating them to two different groups, some of the students did not access the survey as it was voluntary. Since the first task in the scenario was to register, to pick a random user name, and to enter the user name in the survey – we were able to match some survey data with the server-side data collected automatically. After excluding the uncompleted surveys (the reasons for leaving the survey remain unknown) and the ones that could not be matched with automatically collected data, the total number of participants in group A was 22 and in group B 36.

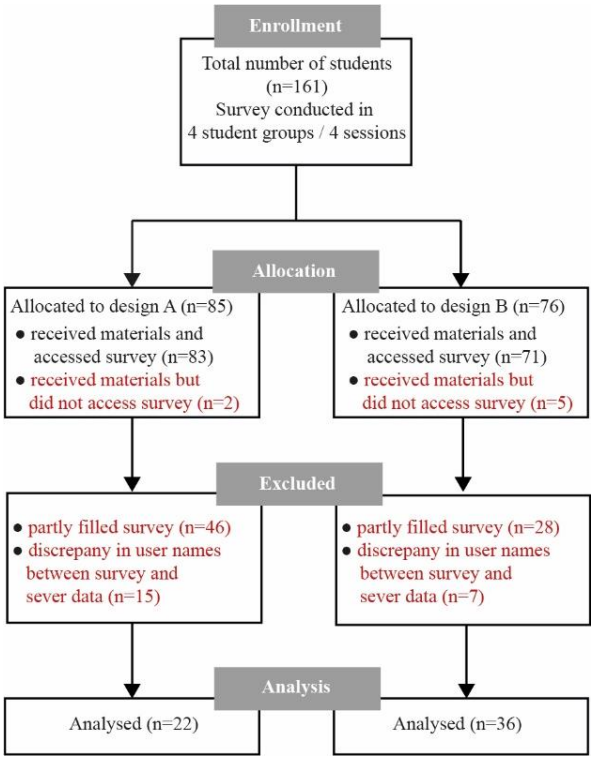


Figure 1. Research procedure with the number of participants (students)

3.2 Website and tasks for the A/B testing

Website we used for the study is from a Croatian start-up company selling natural cosmetics. The original website design is design A. Note that it was designed as a large number of other websites, without consulting the end-users. The website provides a portfolio of products and it has a functional webshop from where the users can buy available products. Additional functionality is the possibility register as a user which is useful if a user is purchasing the products often as they can earn additional discounts. The Figure 2 demonstrates the design A as a very clean design with minimalistic icons. Owners of the website would describe it as "trendy, plain and hipster". In order to make a fair comparison, the design B was built using the same text, the same pictures and the same functionalities, and the only thing that changed was the design. Design B was not plain or minimalistic; at the top of the website all the functionalities were listed. In general, the process of buying products required fewer clicks compared to the original design A. For the purpose of creating the alternative design, end-users were not consulted as well. As mentioned before, the scenarios contained four tasks in total and differed only with respect to the website design. In designing the websites and the tasks it was important to follow the well-established criteria for this kind of experimentation, i.e. the content in both scenarios had to be identical – except for the representational format used (website design); the representations used in both scenarios had to be equivalent in terms of the conveyed information (they had to be "informationally equivalent") and the required time to process these two scenarios had to be equivalent – without time limits.

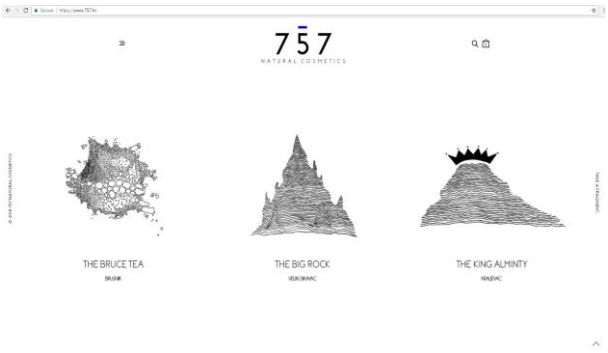


Figure 2. Design A of the website

Figure 3. Design B of the website

The set of four tasks which participants needed to perform are listed and described in Table 4.

Table 4. Task assigned to participants

Task name	Description
Registration	Find where you can register and complete the process. Please select a random username (e.g. NIVI89). For the purpose of completing the registration procedure, you do not need to enter a valid e-mail address. Please enter your username in the allocated field. When you finish the registration, close/leave the website.
Language change	Change the language of the website. When you change the language, close/leave the website.
Information	Find the information about the company and copy/paste it into the allocated field below.
Shopping cart	Find the soap "The Sea Heart – Galešnjak, 60g" and 'buy' it. Before selecting the payment type, take the screenshot of the shopping cart content, save and upload the picture.

3.2 Research participants

The participants of the study were the second-year undergraduate students of the University in Split, Faculty of Economics, Business and Tourism. General statistics of the sample population is presented in Table 5. The participants are from a relatively homogeneous group and mostly share a similar background in terms of education, economic situation etc. Based on the gender and age of the participants it is considered that the sample is representative when compared to the total population and between two groups. Majority of faculty's students are female (almost 80%) and with regards to age from 19 to 22 years old (around 97%).

Table 5. General statistics

		Group A (N=22)		Group B (N=36)		Total (N=58)	
		N	%	N	%	N	%
Gender	Female	19	86,4	27	75	46	79,3
	Male	3	13,6	9	25	12	20,7
Age	19-20	18	81,8	22	61,1	40	69,0
	21-22	3	13,6	13	36,1	16	27,6
	23 and more	1	4,5	4	11,1	2	3,5
Online buying	Yes	17	77,3	33	91,7	50	86,2
	No	5	22,7	3	8,3	8	13,8

4 Research results

As reported, the aim was to investigate whether there are significant differences between subjective (SUS and TAM) and objective data (server side data and mouse data) between groups A and B. The experiment in effect demonstrates the extent to which the objective data is sensitive to changes in design. The results are presented and interpreted hereinafter.

4.1 SUS and TAM results

Results of the SUS survey are presented in Table 6. SUS score does not represent a percentage; a mean score of 68 would represent a grade "C" and anything below a score of 51 is an "F" (putting a website in the bottom 15%) (Mclellan, Muddimer, & Peres, 2012). Following this interpretation, Group A (or design A) got an unacceptable usability score and design B got an acceptable usability score. The Mann-Whitney U test confirms that there is a statistically significant difference between the results of the two groups.

Table 6. System Usability Scale (SUS) results

	Group A (N=22)	Group B (N=36)	Total (N=58)
Mean	**50,7955**	**79,9306**	68,8793
Median	51,2500	80,0000	72,5000
St. dev.	15,34	11,60	19,30
Mode	50,00	75,00*	50,00
Minimum	15,00	50,00	15,00
Maximum	77,50	97,50	97,50
Mean Rank	13,64	39,19	
Mann-Whitney U			745,00
Asymptotic Sig. (2-sided test)			0,000

In analysing TAM results, first internal consistency of the scales of the questionnaire was analysed for both groups to determine the internal consistency. Cronbach alpha coefficient results are presented in Table 7. According to the boundaries for the Cronbach alpha coefficients (0.90/excellent, 0.80/very good and 0.70/ satisfactory), all TAM variables (from 0.875 to 0.922) are considered acceptable for further analysis. Results in table 8 demonstrate that all TAM variables significantly differ between the two groups.

Table 7. Cronbach alpha coefficients

TAM factors	Cronbach's Alpha (N=58)
Ease of use	0,922
Usefulness	0,875
Intention to use	0,919

Table 8. Descriptive statistics for TAM

	Group	Mean	Median	Mode	Min	Max	Mean Rank	p
Ease of use	A	3,15	3,00	3,00	1,00	5,00	15,32	**0,000**
	B	4,60	5,00	5,00	2,00	5,00	38,17	
Usefulness	A	3,29	3,33	3,00	1,00	5,00	19,82	**0,001**
	B	4,26	4,50	5,00	2,67	5,00	35,42	
Intention to use	A	2,30	2,67	1,00	1,00	4,00	20,66	**0,002**
	B	3,38	3,17	3,00	1,33	5,00	39,90	

After performing each task, the part 2 of the questionnaire required from the participants to evaluate the weight of the task execution as well as to evaluate the time spent in comparison with the experience they

had performing similar activities. The results of the Mann-Whitney U test for group A and group B showed statistical difference for each task (the highest p value among differences for each task was 0,039). All the answers from group A were ranked lower than the results from group B. The results regarding weight of the first task (registration) in group A showed mean rank 23,73 and for group B 33,03; as for the time spent – group A has mean rank of 21,23 and group B of 34,56. Second task in group A has mean rank 24,02 and group B 32,85, while for the time spent the group A has mean rank of 22,86 and group B of 33,56. For the third task, both the results for evaluating the weight (A=23,73 and B=33,03) and the time (A= 22,82 and B=33,58) also point that the results for group A have lower mean rank than for group B. Results for the fourth task are also consistent in that regard since mean rank for evaluating the weight of the task is 21,14 for group A and 34,61 for group B, as well as the results regarding time spent (A=21,11 and B=34,62).

4.2 Server side data results

Server side data was data collected automatically while participants interacted with the websites. Even though there was a huge amount of data, for the purpose of this research, several metrics were used:

- visit duration in seconds – is refers to the total amount of time spent on the website,
- number of searches – number of times when participants used the "search" option on the website,
- number of actions – action is every page participants' visit,
- number of extra actions – this metrics was calculated as the difference between the "number of actions" and the "number of actions spent for finishing a task"; therefore, it represents a number of extra actions which a participant performed, probably in order to get to know the website,
- extra time needed.

Results presented in Table 9 include only the metrics with statistically significant difference between the group A and group B. Statistical data analysis was also done at the level of each task, but there were no significant differences due to a small sample.

Table 9. Server side data results

Metrics	Group	Mean	Median	Mean Rank	U	p
Searches	A	0,50	0,00	33,86	300,00	**0,007**
	B	0,06	0,00	26,83		
Extra actions	A	7,54	7,00	35,73	259,00	**0,027**
	B	4,89	4,50	25,69		
Extra time	A	207,45	165,00	37,34	223,50	**0,006**
	B	93,14	86,50	24,71		

It should be noted that even though Google Analytics was set up, due to the controlled conditions of the survey, Google Analytics did not provide any other useful data (apart from number of visits, and time stamp of visits, all consistent with the number of participants and survey feedback).

4.3 Mouse metrics results

Mouse metrics were also collected automatically using the Mousotron tool. It enabled collecting several types of metrics: (1) keystroke, (2) left button, (3) right button, (4) double clicks, (5) mouse wheel, (6) speed, (7) seconds, (8) idle seconds and (9) centimetres. Since the participants needed to upload the results/collected statistics after each task, it was possible to process the data on the level of each task. For the first task (registration) the results of mouse metrics, which represent statically significant difference ($p < .05$), are presented in Table 10. For the second task (language change) only for the 'mouse wheel' movement there was statically significant difference ($p = .004$). For the third task, the collected mouse metrics differ significantly between groups for 'keystroke', 'mouse wheel' and 'centimetres'. Results are presented in Table 11. An examination of the findings of the mouse metrics referring to the last task (buying the product) revealed that average ranks for each metrics are not significantly different.

Table 10. Mouse metrics for the first task

Task Registration	group	Mean	Median	Mode	Min	Max	Mean Rank	U	p
Keystroke	A	148,75	145,50	20,00	20,00	257,00	34,75	235,00	0,033
	B	115,97	114,00	98,00	0,00	244,00	25,03		
Left button	A	66,45	64,00	65,00	15,00	168,00	38,60	158,00	0,001
	B	40,19	39,00	23,00	0,00	99,00	22,89		
Mouse wheel	A	41,20	31,50	0,00	0,00	120,00	34,32	243,50	0,045
	B	24,08	13,00	0,00	0,00	133,00	25,26		
Seconds	A	416,15	410,00	51,00	51,00	667,00	39,10	148,00	0,00
	B	250,47	237,50	8,00	8,00	640,00	22,61		
Idle Seconds	A	62,90	58,00	0,00	0,00	155,00	35,12	227,50	0,022
	B	26,53	22,00	0,00	0,00	122,00	24,82		
Centimeters	A	2159,09	1945,40	2038,94	761,48	4694,00	38,63	140,00	0,000
	B	1260,25	1088,75	40,00	40,00	3900,00	22,39		

Table 11. Mouse metrics for the third task

Information	group	Mean	Median	Mode	Min	Max	Mean Rank	U	p
Keystroke	A	26,90	4,00	0,00	0,00	301,00	35,05	251,00	0,036
	B	16,78	3,00	0,00	0,00	202,00	25,47		
Mouse wheel	A	139,86	92,00	0,00	0,00	527,00	40,00	147,00	0,000
	B	27,64	14,50	12,00	0,00	156,00	22,58		
Centimeters	A	1952,19	767,00	212,48	212,48	9186,36	35,05	251,00	0,036
	B	967,94	496,00	496,00	239,00	6888,50	25,47		

5 Discussion

The general idea of this paper was to examine the potential of data driven approach for (re)designing the digital services. Although there are a number of limitations to this research, especially the number of participants, restricted conditions and not using all the available objective data – the research reveals some interesting results.

Our main question was whether objective data are sensible to changes in design, and the general answer is positive. There was significant statistical difference in the subjective and objective data results. The results are consistent with theory implications that, if something takes more actions or requires more clicking and scrolling, the subjective opinion of users is worse. Even though the two provided designs kept the same functionalities, the results demonstrate the difference in both objective and subjective data.

Both **subjective methods** (SUS and TAM) show that users prefer design B to design A. The results for the overall website, they were not done at the level of each task. Results demonstrate that users' evaluation of the weights of performing each task as well as the time spent on each task, statistically differs for each task in favour of design B.

As mentioned, one of the restrictions of this research is the number of participants, consequently after analysing the results of **server side data** for each task of design A and design B – it turns out that there was no statistically significant difference in server side data logs. However, by analysing the results for four tasks altogether, for example, total number of "extra time", the results were statistically different between design A and design B. Server side data was consistent with above-mentioned results. After analysing the data from Mousotron software, the results also point to statistically different results at the level of each task (task 1, task 2 and task 3). Results of **mouse metrics** were therefore also consistent with the above results.

In the subsequent phases, the framework we mention throughout the paper would be further developed and validated through pilot studies such as this one.

6 Conclusion

To conclude, even though without actually asking the end users how they feel about a digital service, or whether there is something that could be improved – by using and following the data-driven approach enough data could be obtained to use it for (re)designing the digital services. This could be very practical when there is huge amount of users with different backgrounds and in the conditions where it could be hard to collect user specifications.

This approach should not be used in isolation, it can be helpful to provide additional information so the development teams can predict possible issues without waiting for formal users' feedback, as mentioned in proposed methodological framework for user-oriented data-driven information systems modelling that is under development by the authors of the paper.

Acknowledgment

This work has been supported by the Croatian Science Foundation under the project UIP-2017-05-7625: User-oriented process (re)design and information systems modelling – a case of smart city services.

References

Anderson, C. (2015). *Creating a Data-Drivem Organization*. Sebastopol, Ca, SAD: O'Reilly.

Andrica, S., & Candea, G. (2011). WaRR: A tool for high-fidelity web application record and replay. *Proceedings of the International Conference on Dependable Systems and Networks*, 403–410. https://doi.org/10.1109/DSN.2011.5958253

Au, F. T. W., Baker, S., Warren, I., & Dobbie, G. (2008). Automated Usability Testing Framework. *Proceedings of the Ninth Conference on Australasian User Interface - Volume 76*, (January), 55–64. Preuzeto od http://dl.acm.org/citation.cfm?id=1378337.1378349

Bakaev, M., Bakaev, M., & Mamysheva, T. (2016). Current Trends in Automating Usability Evaluation of Websites Can You Manage What You Can ' t Measure ? Current Trends in Automating Usability Evaluation of Websites. *Ieee*, (August), 510–514.

Bao, Y., Wu, H., & Liu, X. (2017). From Prediction to Action: Improving User Experience with Data-Driven Resource Allocation. *IEEE Journal on Selected Areas in Communications*, *35*(5), 1062–1075. https://doi.org/10.1109/JSAC.2017.2680918

Blacksunsoftware. (2018). Blacksunsoftware. Preuzeto od http://www.blacksunsoftware.com/

Brooke, J. (1996). *SUS - A quick and dirty usability scale*. London, UK: Taylor and Francis.

Brooke, J. (2013). SUS : A Retrospective. *Journal of Usability Studies*, *8*(2), 29–40. https://doi.org/10.1074/jbc.R115.675280

Davis, F. D. (1989). Perceived Usefulness , Perceived Ease of Use , and User Acceptance of. *MIS Quarterly*, *13*(3), 319–340.

Frantz, T. L. (2018). Blockmap: an interactive visualization tool for big-data networks. *Computational and Mathematical Organization Theory*, *24*(2), 149–168. https://doi.org/10.1007/s10588-017-9252-6

Garcia, J. E. . c, & Paiva, A. C. R. . c c. (2016). An automated approach for requirements specification maintenance. *Advances in Intelligent Systems and Computing*, *444*, 827–833. https://doi.org/10.1007/978-3-319-31232-3_78

Garrido, A., Firmenich, S., Grigera, J., & Rossi, G. (2017).

Data-driven usability refactoring: Tools and challenges. *SoftwareMining 2017 - Proceedings of the 2017 6th IEEE/ACM International Workshop on Software Mining, co-located with ASE 2017*, (October).https://doi.org/10.1109/SOFTWAREMINING.2017.8100854

Geogy, M., & Dharani, A. (2016). A Scrutiny of the Software Requirement Engineering Process. *Procedia Technology*, *25*, 405–410. https://doi.org/10.1016/j.protcy.2016.08.125

Gordillo, A., Barra, E., Aguirre, S., & Quemada, J. (2014). The usefulness of usability and user experience evaluation methods on an e-Learning platform development from a developer's perspective: A case study. *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*, 1–8. https://doi.org/10.1109/FIE.2014.7044340

Grigera, J., Garrido, A., Rivero, J. M., & Rossi, G. (2017). Automatic detection of usability smells in web applications. *International Journal of Human Computer Studies*, *97*(October 2017), 129–148. https://doi.org/10.1016/j.ijhcs.2016.09.009

Halvorsrud, R., Kvale, K., & Følstad, A. (2016). Improving service quality through customer journey analysis. *Journal of Service Theory and Practice*, *26*(6), 840–867. https://doi.org/10.1108/JSTP-05-2015-0111

Harrati, N., Bouchrika, I., Tari, A., & Ladjailia, A. (2015). Automating the evaluation of usability remotely for web applications via a model-based approach. *NTIC 2015 - 2015 1st International Conference on New Technologies of Information and Communication, Proceeding*. https://doi.org/10.1109/NTIC.2015.7368757

Harrati, N., Bouchrika, I., Tari, A., & Ladjailia, A. (2016). Exploring user satisfaction for e-learning systems via usage-based metrics and system usability scale analysis. *Computers in Human Behavior*, *61*, 463–471. https://doi.org/10.1016/j.chb.2016.03.051

Havice, J. (2017). How to Create Customer Personas with Actual, Real Life Data. Preuzeto od http://bit.ly/2FrhRdx

Huang, Z., & Benyoucef, M. (2014). Usability and credibility of e-government websites. *Government Information Quarterly*, *31*(4), 584–595. https://doi.org/10.1016/j.giq.2014.07.002

Ilbahar, E., & Cebi, S. (2017). Classification of design parameters for E-commerce websites: A novel fuzzy Kano approach. *Telematics and Informatics*. https://doi.org/10.1016/j.tele.2017.09.004

Inversini, A., Cantoni, L., & Bolchini, D. (2011). Design, User Experience, and Usability. Theory, Methods, Tools and Practice, *6770*(May 2014). https://doi.org/10.1007/978-3-642-21708-1

ISO. (1998). ISO 9241-11:1998. Preuzeto od https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en

King, R., Churchill, E. F., & Tan, C. (2017). *Designing with Data*. O'Reilly.

Kujala, S., Kauppinen, M., & Rekola, S. (2001). Bridging the Gap between User Needs and User Requirements. *Advances in Human-Computer Interaction I (Proceedings of the Panhellenic Conference with International Participation in Human-Computer Interaction PC-HCI 2001)*, (February), 45–50.

Lee, Y. Y., Smith, A., Calvert, L., & Snajdr, E. (2016). Innovative Data-Driven Methods to Improve Digital User Experience. *Qualitative & Quantitative Methods in Libraries*, *5*(2), 461. Preuzeto od http://search.ebscohost.com/login.aspx?direct=true&AuthType=ip,cookie,url,cpid,uid&custid=s8863137&db=edb&AN=120969018&site=eds-live&scope=site&authtype=ip,uid

Lemon, K. N., & Verhoef, P. C. (2016). Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing*, *80*(6), 69–96. https://doi.org/10.1509/jm.15.0420

Maalej, W., Nayebi, M., Johann, T., & Ruhe, G. (2015). Towards Data - Driven Requirements Engineering, *33*, 1–6. https://doi.org/10.1109/MS.2015.153

Marangunić, N., & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society*, *14*(1), 81–95. https://doi.org/10.1007/s10209-014-0348-1

Mclellan, S., Muddimer, A., & Peres, S. C. (2012). The Effect of Experience on System Usability Scale Ratings. *Journal of Usability Studies*, *7*(2), 56–67.

Mijač, T., Jadrić, M., & Ćukušić, M. (2018). The Potential and Issues in Data-Driven Development of Web Personas. U *MIPRO 2018* (str. 1427–1432).

Oertel, K., & Hein, O. (2003). Identification of Web usability problems and interaction patterns with the RealEYES-iAnalyzer. *Interactive Systems Design, Specification, and Verification 10th International Workshop, DSV IS 2003 Revised Papers Lecture Notes in Comput Sci Vol 2844*, *2844*, 77–91. Preuzeto s https://www.scopus.com/inward/record.uri?eid=2-s2.0-0142184064&partnerID=40&md5=62bbbc0cdfe9cbbd7cbd4f86da6a4cde

Rodden, K., Hutchinson, H., & Fu, X. (2010). Measuring the User Experience on a Large Scale : User-Centered Metrics for Web Applications. *SIGCHI Conference on Human Factors in Computing Systems*, 2395–2398. https://doi.org/10.1145/1753326.1753687

Rodriguez, M. G. (2002). Automatic data-gathering agents for remote navigability testing. *IEEE Software*, *19*(6), 78–85. https://doi.org/10.1109/MS.2002.1049396

Sengers, P., Boehner, K., Mateas, M., & Gay, G. (2008). The disenchantment of affect. *Personal and Ubiquitous Computing*, *12*(5), 347–358. https://doi.org/10.1007/s00779-007-0161-4

Spiess, J., Joens, Y. T., Dragnea, R., & Spencer, P. (2014). Using Big Data to Improve Customer Experience and Business Performance. *Bell Labs Technical Journal*, *18*(4), 3–17.

Tomitsch, M. (2018). *Making cities smarter*. Berlin: Jovis.

Venkatesh, V., Chan, F. K. Y., & Thong, J. Y. L. (2012). Designing e-government services: Key service attributes and citizens' preference structures. *Journal of Operations Management*, *30*(1–2), 116–133. https://doi.org/10.1016/j.jom.2011.10.001

Quality of Software and Services

Developing a Compliant Cybersecurity Process for Medical Devices

Nadica Hrgarek Lechner
MED-EL Elektromedizinische Geräte GmbH
Fürstenweg 77a, 6020 Innsbruck, Austria
nadica.hrgarek@medel.com

Abstract. *Cybersecurity is growing in importance for medical device manufacturers, health care facilities, clinicians, patients, and regulators. The purpose of this paper is to propose the approach or methodology that can support medical device manufacturers to develop a compliant cybersecurity process as an integral part of their quality management systems to systematically manage cybersecurity risks.*

Keywords. cybersecurity, FDA, medical devices, privacy, quality management system, risk management, security, wireless

1 Introduction

Over the last few decades, medical devices have evolved from isolated equipment to networked devices with wireless communication and remote connectivity (Burns et al., 2016). In the past the focus was primarily on essential performance and safety, and less on security of medical devices. This paper provides some guidelines to develop a compliant cybersecurity process for design and development of medical devices. The paper is divided into four sections. Section 1 gives a brief overview of medical devices and emphasizes importance of cybersecurity for medical device manufacturers. The proposed cybersecurity model is briefly described in Section 2. In Section 3 we give some insights into cybersecurity as an integral part of a quality management system. Some conclusions are drawn in the last section.

Software was first used in medical devices in the 1980's (McHugh, 2015). According to Sarig (2012), the amount of software built into medical devices doubles about every two years. Embedding the software into medical devices can reduce development and maintenance costs. In addition, it can introduce new opportunities such as bringing new innovative products to the market faster, increasing clinical effectiveness, providing better services to end users through predictive and preventive maintenance, improving user experience, etc. As medical devices tend to change over time and become increasingly interconnected, implementation of new features including connections to the cloud, databases, third-party and open source software, IoT, hospital/health care facility networks, and other medical and non-medical devices leads to larger attack surfaces, associated with the increased complexity of the entire system and use models. To prevent cybersecurity incidents, Williams and Woodward (2015) point out that it is important to recognize the complexity of the operational environment as well as to catalog the technical vulnerabilities.

Software incorporated in connected medical devices such as remote-controlled drug infusion systems, defibrillators, cardiac pacemakers, and network-connected X-ray machines is vulnerable to cybersecurity threats. Medical devices that are connected to a public network like the Internet could be exploited by a threat actor through a single cybersecurity vulnerability. Some exploits could affect integrity of health data, availability of patient care, or even how a medical device operates. For example, malware infection can cause a device to slow down and miss critical interrupts and therefore, clinicians cannot trust the integrity of the sensor readings (Fu & Blum, 2013). Some vulnerabilities may cause the system to stop working which is especially dangerous for implantable, life-saving, and life-sustaining medical devices ("Bug can cause deadly failures when anesthesia device is connected to cell phones," 2014). Fu and Blum (2013) have raised some concerns about risks of depending on unsupported software (e.g., some medical devices still rely on Windows XP operating system with service packs and security patches). A compromised medical device may also serve as access point for entry into hospital networks to steal confidential data ("MEDJACK: Hackers hijacking medical devices to create backdoors in hospital networks," 2015). Using a networked medical device as means for intrusion may lead to compromise of other medical devices (e.g., in operating room), loss or exposure of sensitive and confidential patient information, or a safety issue. According to PwC's Health Research Institute (2017) consumer survey, 38% of consumers would be wary of using a hospital associated with a hacked medical device.

2 The Generic Cybersecurity Model for the Medical Device Industry

2.1 Background

Cybersecurity and information security are commonly used interchangeably; however, these terms differ. Cybersecurity is a part of information security (Spremić & Šimunic, 2018). Cybersecurity is defined as "the protection of information assets by addressing threats to information processed, stored, and transported by internetworked information systems" (ISACA, 2016, p. 9). The FDA (2014, p. 3) defines cybersecurity as "the process of preventing unauthorized access, modification, misuse or denial of use, or the unauthorized use of information that is stored, accessed, or transferred from a medical device to an external recipient". Information security "ensures that within the enterprise, information is protected against disclosure to unauthorized users (confidentiality), improper modification (integrity), and non-access when required (availability)" (ISACA, 2016, p. 15).

The cybersecurity process plays a vital role in the field of medical and health technology. The main objective is to design and develop medical devices that are secure throughout the whole life cycle without compromising patient safety. As illustrated in Fig. 1, privacy and security must be considered early in conception and design of medical devices, become a part of medical device architecture, and end with obsolescence of medical devices.

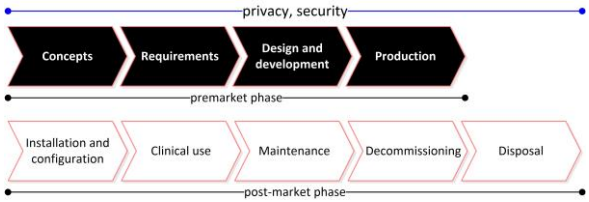


Figure 1. Medical device security life cycle

The model of confidentiality, integrity, and availability (also known as the CIA triad) is illustrated in Fig. 2. This model could be used as a starting point to implement a cybersecurity process. The main goal of the CIA triad is to apply appropriate security controls when data is stored, in processing, or in transit. The CIA triad alone is not enough to develop an effective cybersecurity strategy for medical devices. The ISO/IEC 27000 family of standards could assist the medical device manufacturers to keep key assets secure. Besides that, there are many federal government laws, regulations, standards, technical reports, and guidance combined with industry best practices that deal with information security and cybersecurity vulnerabilities of medical devices. As shown in Table 1, every component of the CIA triad can be mapped to applicable regulations, standards, and guidance documents. We suggest keeping a close eye on cybersecurity regulations and adapting processes accordingly to develop safe, effective, and secure medical devices. A brief overview of cybersecurity regulations and standards for medical devices is provided in (Hrgarek Lechner, 2017).

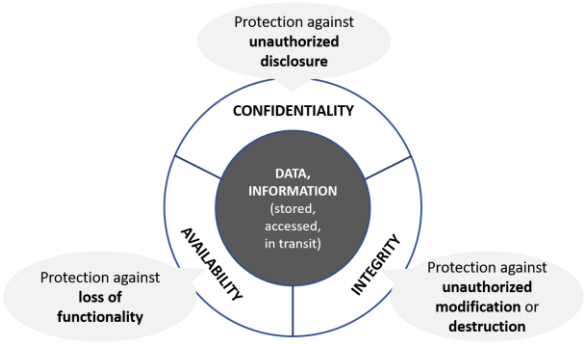


Figure 2. The CIA triad

Table 1. Mapping between CIA triad components and regulations, standards, and guidance documents

CIA triad component	Regulation, standard, guidance
Confidentiality	• AAMI TIR57 • BS EN 45502-1 • Content of Premarket Submissions for Management of Cybersecurity in Medical Devices • IEC TR 80001-2-8 • ISO/IEC 27000 • ISO/IEC 27001 • Medical Device Regulation • NIST cybersecurity framework • NIST SP 800-39 and 800-53 • Postmarket Management of Cybersecurity in Medical Devices • UL 2900-1 • UL 2900-2-1
Integrity	• AAMI TIR57 • BS EN 45502-1 • Content of Premarket Submissions for Management of Cybersecurity in Medical Devices • Health Insurance Portability and Accountability Act (HIPAA) • IEC 60601-1+AMD1 • IEC 62304+AMD1 • IEC 82304-1 • IEC TR 80001-2-8 • ISO/IEC 27000 • ISO/IEC 27001 • NIST cybersecurity framework • NIST SP 800-39 and 800-53 • Postmarket Management of Cybersecurity in Medical Devices • UL 2900-1 • UL 2900-2-1

CIA triad component	Regulation, standard, guidance
Availability	• AAMI TIR57 • BS EN 45502-1 • Content of Premarket Submissions for Management of Cybersecurity in Medical Devices • ISO/IEC 27000 • ISO/IEC 27001 • NIST cybersecurity framework • NIST SP 800-39 and 800-53 • Postmarket Management of Cybersecurity in Medical Devices

The FDA (2014) recommends considering five core functions of the NIST cybersecurity framework to guide cybersecurity activities. When developing a generic cybersecurity model that can be tailored to meet the regulatory requirements for the design and development of medical devices, we adapted the core functions of the NIST cybersecurity framework. The first version of the NIST framework ("Framework for Improving Critical Infrastructure Cybersecurity Version 1.0," 2014) was published in February 2014. According to a Gartner report ("Best Practices in Implementing the NIST Cybersecurity Framework," 2016), this version has been adopted by 30% of US companies and is expected to grow to 50% by 2020. In April 2018, a newer version of the NIST framework ("Framework for Improving Critical Infrastructure Cybersecurity Version 1.1," 2018) was released. The NIST's Framework Core consists of five functions: Identify, Protect, Detect, Respond, and Recover. Functions are subdivided into total 23 categories and 108 subcategories. Informative references are mapped to each subcategory.

2.2 Prerequisites

Our generic cybersecurity model that is further described in upcoming sections pre-assumes that the medical device manufacturer has established the risk management process to address and document all risks, including security risks with safety impact, throughout the whole medical device's life cycle. The elements of a cybersecurity vulnerability and management approach as part of the software validation and risk analysis are listed in FDA's (2014) guidance document. AAMI TIR 57 (2016), NIST SP 800-30 Rev. 1 (2012), HIMSS/NEMA Standard HN 1-2013 (2013), ISO/IEC 27005 (2018), OCTAVE® (Operationally Critical Threat, Asset, and Vulnerability Evaluation) approach (Alberts et al., 2003), and a white paper published by Medical Device Privacy Consortium (2014) can be used to guide the implementation of the security risk management process.

To determine appropriate security controls, the medical device manufacturer should start with identifying the critical assets that need to be protected, threats, and vulnerabilities that expose assets to the threats. As illustrated in Fig. 4, causal chain of security threats begins with a threat source initiating a threat event. If a threat source successfully exploits a device vulnerability and gains access to assets, this may result in an adverse impact due to a compromise of the device confidentiality, integrity, and/or availability.

Interfaces and threats can be identified using threat modeling (Domas & Merdinger, 2017). Threat modeling helps organizations to find security bugs early, understand security requirements, engineer and deliver better products, and address issues that other tools will not find (Shostack, 2014).

The CVSS calculator can be used to support vulnerability assessment ("Common Vulnerability Scoring System Version 3.0 Calculator," 2018). The calculator produces for each identified vulnerability a numerical score with a range between 0.0-10.0 reflecting its severity (i.e., none, low, medium, high, critical).

Following the initial risk identification phase, the manufacturer should perform security risk control activities for each identified risk and evaluate the overall residual security risk acceptability.

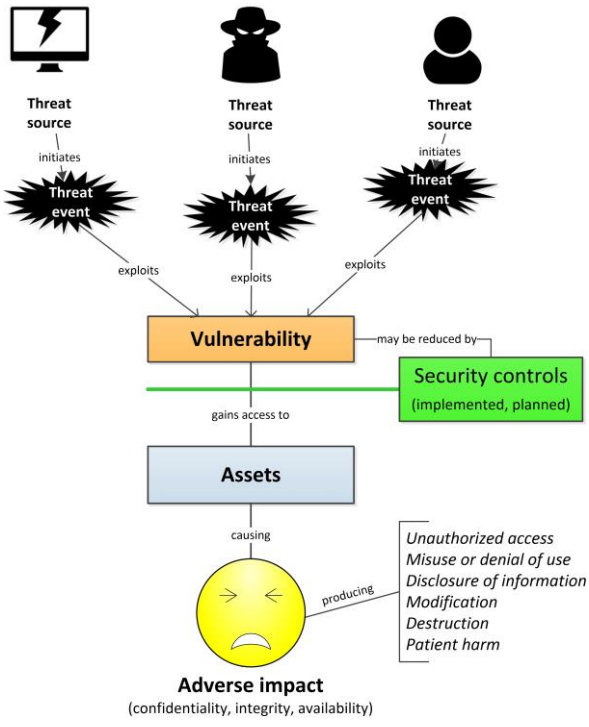


Figure 4. Illustrated causal chain of security threats Adapted from "Security risk analysis and management", by B. D. Jenkins, 1998, p. 4

2.3 Main Components

Our simplified cybersecurity model shown in Fig. 4 is based on the following three questions:

1. How to prevent cybersecurity incidents of medical devices in the first place?

2. How to detect cybersecurity vulnerabilities?
3. When a cybersecurity incident happens, how to respond to it?

The proposed model consists of the following three components: (1) prevention, (2) detection, and (3) incident response and recovery. Each component of the cybersecurity model has a corresponding set of security controls as further described in Section 2.4. Table 1 may be used to provide relevant informative references for the medical device industry.

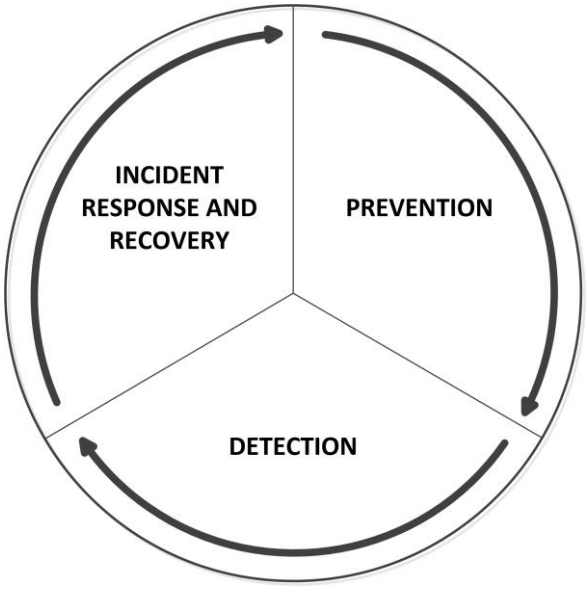


Figure 4. Generic cybersecurity model

Prevention is focused on designing medical devices from the start with cybersecurity in mind (i.e., security by design).

Detection means performing diverse types of security assessments to discover vulnerabilities that could be exploited and applying appropriate security controls to manage risks from cybersecurity threats that could impact the confidentiality, integrity, and/or availability of the medical device or the information processed by the device. It is important to note that cybersecurity needs to be assessed in the context of the larger system in which the medical device operates.

Incident response and recovery is needed to ensure that a medical device manufacturer has policies, procedures, and appropriate controls in place in case of a cybersecurity incident.

2.4 Security Controls

The FDA (2014) recommends developing a set of cybersecurity controls to assure medical device cybersecurity and to maintain medical device functionality and safety. The OWASP Secure Medical Device Deployment Standard (2017) can be used as comprehensive guide to the secure deployment of medical devices within a healthcare facility. This standard provides an overview of security controls that are divided into the following categories: purchasing controls, perimeter defences, network security controls, devices security controls, interface and central station security, security testing, and incident response.

Table 2 lists examples of security controls for each component of our cybersecurity model illustrated in Fig. 4. The table demonstrates that the highest number of security controls can be applied during prevention. More examples of security controls can be found in Annex E of AAMI TIR57 (2016).

Selection of appropriate security controls at various life cycle stages of a medical device depends on:

- Type of the medical device (e.g., device that contains software/firmware, device that contains programmable logic, software that is a medical device, mobile medical app, device that is considered part of an interoperable system, legacy device),
- Device classification,
- Intended use of the device,
- Operating environment in which the device is intended to be used,
- Intended users,
- User interaction with the device,
- Device's interaction with other devices on the network,
- Wired, remote, and wireless (e.g., Bluetooth, WiFi, wireless footswitch, Global System for Mobile Communications) interfaces,
- Communication protocols supported on internal and external interfaces,
- Technology (e.g., mobile, web, desktop, cloud computing, IoT),
- Used third-party components and open source software,
- External file inputs,
- Critical assets that need to be protected,
- Sensitivity levels of certain data (e.g., personally identifiable information, protected health information),
- Device's data flows,
- Data storage, etc.

Table 2. Examples of security controls

Component	Security control
Prevention	• Asset inventory • Code obfuscation • Conducting mock incidents • Cybersecurity policies and procedures • Data encryption technologies

Component	Security control
	• Data integrity controls (e.g., checksums, cryptographic checksums) • Database clusters • Default "deny" firewall policy • De-identification of patient data (e.g., anonymization, pseudonymization) • Established process to download and install security patches • Evaluation of cloud providers with respect to the security controls • Guidelines for secure development (e.g., avoiding exploitable code errors, validating data inputs before using or processing the data, storing local data securely, implementing access controls, etc.) • Instructions for the secure use of the device • Intrusion prevention systems • Malformed input (i.e., fuzz) testing • Network micro-segmentation • Operating system hardening • Partnerships with white-hat hackers and forensic experts to detect vulnerabilities that could be exploited • Physical locks on devices and their communication ports • Policies for classifying and categorizing all device data • Privileged user/account management (i.e., assigning roles using the principle of least privilege) • Publications describing how to avoid introducing common errors into the software that might become a vulnerability • Restricted software/firmware updates to authenticated code (e.g., code signing) • Sandboxing • Secure coding guidelines • Secure data transfer using encrypted connections (e.g., HTTPS, SSL, TLS, FTPS, etc.) • Security audits • Security awareness training for employees • Security code reviews • Security risk assessments • Self-descriptive user interface • Static binary and bytecode analysis • Static source code analysis • Threat intelligence • Threat modeling • User access controls (e.g., use of user ID and password, multi-factor authentication, account lockout after failed login attempts, automatic user logoff, changing default passwords at/prior to installation, password rules requiring use of strong passwords, lock screen function) • User authentication before permitting software/firmware updates • Version control systems • Vulnerability analysis
Detection	• Anti-virus software • Audit trails • Behavioral scanning • Endpoint protection tools (e.g., CrowdStrike Falcon®, Traps™) • Firewalls at the perimeter • Internal firewalls • Log monitoring • Malformed input (i.e., fuzz) testing • Malware testing • Network intrusion detection systems (e.g., Wireshark) • Port scanning tools (e.g., Nmap, Netcat) • Structured penetration testing • Vulnerability scanning tools (e.g., OpenVAS, Nexpose, Metasploit, Greenbone, Nessus) to scan for known vulnerabilities
Incident response and recovery	• Backup of device configuration • Cyber threat intelligence sharing via Information Sharing and Analysis Organizations (ISAOs) • Cybersecurity updates and patches • Data backup and restore • Established process to report detected cybersecurity incidents (e.g., coordinated vulnerability disclosure policy and practice) • Failsafe and recovery procedures • Incident response plan • Reverting to the previously installed version if the cybersecurity update fails

3 Integrating Cybersecurity into the Device Development Life Cycle

3.1 Challenges

The life cycle of medical devices involves design and development, design transfer, risk management, usability engineering, cybersecurity, clinical evaluation, servicing, decommissioning, disposal, and other processes. Cybersecurity shall be addressed throughout the whole lifecycle of a medical device.

According to a recent Deloitte's (2017) online poll, identifying and mitigating the risks of fielded and legacy connected devices presents the biggest challenge facing the medical device industry with respect to cybersecurity (30,1%). Additional challenges that connected medical devices presented to respondents included embedding vulnerability management into the design phase of medical devices (19,7%), monitoring and responding to cybersecurity incidents (19,5%), lack of collaboration on cyber threat management throughout connected medical device supply chain (17,9%), and meeting regulatory requirements (8,4%).

3.2 Integrating the Cybersecurity Process within a Quality Management System

According to the FDA guidance document (2016), cybersecurity shall be addressed in the following aspects of quality management systems: complaint handling, quality audit, corrective and preventive action, software validation and risk analysis, and servicing.

Integration of cybersecurity into the product development life cycle as part of a quality management system is not easy. The cybersecurity process spreads throughout the premarket and post-market phases of a medical device. The process is not isolated and the interfaces to the other processes within the organization and beyond it must therefore be identified and considered. Fig. 5 shows how the cybersecurity process is linked with other processes within a quality management system.

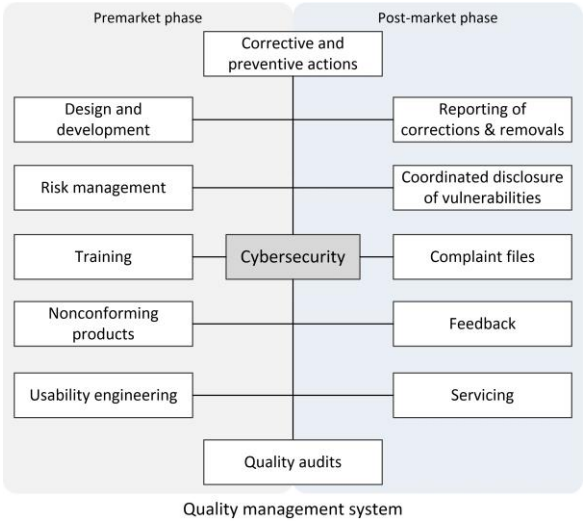


Figure 5. Position of the cybersecurity process within a quality management system

3.3 Implementation Considerations

Implementation of a cybersecurity program is a significant investment for any organization (Hrgarek Lechner, 2017). It requires support and endorsement from the top management to ensure availability of adequate resources and trained personnel for the cybersecurity process.

HITRUST ("Healthcare Sector Cybersecurity Implementation Guide," 2016) recommends the following seven steps to implement a cybersecurity framework in the healthcare sector: (1) prioritize and scope organizational components for framework adoption, (2) identify systems and existing risk management approaches within the scope, (3) create a desired risk management profile based on the organization's risk factors (Target Profile), (4) conduct a risk assessment, (5) create a current risk management profile based on assessment results (Current Profile), (6) develop a prioritized action plan of controls and mitigations (Action Plan), and (7) implement the Action Plan.

When developing a compliant cybersecurity process as an integral part of the quality management systems, the medical device manufacturers should consider the following:

- Gaining executive management support,
- Performing a gap analysis to identify discrepancies between the quality management system and the requirements set forth in the regulations and the organization's existing cybersecurity program (Hrgarek Lechner, 2017),
- Building an appropriate structure: e.g., a cross-functional cybersecurity team, a network of security champions ("Build a Network of Champions to Increase Security Awareness," 2017),
- Implementing an effective training and cybersecurity awareness program (Death, 2017),
- Working with external companies and security consultants providing security consulting services and performing penetration tests,
- Continuously monitoring regulatory requirements on medical device cybersecurity.

4 Conclusion

Historically seen, medical devices were designed and developed without design inputs related to cybersecurity. The next decade is likely to witness a considerable rise in cybersecurity threats of networked medical devices, wearable sensors, and other IoT devices. Since poor cybersecurity implementation may lead to data breach incidents and could have an adverse effect on patients, cybersecurity will play an increasingly significant role in operational safety and performance of medical devices. Cybersecurity also impacts business and top management support is needed to build a security culture within an organization.

Our generic cybersecurity model shows that an effective cybersecurity program is necessary at both

the premarket and post-market phases. Security must be built in from the start as part of the device development. A compliant cybersecurity process for medical devices requires addressing cybersecurity from design to obsolescence of medical devices being developed, marketed, and distributed. The selected security controls should mitigate cybersecurity risks early and prior to exploitation.

The cybersecurity process should be integrated into a quality management system considering the interfaces to the other processes within the organization and beyond it.

Disclaimer

The views and opinions expressed in this paper are those of the individual author and do not represent the approach, policy, or endorsement of the organization that is currently affiliated with the author.

Acknowledgments

The author would like to thank the anonymous reviewers for their valuable comments and suggestions which helped to improve the clarity and quality of this paper.

References

AAMI TIR57: Principles for medical device security—Risk management. (2016).

Alberts, C., Dorofee, A., Stevens, J., & Woody, C. (2003). Introduction to the OCTAVE® Approach. Retrieved from https://resources.sei.cmu.edu/asset_files/UsersGuide/2003_012_001_51556.pdf

Best Practices in Implementing the NIST Cybersecurity Framework. (2016). Retrieved from https://www.gartner.com/doc/3188133/best-practices-implementing-nist-cybersecurity

BS EN 45502-1: Implants for surgery – Active implantable medical devices. Part 1: General requirements for safety, marking and for information to be provided by the manufacturer. (2015).

Bug can cause deadly failures when anesthesia device is connected to cell phones. (2014). Retrieved from https://arstechnica.com/information-technology/2014/04/bug-can-cause-deadly-failures-when-anesthesia-device-is-connected-to-cell-phones/

Build a Network of Champions to Increase Security Awareness. (2017). Retrieved from https://www.gartner.com/smarterwithgartner/build-a-network-of-champions-to-increase-security-awareness/

Burns, A. J., Johnson, M. E., & Honeyman, P. (2016). A Brief Chronology of Medical Device Security. *Communications of the ACM*, 59(10), 66–72. doi:10.1145/2890488

Common Vulnerability Scoring System Version 3.0 Calculator. (2018). Retrieved from https://www.first.org/cvss/calculator/3.0

Death, D. (2017). *Information Security Handbook. Develop a threat model and incident response strategy to build a strong information security framework*. Birmingham: Packt Publishing.

Deloitte. (2017). Medical devices and the Internet of Things: A three-layer defense against cyber threats. Retrieved from https://de.slideshare.net/DeloitteUS/medical-devices-and-the-internet-of-things-a-threelayer-defense-against-cyber-threats

Domas, S., & Merdinger, S. (2017). Designing Robust Medical Devices that Are Ready for Enterprise Security Scanning. *Biomedical Instrumentation & Technology: Cyber Vigilance: Keeping Healthcare Technology Safe and Secure in a Connected World*, 51(6), 26–29. doi:10.2345/0899-8205-51.s6.26

FDA. (2014). Content of Premarket Submissions for Management of Cybersecurity in Medical Devices – Guidance for Industry and Food and Drug Administration Staff. Retrieved from https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm356190.pdf

FDA. (2016). Postmarket Management of Cybersecurity in Medical Devices – Guidance for Industry and Food and Drug Administration Staff. Retrieved from https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm482022.pdf

Framework for Improving Critical Infrastructure Cybersecurity Version 1.0. (2014). Retrieved from https://www.nist.gov/sites/default/files/documents/cyberframework/cybersecurity-framework-021214.pdf

Framework for Improving Critical Infrastructure Cybersecurity Version 1.1. (2018). Retrieved from https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.04162018.pdf

Fu, K., & Blum, J. (2013). Controlling for Cybersecurity Risks of Medical Device Software. *Communications of the ACM*, 56(10), 35–37. doi:10.1145/2508701

Health Insurance Portability and Accountability Act of 1996. (1996).

Healthcare Sector Cybersecurity Implementation Guide. (2016). Retrieved from https://www.us-cert.gov/sites/default/files/c3vp/framework_guidance/HPH_Framework_Implementation_Guidance.pdf

HIMSS/NEMA Standard HN 1-2013: Manufacturer Disclosure Statement for Medical Device Security. (2013).

Hrgarek Lechner, N. (2017). An Overview of Cybersecurity Regulations and Standards for Medical Device Software. In *Proceedings of the Central European Conference on Information and Intelligent Systems (CECIIS 2017)* (pp. 237–249). University of Zagreb, Faculty of Organization and Informatics Varaždin.

IEC 60601-1+AMD1: Medical electrical equipment – Part 1: General requirements for basic safety and essential performance. (2012).

IEC 62304+AMD1: Medical device software – Software life cycle processes. (2015).

IEC 82304-1: Health software – Part 1: General requirements for product safety. (2016).

IEC TR 80001-2-8: Application of risk management for IT-networks incorporating medical devices – Part 2-8: Application guidance – Guidance on standards for establishing the security capabilities identified in IEC TR 80001-2-2. (2016).

ISACA. (2016). Cybersecurity Fundamentals Glossary. Retrieved from https://www.isaca.org/Knowledge-Center/Documents/Glossary/Cybersecurity_Fundamentals_glossary.pdf

ISO/IEC 27000: Information technology – Security techniques – Information security management systems – Overview and vocabulary. (2018).

ISO/IEC 27001: Information technology – Security techniques – Information security management systems – Requirements. (2013).

ISO/IEC 27005: Information technology – Security techniques – Information security risk management. (2018).

Jenkins, B. D. (1998). *Security risk analysis and management* [White paper]. Retrieved from https://www.nr.no/~abie/RA_by_Jenkins.pdf

McHugh, M. (2015). Medical Device Software and Technology: the past, present and future. *BEAI Spectrum*, Spring, 28–32.

Medical Device Privacy Consortium. (2014). Security Risk Assessment Framework for Medical Devices.

MEDJACK: Hackers hijacking medical devices to create backdoors in hospital networks. (2015). Retrieved from https://www.computerworld.com/article/2932371/cybercrime-hacking/medjack-hackers-hijacking-medical-devices-to-create-backdoors-in-hospital-networks.html

NIST Special Publication 800-30 Revision 1: Guide for Conducting Risk Assessments. (2012). Retrieved from https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-30r1.pdf

NIST Special Publication 800-39: Managing Information Security Risk: Organization, Mission, and Information System View. (2011). Retrieved from https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-39.pdf

NIST Special Publication 800-53 Revision 4: Security and Privacy Controls for Federal Information Systems and Organizations. (2013). Retrieved from https://nvlpubs.nist.gov/nistpubs/specialpublications/nist.sp.800-53r4.pdf

OWASP. (2017). OWASP Secure Medical Device Deployment Standard Version 1.0. Retrieved from https://www.owasp.org/images/c/c3/SecureMedicalDeviceDeployment.pdf

PwC Health Research Institute. (2017). Top health industry issues of 2018: A year for resilience amid uncertainty. Retrieved from https://www.pwc.com/us/en/health-industries/assets/pwc-health-research-institute-top-health-industry-issues-of-2018-report.pdf

Sarig, I. (2012). Meet Embedded Software Quality Challenges in Medical Device Development. *MEDS Magazine*, Jan., 24–28.

Shostack, A. (2014). *Threat Modeling: Designing for Security*. Indianapolis: John Wiley & Sons.

Spremić, M., & Šimunic, A. (2018). Cyber Security Challenges in Digital Economy. *Proceedings of the World Congress on Engineering 2018 (WCE 2018) Vol I* (pp. 341–346). London.

UL 2900-1: UL Standard for Safety for Software Cybersecurity for Network-Connectable Products, Part 1: General Requirements. (2017).

UL 2900-2-1: UL Standard for Safety for Software Cybersecurity for Network-Connectable Products, Part 2-1: Particular Requirements for Network Connectable Components of Healthcare and Wellness Systems. (2017).

Williams, P. A. H., & Woodward, A. J. (2015). Cybersecurity vulnerabilities in medical devices: a complex environment and multifaceted problem. *Medical Devices: Evidence and Research*, 2015(8), 305–316. doi:10.2147/MDER.S50048

Software Engineering

Domain Specific Honeytokens Based on Natural Language Processing – A Conceptual Model

Tomislav Turek, Tonimir Kišasondi
Laboratory for Open Systems and Security
Faculty of Organization and Informatics
Pavlinska 2, 42000 Varaždin, Croatia
{tomislav.turek, tonimir.kisasondi}@foi.hr

Markus Schatten
Artificial Intelligence Laboratory
Faculty of Organization and Informatics
Pavlinska 2, 42000 Varaždin, Croatia
markus.schatten@foi.hr

Abstract. *This paper presents the idea and conceptual model for keyword modelling by using Natural Language Processing (NLP) in a specific domain. The paper shows that keywords for Honeytokens (HTs) derived from a specific website domain can be generated automatically by extracting concepts of interest from security related or other domain specific texts. A conceptual model for generating such tokens is presented and guidelines for implementation are given. It is argued that such domain specific HTs are a better form of deception technology, that provides a harder challenge for detection from automated attacks and thus improves early detection and incident response procedures in modern complex systems.*

Keywords. honeytokens, honeynets, information security, natural language processing, domain modelling, keyword generation

1 Introduction

Deception technology is designed to detect and prevent an attack on a particular network or an application by deceiving the attacker and notifying the defenders that the attack has happened or is currently taking place. Deception technologies are useful when we want to detect that an attack or active reconnaissance is happening against a system. This helps incident responders and local CERT staff to screen a possible incident or to detect that a future incident might happen. A basic concept in deception technologies is a honeypot system which is a system that doesn't contain any real business data and whose main purpose is to collect intelligence and detect intrusions (Spitzner, 2003a). At the same time similar concepts such as honeytokens, honeynets and honeyfarms emerged that used the honeypot concept as a foundation. After (Spitzner, 2003b) defined HTs as a honeypot that is not a computer, there were a lot of different use cases for HTs in production systems and the problem of their generation was also researched by different scientists.

Published work by (Bercovitch et al., 2011) on an automated HTs generator called *HoneyGen* conducts data mining and extraction of characteristics and properties from real data items in order to generate an artificial relational database based on extracted rules. This method can generate artificial data items that can be planted into production resources in order to detect the unauthorized use of information. In a more specific use case, (Erguler, 2016) proposed a honeyword generation method by using existing user passwords as a solution for detecting password disclosures which was later improved by (Akshaya and Dhanabal, 2017) where graphical passwords were used to mitigate the ethic issues on using users real credentials. The research on generating honeywords is specifically used in securing passwords against disclosure and as an alarm system in case the honeyword is used as a users password.

Shown related work regarding HTs is used in very specific use cases. This means that generated HTs cannot be used in different situations other than for which they were designed. Main issue in HT generation is that HTs can be used in multiple places and, at this moment, to fully utilize them in the whole system, we need to use different methods of generation for different parts of the system.

In order to automate the implementations of HTs into applications and systems, we are introducing a conceptual model for generating HTs via NLP methods by using domain related content to enable the generation of HTs of any type with high probability of luring attackers and automated scanners but low probability of detecting which of the elements is an actual HT. By solving the problem of generating HTs that cannot be distinguished between real elements of the system, we would be one step closer to automate the process of implementing intrusion detection and prevention into an application or a system and raise the security level of every system by default along with escaping the cumbersome actions of manually implementing the same security defenses for each system separately.

The rest of this paper is organized as follows: firstly in section 2 we provide an introduction into HTs. Afterwards in section 3 we give an overview of NLP techniques related to the research at hand. In section 4 we present the developed conceptual model that can enable

us to use NLP techniques for generating HTs. Then we discuss the approach in section 5 and draw our conclusions in 6.

2 Honeytokens

The first usage of deception technology was described by (Stoll, 1988) where Clifford Stoll created an alert system and decoy files to gather information and identify a persistent hacker that accessed Lawrence Berkeley Laboratory systems. Afterwards, the concept of a honeypot emerged, which are specialized computers in a network that are not related to production systems. Honeypots are security components in a system designed to serve multiple purposes such as (1) luring attackers in order to alarm the administrators, (2) learning more about attackers behaviour and tactics, (3) stopping an attack or (4) recognizing new threats that are in the wild but still unknown to the public (Spitzner, 2003a). From a defensive perspective, honeypots can extend the knowledge of methods attackers use, and serve intelligence as a foundation for setting up defenses on production systems. Additionally, their implementation allows for automating the production system breach recovery process in case one of the production system gets attacked and infected prior to gathering intelligence via honeypots (I. Kim and M. Kim, 2012).

Spitzner (Spitzner, 2003b) also described a HT concept where he defined HTs as an implementation of a honeypot that is not a computer system. For example HTs can be anything from whole documents, URL parameters, database records, a sentence or a word in a legitimate business document. Basic usage for HTs is the same as for honeypots. You can use HTs as a decoy URL parameter to recognize that a service is being attacked and raise an alarm in the system, you can use them to analyze attackers tactics and methods, list a unique fake database record along with legitimate records to be able to recognize data leaks or disclosures, insert a unique fake sentence that does not change the meaning of a text to be able to recognize if a user has disclosed or leaked information or trace which profile was breached to steal data.

The definition of HTs for a specific system, can be hard and cumbersome manual work, since for every given application domain (AD) a new set of tokens has to be established and positioned in adequate places to serve the outlined purpose. In the following chapters we will try to address the first part of this process: how to generate honeytokens by using NLP techniques.

3 Natural Language Processing

NLP is a field of research that deals with various computational methods with the main objective how to make sense out of human language in order to be able to use it for various applications (Chowdhury, 2003). In particular, NLP uses various approaches for synonym and concept extraction, taxonomic or non-taxonomic relationship extraction, or even ontology construction, by building upon well established symbolic and statistical methods including but not limited to Lexico-Syntactic Patterns (LSPs), Hidden Markov Models (HMMs), Neural Networks (NNs), Support Vector Models (SVMs), Conditional Random Fields (CRFs), Compound Noun Information (CNI), various clustering techniques, co-occuring information, association rule mining, dependency triples, and of course machine– and more recently deep–learning (Liu et al., 2011).

One special technique we would like to outline here is the technique of concept/keyword extraction. In this technique, most relevant keywords from a given text are extracted, usually using statistical techniques with a few additions from lexical analysis (e.g. stop word, common word elimination, text vectorization, stemming and similar) as has been described in (Schatten, Seva, et al., 2015; Voinea and Schatten, 2014).

Another important technique is synonym extraction which we deem especially useful for the research at hand. Wordnet (Miller, 1995) is a lexical database for the English language that formally models various semantic and conceptual relationships between English words. Some of these relationships include synonyms and can be used to extract synonymous concepts (as has been done in (Schatten, Magdalenic, et al., 2011; Schatten, Ševa, et al., 2015), albeit for a different purpose).

By combining these two techniques, keyword extraction (function ke) and synonim extraction (relation se), one is able to construct a set of relevant keywords K from a given domain document D, as well as a superset K^+ of concepts synonymous to concepts in K by applying:

$$\begin{aligned} K &= ke(D) \\ K^+ &= \{s \mid \forall k \in K \wedge (k,s) \in se\} \end{aligned}$$

4 A Conceptual Model

Most important part in HT generation is to achieve several characteristics in order for the HT to serve its purpose. We propose generation of HTs which are related to the security part of the system but at the same time are mixed with the domain of the system where it is applied. If HT is a generated keyword related to the information security domain, they would be much more attractive to exploit than the standard generated domain keyword. On the other hand, if the generated HT is related only to information security but not to the domain, it could be exposed as a HT.

Therefore, every HT needs to have several properties:

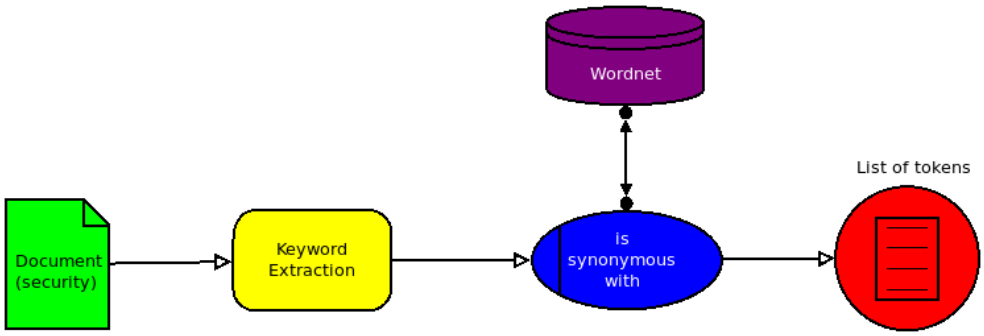


Figure 1: Flavor 1: Extraction of security related tokens

- HT should be attractive to exploit
- HT should be related to its domain
- HT should be adequate to its placement
- HT should not be easy to discover

In a HT environment, the established deception technology establishes one or more parameters (tokens) which in their nature are keywords related to the information security domain and/or the domain of the actual web system that is being protected with the HT. We propose two flavors of a method for generating such keyword parameters through the application of NLP techniques introduced in the previous section.

The first flavor is the straightforward extraction of tokens from the security AD, and generation of a synonymous token set. The next phase is usage of this set for establishing a HT. In particular, let D^{sec} be a document (or set of documents) dealing with a information security domain related to web systems. Further, let P^u be a set of usual parameters of the system being secured. The token set P^{sec} for the HT is then defined as:

$$P^{sec} = \{s \mid \forall k \in ke(D^{sec}) \wedge (k, s) \in se\}$$

Since P^{sec} and P^u might overlap, we define a mapping δ, that will exchange all keywords from P^u that are in P^{sec} with other, arbitrary words not in P^{sec}, therefore:

$$\begin{aligned} P^u_\delta &= \{\delta(k) \mid \forall k \in P^u \cap P^{sec}\} \\ &\cup \{k \mid \forall k \notin P^u \cap P^{sec}\} \end{aligned}$$

Figure 1 shows a graphical workflow of the proposed flavor.

The second flavor would use documents related to the actual AD of the web system being protected (these could be the actual web pages for this particular case), and try to find tokens that are either equal or synonymous to tokens extracted from the information security domain. This way, the set of tokens to be used would reflect the actual domain of the system. Let D^{dom} be a domain specific document (or set of documents), then the token set P^{dom} for the HT is defined as:

$$\begin{aligned} P^{dom} = \{s \mid & \\ & \forall \quad s \in ke(D^{dom}), \\ & \forall \quad s \in ke(D^{sec}) \\ & \wedge \quad (k, s) \in se\} \end{aligned}$$

Figure 2 shows the graphical work-flow for the second flavor.

5 Discussion

The proposed mechanism of automatically generating HTs via NLP would be the first step in automating the implementation of security defenses into an application or a system. In order for those security defenses to actually serve a purpose, they need to be generated in a way that the attacker or a scanner will not be able to distinguish between real data items or HTs and at the same time increase the chances of luring attackers into a trap. If we can generate HTs that meet these requirements, we will be able to increase the efficacy of intrusion detection and thus prevent possible attacks. Furthermore, we can use any system or application to gather intelligence on various tactics and methods that attackers use by implementing HTs as a security mechanism by default via automating the implementation of security defenses into systems.

To measure the effects of generated HTs, we can use different methods. For example, vulnerability scanners could be used as a tool to recognize vulnerabilities in the system by using different heuristics. Scanners can scan different elements of an application such as cookies and cookie flags, URL paths and parameters, HTTP headers, SSL certificates and similar. By using these elements as a placement for HTs, we could measure the scanners affinity towards certain parameters or data items and the probability of luring the scanners into actively exploiting the HT or recommending the HT as a potential attack vector. On the other hand, to measure the affinity of a human attacker towards certain elements of an application, we could conduct a case study in a Capture the Flag (CTF) competition where

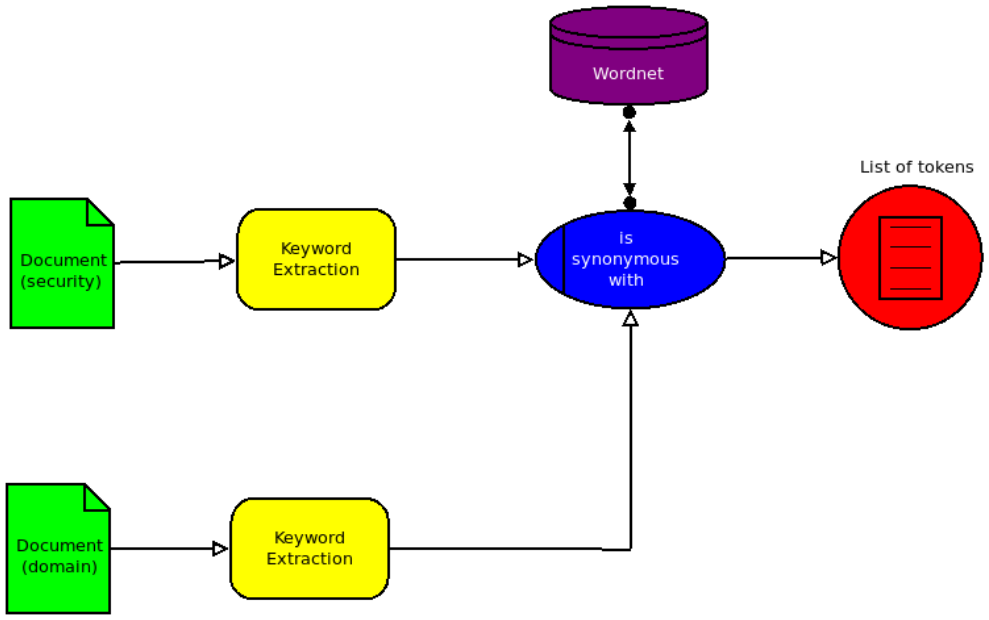


Figure 2: Flavor 2: Mapping of domain-specific to security related tokens

we could measure the affinity of the competitors towards certain elements in an application when trying to solve the challenge by analyzing their initial attack vector or how many vectors they tried to exploit before they started exploiting a HT. Additionally, we could measure if humans can recognize the difference between real data items and HTs by conducting a Turing test in order to see if there is a possibility to distinguish between items generated by a computer and items that are placed in the system for legitimate purposes.

6 Conclusion

The easiest way to increase the security of a system, without spending a lot of time in designing and implementing security defenses into an existing system, is to automate the process. One of the most convenient ways to implement intrusion detection is to use HTs which need to be generated in a way that look relevant to the domain and look attractive to exploit but at the same time do not raise suspicion that they are actually placed to lure the attacker. To solve this problem we propose two flavors for generating HTs by implementing extraction of security related tokens and mapping of domain-specific to security related token. To analyze the effectiveness of generated HTs we can use vulnerability scanners to analyze its affinity towards certain parameters or, to include the human element, conduct a case study by analyzing competitors behaviour during CTF competitions when solving challenges.

Our future research will be aimed at identifying possible interfaces to integrate the proposed approach into existing HT systems especially the placement of HTs in a given system. This task is non-trivial and depends on the system architecture, use-cases as well as best practices.

On the other hand, we will be testing different NLP techniques for generating tokens by using some of the outlined methods in order to find the most suitable ones for a wide range of systems. An open question is, would it be possible to generate phrases instead of, or in conjunction with, keywords?

In the end, we will try to implement an automated system described in the conceptual model that will make use of both placement and automated HTs generation techniques.

References

Akshaya, K. & Dhanabal, S. (2017). Achieving flatness from non-realistic honeywords. In *Innovations in information, embedded and communication systems (iciiecs), 2017 international conference on* (pp. 1–3). IEEE.

Bercovitch, M., Renford, M., Hasson, L., Shabtai, A., Rokach, L., & Elovici, Y. (2011). Honeygen: An automated honeytokens generator. In *Intelligence and security informatics (isi), 2011 ieee international conference on* (pp. 131–136). IEEE.

Chowdhury, G. G. (2003). Natural language processing. *Annual review of information science and technology*, *37*(1), 51–89.

Erguler, I. (2016). Achieving flatness: Selecting the honeywords from existing user passwords. *IEEE Transactions on Dependable and Secure Computing*, *13*(2), 284–295.

Kim, I. & Kim, M. (2012). Agent-based honeynet framework for protecting servers in campus networks. *IET Information Security*, *6*(3), 202–211.

Liu, K., Hogan, W. R., & Crowley, R. S. (2011). Natural language processing methods and systems for biomedical ontology learning. *Journal of biomedical informatics*, *44*(1), 163–179.

Miller, G. A. (1995). Wordnet: A lexical database for english. *Communications of the ACM*, *38*(11), 39–41.

Schatten, M., Magdalenic, I., & Vrdoljak, B. (2011). Towards ontology alignment of e-business standards using owl and f-logic. *International Journal of Metadata, Semantics and Ontologies*, *6*(3-4), 207–218.

Schatten, M., Seva, J., & Đuric, B. O. (2015). An introduction to social semantic web mining & big data analytics for political attitudes and mentalities research. *European Quarterly of Political Attitudes and Mentalities*, *4*(1), 40.

Schatten, M., Ševa, J., & Okreša-Đurić, B. (2015). Big data analytics and the social web: A tutorial for the social scientist. *European Quarterly of Political Attitudes and Mentalities*, *4*(3), 30–81.

Spitzner, L. (2003a). Honeypots: Catching the insider threat. In *Computer security applications conference, 2003. proceedings. 19th annual* (pp. 170–179). IEEE.

Spitzner, L. (2003b). *Honeytokens: The other honeypot*. SecurityFocus.

Stoll, C. (1988). Stalking the wily hacker. *Communications of the ACM*, *31*(5), 484–497.

Voinea, C. F. & Schatten, M. (2014). Recovering the past. eastern european web mining platforms for reconstructing political attitudes. In *European conference on political attitudes and mentalities ecpam*.

Strategic Planning and Decision Making

Multi-Task Learning for Ski Injury Predictions

Sandro Radovanović, Boris Delibašić, Milija Suknović
University of Belgrade – Faculty of Organizational Sciences
Jove Ilića 154, 11000 Belgrade, Serbia
{Sandro.radovanovic, boris.delibasic, milijas}@fon.bg.ac.rs

Abstract. *Predicting ski injuries is a very hard classification problem. This is due to the high class imbalance of injured vs. non-injured skiers and the lack of demographic information about skiers. Additional problems are the intrinsic properties of the ski lifts. Ski lifts differ in width, the difficulty degree, geographical position on the mountain etc. which results in different patterns for ski injury. In most researches, this information is not included. Aim of this paper is to develop multi-task classification models, which account for the uniqueness of ski lifts, taking into consideration information from other ski lifts. The proposed models were created on Mt. Kopaonik, Serbia ski resort and they show that ski injury in the following hour can be predicted with AUC ~0.64, or 3-4% better compared to the classical approaches.*

Keywords. Ski injury, logistic regression, multi-task learning

1 Introduction

Ski sports and leisure industry is considered as a big industry with steady 60 million average skier visits since 2002/2003 in the United States only. However, the United States is estimated to have 15% of worldwide skiers. Therefore, it is expected to have 400 million skier visits worldwide (The National Ski Areas Association, 2018). Additionally, every country with mountain regions tends to have ski resorts because they will be a major source of income and sustainability not only for the mountain region but for the whole country as well.

Although skiing is very popular, especially in the winter, the decision-making process in the ski industry is in most cases not data driven, and ski resorts often face problems with sustainability and profitability. This means that decisions are made based on views, opinions and experience of top management. Without adequate support of data-driven decision making it is nowadays not possible to reach the KPIs of an organization.

One way for improvement of the decision-making process in the ski industry is by inspecting the data about skiing behavior. This data is already available in most ski resorts. Namely, most ski resorts utilize RFID ski passes for entering ski lift gates. Based on this, a huge amount of data about skiers are generated which can be used for informed decision making.

One problem which is often analyzed in ski resorts is ski injury. This problem is considered as a highly challenging one (Ruedl et al., 2014; Bianchi et al., 2017) because incidence rate is very low. Namely, ski injuries are very rare events with 0.2% or fewer injuries per skier day. This number may seem low, but it is expected to have yearly 800,000 ski injuries worldwide, which is a major cost for insurance companies, and a global public health problem. Additionally, ski injuries come with a high cost, i.e. broken arm or leg, temporary or permanent movement disability and sometimes even death.

In this paper, we created a prediction model which predicts whether an injury will occur in the following hour on the ski lift on Mt. Kopaonik, Serbia. This research setup is important, as it is relevant to find-out real-time predictors of ski injury occurrence which could help in real-time prevention of injuries, and therefore reduce the occurrence of ski injuries. From a machine learning perspective, this problem presents a binary classification problem which we evaluated using the area under the ROC curve (AUC). An additional problem which is present in the data at hand is the fact that ski lifts are very different among themselves. Therefore, using all data for classification model could generate a too general model. On the other side, creating a classification model for each ski lift would require creating as many classification models as there are ski lifts (in this case 14). Additionally, we might not have enough data for some ski lifts as some ski lifts are less utilized, meaning that fewer data are available. As a solution, we propose multi-task classification (Pan & Yang, 2010) which generate multiple classification models at the same time, while exploiting similarities and differences between models, in this case, ski lifts. In other words, we want to extract and utilize shared information from all available ski lifts but also account for the differences due to the uniqueness of ski lifts. Since each ski lift is learned independently it is expected to have better performances compared to a model using all data. Additionally, since information is shared with other ski lifts it is expected to have better performance compared

to independent classification models for each ski lift. We will train three different multi-task logistic regression models. Namely, Lasso regularized multi-task logistic regression, Trace norm regularized multi-task logistic regression and $l_{2,1}$ norm regularized multi-task logistic regression.

The aim of this paper is a proposal of multi-task logistic regression model for prediction of ski injuries. We argue that data mining and machine learning techniques could be beneficial for the whole ski industry, which means that effects could be seen not only by the ski resort but also by the final users of the ski resort. Also, information about possible ski injury could be of great help especially for mountain rescue service which can be used to reduce the time needed for intervention, or even act preventively.

The remainder of the paper is structured as follows. Section 2 provides a literature review on ski injury predictions. Section 3 provides a methodology of the research providing a brief description of the data, multi-task learning, and experimental setup. Section 4 presents results and discussion of the results, while Section 5 concludes the paper.

2 Literature Review

Ski injuries are most often analyzed in small-scale, case-control studies. The goal of such analyses is to compare injured skiers population to a small subsample of the non-injured population in order to discover what the differences were between subpopulations. As a result, one could see odds ratios or risk ratios for different properties of skiing. Those properties could be physical, i.e. gender (Ruedl et al., 2016a) and age (Ruedl et al., 2016b, Chamarro & Fernández-Castro, 2009). Properties could refer to skiing behavior, i.e. the speed of skiing (Ruedl et al., 2016a) and skiing experience (Hume et al., 2015). Also, properties could be ski lift related, i.e. quality of ski lifts (Chamarro & Fernández-Castro, 2009) and snow condition (Ruedl et al., 2013). Finally, properties could represent weather (Hume et al., 2015).

Although information about odds ratio or risk ratio are of interest and could be useful for injury prevention and education they could hardly be used in real-time analysis and decision making. To the best of our knowledge, the first data mining model for ski injury prediction is presented in (Bohanec & Delibašić, 2015). In that paper Decision Expert (DEX) model was combined with data mining model to predict global daily prediction of ski injuries. Namely, it combined domain knowledge with data to improve predictive performance. Although it dealt with regression problem remark that combination of expert modeling and data mining does improve the predictive performance of the learning algorithm. Further improvement using expert knowledge and data mining is presented in (Delibašić et al., 2018a), with a framework that allows extension of logistic regression models with DEX hierarchies. The framework is based on a stacking like approach to logistic regression. The proposed framework can be seen as a feature extraction model (resembling neural networks) where DEX model provides the structure. It has been shown that utilizing DEX models in a hierarchical manner in combination with the logistic regression improved performance of the predictive model. Both abovementioned papers present an introduction of knowledge into data mining and machine learning algorithms. The goal of knowledge is to enhance evidence available in the data. In this paper, we would like to enhance data mining models by sharing knowledge between classification models. This can be done using multi-task learning.

Another example of the application of data mining models for prediction of ski injuries is presented in (Delibašić et al., 2017a). The idea of the paper was that classical analysis is not suitable and that interaction of attributes are of great importance for the prediction model, namely for prediction of ski injuries. Therefore, logistic regression models are not suitable because they assume a linear dependency between attribute and ski injury. Therefore, the CHAID algorithm was used. It has been shown that performance of the data mining models was significantly better compared to univariate analysis and that performances of logistic regression and CHAID algorithm were comparable. However, CHAID decision tree model could be useful for identification of injury risk subpopulations because decision tree is much more interpretable compared to logistic regression.

Prediction whether a ski injury will occur or not is an information of high value. However, mountain ski rescue service could use information about an injury, i.e. what type of injury and what part of the body was injured in order to better allocate resources before the intervention. This further complicates classification model because instead of one label (whether an injury occurred or not) multiple are available (a type of the injury and part of the body). Therefore, the multi-label prediction must be applied. This is presented in (Radovanović et al., 2018). It has been shown that the performance of multi-label models could be utilized. However, some types of injuries and some part of the body are hard for prediction models. Namely, those are types of injuries and part of the body which are seldom injured.

Another interesting approach to ski injuries can be found in (Delibašić et al., 2017b). Instead of using classification models to predict whether an injury will occur or not we might use different data mining methodology which will return the same output (whether an injury will occur or not), namely recommender systems. Although this seems less intuitive it has been shown that predictive performance of recommender systems was comparable or better than data mining and machine learning algorithms.

Analysis of ski industry is not limited to ski injuries. One can find analysis and recommendations

for ski lift pricing tickets, clustering of skiers etc. Interested readers in the analysis of the ski industry related researches are referred to (Delibašić et al., 2018a).

Authors observed that missing part of the ski injury prediction models is the account for unique characteristics of the ski lift, but also share information between ski lifts. Namely, every above-mentioned paper creates a classification model using data from all ski lifts. Because of that, we propose a classification model which is able to use information from other ski lift if available or data is missing, but also utilize the uniqueness of ski lift.

3 Methodology

In this section, we will present data, short explanation and motivation for multi-task and logistic regression, and experimental setup.

3.1 Data

The data used in this research was obtained from Mt. Kopaonik, Serbia ski resort. Mt. Kopaonik has 20 ski lifts with different degrees of difficulty, from which 14 are used in this research due to the lack of observed injury data on a specific lift on a specific hour. Data include all ski lift gate entrances from season 2005/2006 to season 2011/2012. Ski lift entrances are obtained using RFID checkouts of ski tickets which are needed in order to start skiing on the lift. In order to prevent confusion, term *skier* is used for all ski participants, i.e. skiers, snowboarders. Ski injuries are available in the other data source, namely, ski mountain rescue service database. These data sources are joined using ski ticket number and ski lift gate checkout. Finally, data about the weather is obtained from Republic hydro-meteorological service of Serbia. The whole dataset has over 20,000,000 observations.

The goal of the paper is to create a prediction model for ski injuries. Namely, we want to predict whether an injury will occur in the following hour on a ski lift. Because of that, data was aggregated on the hour level. Due to aggregation of the data instead of 20,000,000 observations we have 44,941 observations across 14 ski lifts and for each ski lift we have 19 input attributes, and one output attribute (label) which present binary signal whether an injury occurred in the following hour or not. Input attributes can be roughly divided into three categories. First set of attribute present ski lift crowd. Attributes which represent this group are the *hour* of observation, *number of ski lift checkouts* in one hour and *number of skiers* in one hour. The second group would present skier behavior on a ski lift. Attributes in this group are *average time on tracks* skier spent on ski resort up to that hour, *number of local maxima in average time on tracks* (Delibašić et al., 2017a), *average vertical distance* skier spent on ski resort up to that hour, *number of local maxima for vertical distance*, *average number of lifts* skier skied up to that hour, *number of local maxima of number of lifts*, *number of distinct lifts* skier skied up to that hour and *number of local maxima of distinct lifts*. The third group represents weather attributes. Those are *temperature*, *dew point*, *humidity*, *wind speed*, *visibility*, *fog*, *rain*, and *snow*.

It is worth to mention that class imbalance is present in data. This means that the majority of observation are non-injured. The overall class imbalance is 3.73%.

3.2 Multi-task logistic regression

Many data mining and machine learning applications are related to classification tasks. This means that same input attributes and same output attributes are used but with different observations. One instance of observation is called task. Most often task are related between themselves. The simplest approach is to solve these tasks independently, ignoring the relations between them. However, we would like to utilize relations and create a classification model simultaneously. This is called multi-task model learning. Difference between traditional and multi-task learning is presented in Figure 1.

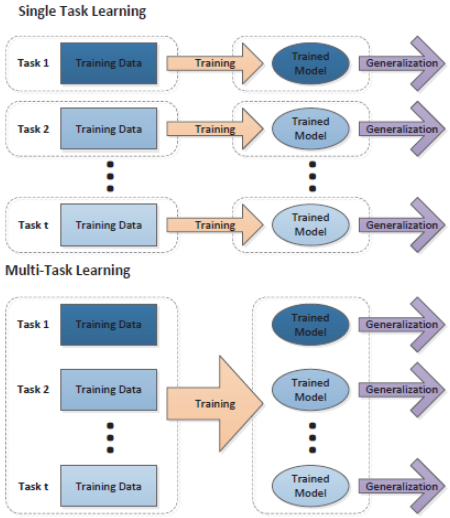


Figure 1. Illustration of traditional data mining model learning and multi-task model learning (Zhou et al., 2011)

A common mathematical model in data mining and machine learning is to minimize loss function and a regularization term. Namely, we want to:

$$\min_{W} L(W) + \lambda(W) \qquad (1)$$

where W is the parameter to be estimated from the data, namely coefficients of logistic regression, $L(W)$ empirical loss from the data, namely logistic loss and $\lambda(W)$ regularization term with is used to restrict overfitting and in case of multi-task learning share knowledge between classification models. In this paper we will utilize three multi-task approaches.

The first multi-task approach is called Lasso regularized multi-task learning. Lasso regularization is one of the most popular method for reducing generalized linear model complexity (Tibshirani, 2011). Complexity is reduced by forcing some coefficients of logistic regression to be zero. In this setting zero coefficients will be distributed across the tasks. Namely, some tasks might have several zero coefficients while other tasks would have none. It is easily extended into multi-task formulation by providing a matrix of logistic regression coefficients instead of a vector (Zhou et al., 2011) providing the following mathematical model.

$$\min_{W,\ c} \sum_{i=1}^{t} \sum_{j=1}^{n_i} \log(1 + \exp(-Y_{i,j}(W_j^T X_{i,j} + c_i)) + \lambda ||W||_1 \qquad (2)$$

where $X_{i,j}$ present observation j of the task i, $Y_{i,j}$ output attribute j of the task i, W_j and c_i represent the logistic regression models. Parameter λ controls the strength of the regularization.

The second approach used in this paper is called trace norm regularized multi-task learning (Grave et al., 2011). The mathematical model can be presented as:

$$\min_{W,\ c} \sum_{i=1}^{t} \sum_{j=1}^{n_i} \log(1 + \exp(-Y_{i,j}(W_j^T X_{i,j} + c_i)) + \lambda ||W||_* \qquad (3)$$

where $Y_{i,j}$ present observation j of output attribute for task i, $X_{i,j}$ input attributes vector of a j-th row of task i, while W_j and c_i represent the logistic regression models. Finally, the trace norm is presented as $||W||_* = \sum_i \sigma_i(W)$. Parameter λ controls the strength of the regularization. This means that trace norm tries to capture the relationship between tasks by assuming that every classification model share a common low-dimensional subspace.

Finally, we utilized $l_{2,1}$ norm regularization. This regularization forces all tasks to share a common set of features (Nie et al., 2010). This can be interpreted that each task would try to reduce same set of coefficients to be zero. In order to get the following properties, one needs to solve the following mathematical model.

$$\min_{W,\ c} \sum_{i=1}^{t} \sum_{j=1}^{n_i} \log(1 + \exp(-Y_{i,j}(W_j^T X_{i,j} + c_i)) + \lambda ||W||_{2,1} \qquad (4)$$

where $Y_{i,j}$ present row j of output attribute for task i, $X_{i,j}$ input attributes vector of a j-th row of task i, while W_j and c_i represent the logistic regression models. Parameter λ controls the strength of the regularization.

Finding global minima of such models can be done in the same manner as in for simple logistic regression by using gradient descent. We expect that multi-task formulation would yield better performing models compared to baseline methods which will be explained in the Experimental setup section. Additionally, we expect Lasso logistic regression to be the best performing one because it allows different representation of logistic regression coefficients, i.e. one attribute is selected by one ski lift model, but not by the other ones. Also, we expect that $l_{2,1}$ norm obtain results better than baselines but lower compared to other multi-task since it forces all models to select the same attributes for predictions, i.e. it forces every ski lift to use same attributes.

Finally, each of the presented models needs to provide a probability of injury occurrence. Since every task is trained at the same time, we must update formula for obtaining log-odds. Therefore, we would use:

$$\log\left(\frac{p}{1-p}\right) = X_t * W_t^T \qquad (5)$$

This will give us for each task t a vector of predictions which can be converted to probabilities using the sigmoid function.

3.3 Experimental setup

In order to test our hypothesis that sharing knowledge between classification models using multi-task logistic regression would yield better predictive performance compared to single classification model per ski lift and model which uses all the data, we will use split validation of proposed models. Namely, we will use random 70% for model training, and the remaining 30% for model testing. The procedure is repeated 10 times and the average value of performance measure is reported alongside the standard deviation.

We will train two baseline models. Those are single model per ski lift and model using all available data. We expect that model which uses only data from that ski lift would obtain good performances only for those ski lifts which have a high number of examples and a high number of injuries while performing badly for ski lifts with a small number of observations and huge class imbalance. Also, we expect the model which utilizes all available data to be too general.

As a performance metric, we selected the area under the ROC curve (AUC) since it is a common binary classification measure. Additionally, AUC is a decision threshold independent measure which means that the value of AUC present the overall goodness of the model. AUC is calculated by calculating the true positive rate and false positive rate for every possible decision threshold available in the data and calculating the area under the curve which is created by those two values. However, it can be calculated more easily using Mann-Whitney U test. AUC ranges from 0 to 1, where 1 present perfect classifier, while the value of 0.5 present classifier which is equal to the random classifier.

An additional challenge is a selection of the regularization parameter λ. In order to get a best possible estimate of the parameter λ, we performed 10-fold cross validation on a specified vector of possible values of λ. Vector of possible values is implemented

such that biggest value is smallest possible λ which yield intercept model only (all coefficient of logistic regression are equal to zero). Value is reduced by ~8 until λ reaches zero (James et al., 2013). Value λ which obtained the best AUC on inner 10 fold cross validation is then selected. One additional challenge comes from multiple tasks. Since we have 14 models we would obtain 14 AUCs. We will present the micro and macro values of AUCs. Micro value of AUC would account for a number of observations in the dataset by presenting weighted average, while the macro value of AUC would take average values of AUCs regardless of the number of observations in each task. In process of inner cross-validation, we need to select what will be optimized. We will have two results, one with optimized macro AUC and other with optimized micro AUC.

4 Results and Discussion

The experimental results are shown in Table 1 and Table 2. We report AUCs for baseline methods, which are all data logistic regression (AD-LR) and a logistic regression model for each ski lift which we can call task independent logistic regression (TI-LR), and also for multi-task models. Those are lasso logistic regression (MT-LN-LR), trace norm logistic regression (MT-TN-LR) and $l_{2,1}$ norm logistic regression (MT-L21-LR). Since the experiment is repeated 10 times we report the average value of AUC on the test set with the standard deviation, but also the lowest value and the highest value.

Table 1. Macro AUCs of the ski injury prediction models

Method	Average +/- S.D.	Min	Max
AD-LR	0.601 +/- 0.008	0.587	0.611
TI-LR	0.612 +/- 0.023	0.579	0.647
MT-LN-LR	**0.644 +/- 0.006**	**0.632**	0.657
MT-TN-LR	0.643 +/- 0.011	0.630	**0.661**
MT-L21-LR	0.641 +/- 0.010	0.627	0.655

As we can observe from Table 1 the best performing algorithm is multi-task Lasso regularized logistic regression with AUC 0.644, but other two multi-task algorithms performed similarly with 0.643 and 0.641 for Trace norm and $l_{2,1}$ norm, respectively. As expected, logistic regression using all data and task independent logistic regressions were 3% or 4% worse compared to multi-task algorithms. This is an indicator that sharing knowledge between tasks improve the performance of the algorithm. There are many examples where the similar finding is found, i.e. medicine (Zhou et al., 2012) and protein-protein interactions (Kshirsagar et al., 2013).

Table 2. Micro AUCs of the ski injury prediction models

Method	Average +/- S.D.	Min	Max
AD-LR	0.607 +/- 0.010	0.590	0.618
TI-LR	0.615 +/- 0.013	0.597	0.637
MT-LN-LR	0.642 +/- 0.011	**0.626**	0.658
MT-TN-LR	0.644 +/- 0.014	0.621	0.660
MT-L21-LR	**0.647 +/- 0.011**	0.625	**0.661**

Similar results are obtained using micro AUC (Table 2). The best performances are presented in bold letters. Performance is better on multi-task models compared to the baselines. However, Lasso norm logistic regression was the worst performing algorithm and $l_{2,1}$ norm was the best. Since, the difference is on the third decimal place we can state that performances are similar. Before going to further analysis we present information about ski lifts (Table 3).

Table 3. Information about ski lifts

Ski lift	# Observations	% Injuries
Centar	1735	4.84
Duboka 1	2436	4.68
Duboka 2	2881	6.42
Gobelja relej	1676	2.21
Gvozdac	1880	2.39
Karaman	2624	2.06
Karaman greben	3516	12.32
Kneževe bare	2341	1.84
Mali karaman	3244	8.69
Malo jezero	2884	1.70
Marine vode	2679	1.42
Mašinac	3216	1.74
Pančićev vrh	3110	3.95
Sunčana dolina	1763	1.87

Having abovementioned results in mind we would like to inspect performances on ski lifts and discuss effects of a number of observations and class imbalance on performances. This part of the analysis will be done on MT-LN-LR, MT-TN-LR, and MT-L21-LR methods because these methods performed well in the previous part of the experiment. The results are presented in Table 4 and 5 for macro AUC and micro AUC, respectively.

In Table 4 we can see average macro AUCs by ski lifts. Unfortunately, there aren't clear patterns about performances based on a number of observations and class imbalance.

Table 4. Macro AUC by ski lifts

Ski lift	MT-LN-LR	MT-TN-LR	MT-L21-LR
Centar	0.586	**0,609**	0,584
Duboka 1	0.664	0,664	**0,687**
Duboka 2	**0.663**	0,645	0,637
Gobelja relej	0.615	0,630	**0,655**
Gvozdac	0.651	**0,700**	0,674
Karaman	0.639	**0,649**	0,639
Karaman greben	**0.671**	0,658	0,664
Kneževe bare	0.648	**0,658**	0,634
Mali karaman	**0.714**	0,685	0,710
Malo jezero	**0.588**	0,559	0,544
Marine vode	0.613	0,613	**0,634**
Mašinac	0.587	0,584	**0,593**
Pančićev vrh	0.652	0,647	**0,663**
Sunčana dolina	0.681	0,675	**0,715**

Similar is observed in micro AUCs (Table 5).

Table 5. Micro AUC by ski lifts

Ski lift	MT-LN-LR	MT-TN-LR	MT-L21-LR
Centar	0.586	**0.609**	0.600
Duboka 1	0.664	0.664	**0.674**
Duboka 2	**0.644**	0.633	0.640
Gobelja relej	0.621	**0.665**	0.649
Gvozdac	0.659	0.654	**0.679**
Karaman	**0.670**	0.613	0.646
Karaman greben	0.655	**0.664**	0.659
Kneževe bare	**0.699**	0.684	0.648
Mali karaman	0.719	0.712	**0.728**
Malo jezero	**0.584**	0.558	0.551
Marine vode	0.618	0.624	**0.628**
Mašinac	0.572	0.581	**0.591**
Pančićev vrh	0.659	0.647	**0.660**
Sunčana dolina	0.658	**0.696**	0.685

However, we can observe that lower class imbalance (bigger percentage of ski injuries) do tend to have greater AUCs. Based on the performances we, unfortunately, cannot say which model is the best one.

Finally, we can present logistic regression coefficients for multi-task algorithms. Due to the big number of elements, these images are presented in the Appendix (Figure A1, Figure A2, and Figure A3). We can observe that Trace norm regularization included all available data and, therefore, every classification model has non-zero coefficients for every input attribute. Lasso norm regularization influenced the classification model to select only some of the input attributes. We can observe that the majority of features are zero, especially weather-related ones. A similar pattern is seen for $l_{2,1}$ norm. Since this norm forces classification models to select the same set of features we see that weather related features are omitted, i.e. coefficients are forces to be zero. This is an indicator that weather does not have an effect to ski injuries. More specifically, skiers do not tend to go skiing when the weather is bad. This finding is in accordance to the (Delibašić et al., 2017b). Also, we can observe that coefficient for average time on the track is negative. This is an indicator that the beginning of the skiing session can represent risky skiing behavior due to lack of preparation as noticed by (Hume et al., 2015).

5 Conclusion

This paper proposes a multi-task approach for ski injury prediction. This is motivated by the fact that ski lifts have different properties (width of adjacent slopes, difficulty of adjacent slopes etc.) and ski injury patterns may differ from ski lift to ski lift. However, knowledge about patterns should be exchanged from one ski lift to another ski lift, and even, between ski resorts. Therefore, we utilized three multi-task methods or regularizations. One of them is Lasso norm extension for multi-task setting, which forces coefficients of logistic regression to zero. This way algorithms are less prone to overfitting and therefore have greater generalizability. Another regularization used in this paper was Trace norm which tries to capture the relationship between tasks by assuming that every classification model shares a common low-dimensional subspace. Finally, $l_{2,1}$ norm forces tasks to select the same set of attributes. These methods were compared to baseline methods which are using all data for logistic regression and creating logistic regression for each ski lift independently.

Predictive performance of multi-task was better for 3% or 4% compared to baseline methods using macro AUC and micro AUC as a performance measure. Namely, AUCs of multi-task methods were ~0.64 for both macro and micro version of AUCs.

For a further research, we plan to employ class imbalance techniques in order to reduce class imbalance and to create decision tree multi-task algorithms.

References

Bianchi, G., Brügger, O., & Niemann, S. (2017). Skiing and snowboarding in Switzerland: Trends in injury and fatality rates over time. In *Snow Sports Trauma and Safety* (pp. 29-39). Springer, Cham.

Bohanec, M., & Delibašić, B. (2015, May). Data-mining and expert models for predicting injury risk in ski resorts. In *International Conference on*

Decision Support System Technology (pp. 46-60). Springer, Cham.

Chamarro, A., & Fernández-Castro, J. (2009). The perception of causes of accidents in mountain sports: a study based on the experiences of victims. *Accident Analysis & Prevention*, *41*(1), 197-201.

Delibašić B., Radovanović, S., Jovanović, M., & Suknović, M. (2018b, May). Improving decision making in ski resorts by analysing ski lift transportation. In *Proceedings of the XIII Balkan Conference on Operational Research – BALCOR 2018* (pp. 16). Belgrade, Serbia, 2018.

Delibašić, B., Radovanović, S., Jovanović, M., Bohanec, M., & Suknović, M. (2018a). Integrating knowledge from DEX hierarchies into a logistic regression stacking model for predicting ski injuries. *Journal of Decision Systems*, 1-8.

Delibašić, B., Radovanović, S., Jovanović, M., Obradović, Z., & Suknović, M. (2017a). Ski injury predictive analytics from massive ski lift transportation data. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology* (2017), 1754337117728600.

Delibašić, B., Radovanović, S., Jovanović, M., Vukićević, M., & Suknović, M. (2017b). An Investigation of Human Trajectories in Ski Resorts. In *International Conference on ICT Innovations* (pp. 130-139). Springer, Cham.

Grave, E., Obozinski, G. R., & Bach, F. R. (2011). Trace lasso: a trace norm regularization for correlated designs. In *Advances in Neural Information Processing Systems* (pp. 2187-2195).

Hume, P. A., Lorimer, A. V., Griffiths, P. C., Carlson, I., & Lamont, M. (2015). Recreational snow-sports injury risk factors and countermeasures: a meta-analysis review and Haddon matrix evaluation. *Sports Medicine*, *45*(8), 1175-1190.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning* (Vol. 112). New York: Springer.

Kshirsagar, M., Carbonell, J., & Klein-Seetharaman, J. (2013). Multitask learning for host–pathogen protein interactions. *Bioinformatics*, *29*(13), i217-i226.

Nie, F., Huang, H., Cai, X., & Ding, C. H. (2010). Efficient and robust feature selection via joint ℓ2, 1-norms minimization. In *Advances in Neural Information Processing Systems* (pp. 1813-1821).

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345-1359.

Radovanović, S., Delibašić, B., & Suknović, M. (2018). A multilabel prediction model for predicting part of the body and type of ski injury. In *Proceedings of the 4th International Conference on Decision Support System Technology – ICDSST 2018 & PROMETHEE DAYS 2018* (pp. 41). Heraklion, Greece.

Ruedl, G., Benedetto, K. P., Fink, C., Bauer, R., & Burtscher, M. (2016b). Factors associated with self-reported failure of binding to release among recreational skiers: an epidemiological study. *Current Issues in Sport Science* (CISS).

Ruedl, G., Helle, K., Tecklenburg, K., Schranz, A., Fink, C., & Burtscher, M. (2016a). Factors associated with self-reported failure of binding release among ACL injured male and female recreational skiers: a catalyst to change ISO binding standards?. *British Journal of Sports Medicine*, *50*(1), 37-40.

Ruedl, G., Kopp, M., Sommersacher, R., Woldrich, T., & Burtscher, M. (2013). Factors associated with injuries occurred on slope intersections and in snow parks compared to on-slope injuries. *Accident Analysis & Prevention*, *50*, 1221-1225.

Ruedl, G., Philippe, M., Sommersacher, R., Dünnwald, T., Kopp, M., & Burtscher, M. (2014). Current incidence of accidents on Austrian ski slopes. *Sportverletzung Sportschaden: Organ der Gesellschaft fur Orthopadisch-Traumatologische Sportmedizin*, *28*(4), 183-187.

The National Ski Areas Association. National skier/snowboarder visits, 1979–2016 (2018). Retrieved from http://www.nsaa.org/media/275017/1516_visits.pdf

Tibshirani, R. (2011). Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(3), 273-282.

Zhou, J., Chen, J., & Ye, J. (2011). MALSAR: Multi-task learning via structural regularization. *Arizona State University*, 2011.

Zhou, J., Liu, J., Narayan, V. A., & Ye, J. (2012, August). Modeling disease progression via fused sparse group lasso. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1095-1103). ACM.

Appendix

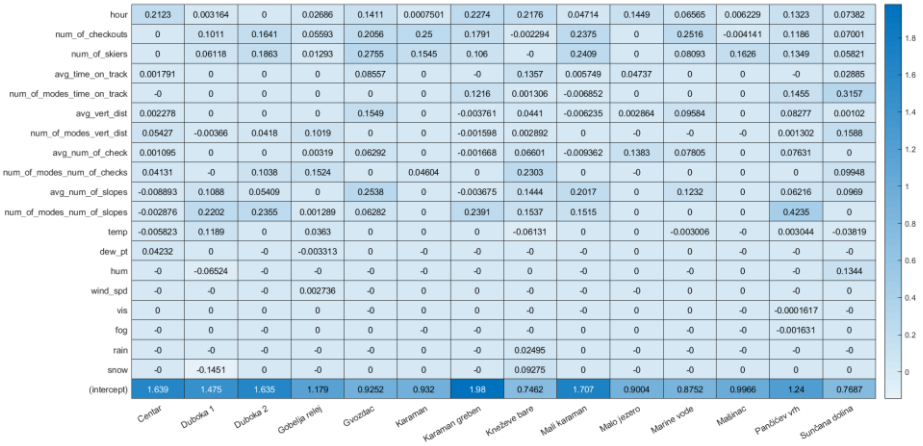


Figure A1. Coefficients of MT-LN-LR

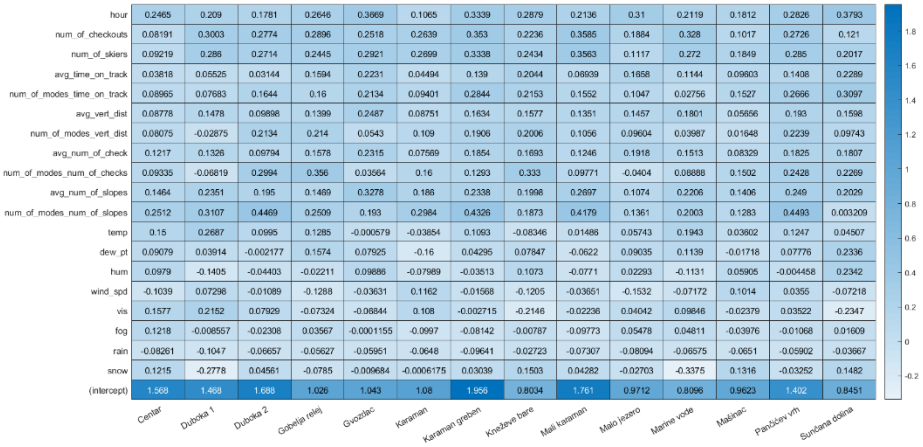


Figure A2. Coefficients of MT-TN-LR

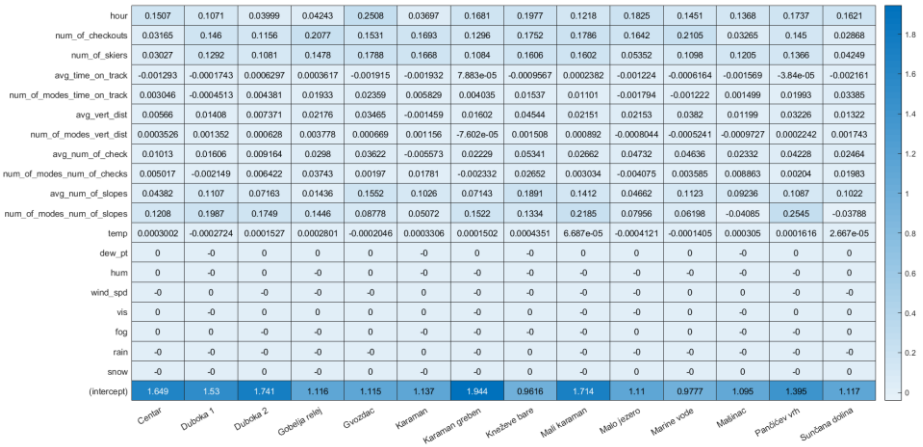


Figure A3. Coefficients of MT-L21-LR

Digital Maturity of Higher Education Institution: A Meta Model of the Analytical Network Process (ANP) and Decision EXpert (DEX)*

Valentina Đurek
Grad Zabok
ZIVTOV trg 10, 49210 Zabok
valentina@zabok.hr

Nikola Kadoić
Faculty of Organization and Informatics
Pavlinska 2, 42000 Varaždin
nkadoic@foi.hr

Željko Dobrović
The University of Applied Sciences Velika Gorica
Zagrebačka 5, Velika Gorica
zeljko.dobrovic@vvg.hr

Abstract. *Digital maturity of higher education institutions (HEI) becomes more and more important as the influence of ICT grows. In this paper, the methods Analytic Network Process (ANP) and Decision EXpert (DEX) are presented and demonstrated in the example of domains for digital maturity of HEIs. The ANP is a quantitative method, DEX is a qualitative method and digital maturity level (DML) is a two-component combination of results for the two methods' application. Additionally, the ERA model of those methods combined to calculate two-component DML of HEIs is designed, and its process will be specified and evaluated in future research.*

Keywords. ANP, DEX, metamodeling, digital maturity, framework, higher education institutions

1 Introduction

This research is a part of wider study that aims to create an instrument to measure overall DML of a certain HEI. Here, the DML is modelled as a two-component measure. One component calculates the DML by application of ANP, and the other determines the DML by application of DEX.

A high-quality higher education institution (HEI) and research excellence are not possible without information and communication technology (ICT). ICT could be a foundation for brand new achievements in analysis and cooperative atmosphere. The employment and integration of ICT in learning, teaching, research and technology transfer contribute to digital maturity of HEIs. The conception of digital maturity is critical for HEIs that, thanks to the fast development of ICTs, have a growing need to develop new teaching and business processes to realize changes in society, the market and organizations (Kampylis, Punie, & Devine, 2015; SCALE CCR, 2012).

The qualitative analysis of the literature analysed several maturity models with the application in education and 16 digital maturity frameworks in education. The results of this analysis are established such that there is no developed comprehensive Digital Maturity Framework for Higher Education Institutions (DMFHEI) and Instrument for the Assessment of Digital Maturity of Higher Education Institutions (IADMHEI) (Đurek, Begičević Ređep, & Divjak, 2017).

In the development of DMFHEI and IADMHEI, a complex methodology was applied, together with a set of methods, techniques and instruments, including qualitative analysis and comparison of comparable frameworks for describing digitally mature organizations with strategic documents at the national and international level and analysis of existing project documentation. DEMATEL (Decision Making Trial and Evaluation Laboratory) (Shih-Hsi Yin, 2012), the ANP Method (Analytic Network Process) (Divjak & Ređep, 2015), the Q-sorting method (Watts & Stenner, 2005), focus groups (Hines, 2000), composite index (Hines, 2000), questionnaires and interviews were also used during the development phase. The qualitative analysis method, Q-sorting method, focus group, and Delphi method—as well as the content validation ratio method (Lawshe, 1975)—were used by experts in the field of HEI and digital technologies to identify and match the domains and elements of the DMFHEI.

The DMFHEI identifies seven areas, within which there are 43 elements. Due to space limitations, we are not able to show the elements and descriptors of all 43 elements. The questionnaire and interview methodology was used in the description phase of the DMFHEI section, the IADMHEI section, and the revision of the first version of DMFHEI and IADMHEI based on qualitative analysis and focus groups. Developed DMFHEI is the basis for strategic planning and decision-making in the application of digital technologies at HEIs based on relevant domain elements' maturity (Đurek i ostali, 2017).

Since digital maturity is a multicomponent concept, it is possible to analyse it through multi-criteria decision-making methods. Multi-criteria decision analysis (MCDA) is a discipline concerned with

* This paper is published and available in Croatian language at: http://ceciis.foi.hr

solving decision problems that include presumably conflicting criteria. MDCA employs a variety of methods to create preference models by using information provided by the decision maker (Figueira, Greco, & Ehrgott, 2005). During the research, information can be given in different forms and representations. Converting representations from one form to another is usually very welcome, as it can bridge the gap between different methodological approaches and enrich the capabilities of individual methods. The DEMATEL method, ANP method, composite index and DEX method were used in the development of IADMHEI and the methodology of calculating the digital maturity of the HEI. DEMATEL was used to structure and determine the relationship between the elements. ANP was used to determine the weighting coefficients of domains and elements in IADMHEI, and the composite index and DEX method (Bohanec, Žnidaršič, Rajkovič, Bratko, & Zupan, 2013) was used for the integration of estimation and determination of overall maturity level and for the needs of ranking a HEI.

In this paper, we will present a meta model of a quantitative method for multi-criteria decision-making. The Analytical Network Process (ANP)—and the qualitative multi-criteria decision-making method, Decision EXpert (DEX)—were both applied in the assessment of the digital maturity of higher education institutions (HEIs).

This paper is divided into the following sections: quantitative method Analytic Network Process in Section 2; Qualitative method Decision Expert in Section 3; Two-component measure of digital maturity level of HEI in Section 4; and meta model of ANP-DEX integration in a two-component measure of digital maturity level of HEI in Section 5. The paper concludes with a discussion about data and future research.

2 Analytic Network Process (ANP)

The method specific for decision-making and human judgment is the multiple criteria decision-making method the Analytic Network Process (ANP). ANP can be described as a method which decomposes decision problems into a network consisting of smaller parts (Saaty, 1999).

In the ANP methodology, the structure of the decision problem is bestowed as a network that presents a system of parts vital for the matter in question. The network can be expanded by introducing the relationships between groups of elements and feedback. The standard of connections depends on the outlined degree of mutual impact of the elements on individual parts. ANP is the extension of Analytic Hierarchy Process (AHP) (Saaty, 1999) that enables networks to be created from the hierarchy as an end result of the gradual enlarge in the quantity of hierarchical connections. The pair comparisons are made in reference to all mixtures of mutual connections between the factors and their groups (Saaty, 1999). The AHP is the most-used multi-criteria decision-making method in HEIs (Kadoić, Begičević Ređep, & Divjak, 2016). It is based on pairwise comparisons of decision-making elements. In pairwise comparisons, the Saaty scale is used. The scale consists of nine degrees (1–9). Value 1 means that two elements in the pair are equally important. Value 3 means weak domination of one element over other. Value 5 means strong domination of one element over other. Value 7 means very strong domination of one element over other. Value 9 means absolute domination of one element over other (Begičević, 2008; Saaty, 2008).

When pairwise comparisons are completed, the inconsistency ratio is calculated. There are four basic steps in the AHP (Begičević, 2008; Saaty, 2008). The first step is the creation of hierarchy structure, followed by the completion of pairwise comparisons of elements from the same level in the structure with respect to superior elements in the hierarchy. The third step is calculating the priorities, and the final step requires performing sensitivity analysis.

Network design is one of the most important steps of the method because it forces the decision maker to conduct a fundamental analysis of the problem. The design of the network in a decision problem is a key factor in finding an appropriate solution. There are no clear directions in the literature on how to design the network (Saaty & Vargas, 2006). To conclude ANP method, several steps have to be followed (Saaty & Cillo, 2008):

1. In the first step, identification of the components, network elements and their relationships should be done. This step can be divided into three basic tasks: identification of the network elements that are decision criteria and alternatives; grouping the elements based on some common feature; and finally, analyzing the relationships between network elements. The third task can be supported by using the DEMATEL method.

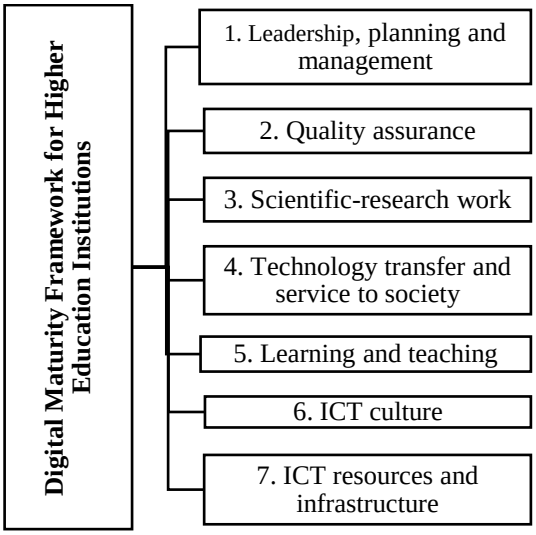


Figure 1. Network elements of DMFHEI

2. The second step consists of calculating the priorities between elements of the same cluster and determining which element is more influential and to what extent.

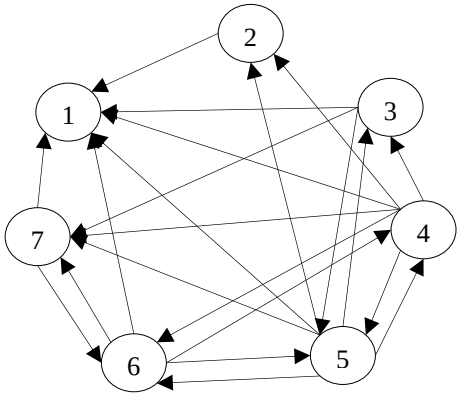


Figure 2. Relationships between network elements of DMFHEI (numbers 1-7 present the domains of DMFHEI in Figure 1.)

3. This step performs pairwise comparison matrices between clusters and calculates the priorities between clusters.
4. Next, it is necessary to do weighting of the unweighted supermatrix blocks using the priorities of each cluster, so that the resulting supermatrix, or *weighted supermatrix*, is column-stochastic.
5. The final step obtains the limit supermatrix where the elements of each column represent the final weightings of the different elements considered.

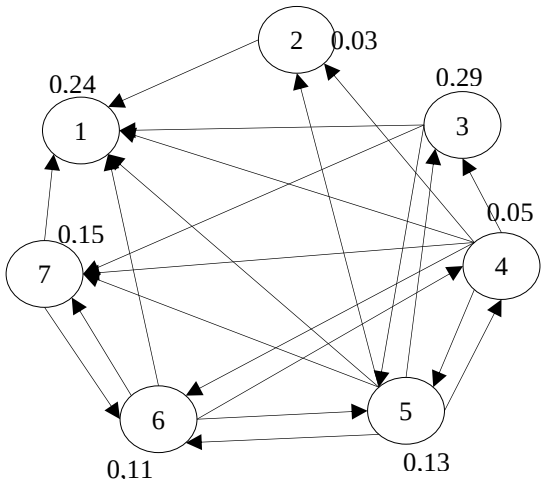


Figure 3. Weightings of the network elements of DMFHEI

The result of steps 2-5 are presented in Figure 3. The domain weights presented in the figure are only demonstrative.

Limitations of the ANP method include the high number of pairwise comparisons, lengthy implementation process, and high potential for misunderstanding some of the pairwise comparisons that have to be done. The complexity of the pairwise comparisons on the cluster level will also be decreased when integrating the ANP with the Decision-Making Trial and Evaluation Laboratory (DEMATEL) (Kadoić, Begičević Ređep, & Divjak, 2017; Đurek, Kadoić, & Begičević Ređep, 2018). This approach has been applied in the example of calculating the priorities of the DMFHEI (Figure 2). DEMATEL was used to identify the strongest relationships in the network that decreased the number of pairwise comparisons that had to be made. Inputs regarding the weights of influences between the domains, as well as the related pairwise comparisons, were obtained from managers of HEIs who participated in workshops organized under the scope of the Higher Decision project. The results are only demonstrative. For the complete research, level of elements (not only domains) also have to be included. Additionally, a higher number of respondents will be included in complete research.

3 Decision EXpert (DEX)

Decision EXpert (Bohanec i ostali, 2013) is a multi-criteria decision modelling method. The DEX method is a qualitative, multi-criteria decision analysis approach that provides support to decision makers in evaluating and choosing decision alternatives by using discrete attributes and rule-based utility functions (Mihelčić & Bohanec, 2017).

The DEX method consists of a set of decision alternatives that are fundamental for the evaluation and analysis. Alternatives are described with a set of variables called attributes, which represent some observed or evaluated property of alternatives (Bohanec i ostali, 2013).

DEX is a hierarchical method, meaning the attributes are organized in a hierarchy that represents a decomposition of the decision problem into sub-problems. The bottom-up direction denotes dependence, so that higher-level attributes depend on the lower-level, more elementary ones. The most elementary attributes—called basic attributes—appear as terminal nodes of the hierarchy and represent the basic observable characteristics of alternatives. Higher-level attributes, which depend on one or more lower-level ones, are called aggregated attributes that represent evaluations of alternatives. The topmost nodes (usually, there is only one such node) are called roots, and they represent the final evaluation(s) of alternatives (Mihelčić & Bohanec, 2017).

Furthermore, DEX is a qualitative method. While most of MCDM methods are quantitative and thus use numeric variables, qualitative methods use symbolic ones. In DEX, each attribute has a value scale that is represented with some ordinary word, such as ‘low’, ‘medium’, ‘high’ and ‘very high’. Scales are usually small, containing two to five values, and scales are also usually preferentially ordered. Attributes that have

preferentially ordered scales are called criteria (Figueira i ostali, 2005). Finally, DEX is a rule-based method. The bottom-up aggregation of alternatives' values is defined in terms of decision rules, which are specified by the decision maker.

In this paper we will present the DEX method in three steps using the two domains (Technology transfer and service to society and Scientific-research work) of DMFHEI due to space limitation (Bohanec et al., 2013):

1. Creating a hierarchical tree – the decision-making problem is modelled through a decision tree that can be interpreted in three ways: decomposition, dependence and aggregation. A qualitative scale is defined for each tree element. The scale consists of several elements. On the leaf level of tree, there are many criteria, which are being aggregated to one goal at the root of the tree. Hierarchical tree for case of domains of digital maturity is given in Figure 1. There are 7 elements on the leaf level that are aggregated into one element at the root. Like being said in ANP section, this is only a demonstrative example, because real leaves (elements of the maturity model) are not currently included in the research.
2. Decision rules – decision-making rules represent the basic mechanism of conclusion and decision-making in the DEX method (Mihelčić & Bohanec, 2017). At the elementary level, there are uniquely measurable criteria for each alternative to the scale of each criterion on the list. Presented case values that are used include: low and high (Table 1). Functions are defined at the level of aggregated criteria (low, medium, high) and at root level decisions that describe which value will take the criterion (on its scale) for each combination of criteria values from the level below (low, medium, high).

Table 1. Domain values

Domain	Value 1	Value 2
1.Leadership, planning and management	low	high
2.Quality assurance	low	high
3.Scientific-research work	low	high
4.Technology transfer and service to society	low	high
5.Learning and teaching	low	high
6.ICT culture	low	high
7.ICT resources and infrastructure	low	high

Figure 4 represents the decision rules of DMFHEI. Columns represent the DMFHEI domain, which are presented in Figure 1. In the presented case, it is necessary to make 128 decision rules.

	1	2	3	4	5	6	7	DMFHEI
1	low	low	low	low	low	low	low	low
2	low	low	low	low	low	low	high	low
3	low	low	low	low	low	high	low	low
4	low	low	low	low	low	high	high	low
5	low	low	low	low	high	low	low	low
6	low	low	low	low	high	low	high	low
7	low	low	low	low	high	high	low	low
8	low	low	low	low	high	high	high	medium
9	low	low	low	high	low	low	low	low
10	low	low	low	high	low	low	high	low
11	low	low	low	high	low	high	low	low
12	low	low	low	high	low	high	high	medium
13	low	low	low	high	high	low	low	low
14	low	low	low	high	high	low	high	medium
15	low	low	low	high	high	high	low	medium
16	low	low	low	high	high	high	high	>=medium

.
.

	1	2	3	4	5	6	7	DMFHEI
57	low	high	high	high	low	low	low	medium
58	low	high	high	high	low	low	high	>=medium
59	low	high	high	high	low	high	low	>=medium
60	low	high	high	high	low	high	high	>=medium
61	low	high	high	high	high	low	low	>=medium
62	low	high	high	high	high	low	high	>=medium
63	low	high	high	high	high	high	low	>=medium
64	low	high	high	high	high	high	high	>=medium
65	high	low	low	low	low	low	low	low
66	high	low	low	low	low	low	high	low
67	high	low	low	low	low	high	low	low
68	high	low	low	low	low	high	high	medium
69	high	low	low	low	high	low	low	low
70	high	low	low	low	high	low	high	medium
71	high	low	low	low	high	high	low	medium
72	high	low	low	low	high	high	high	>=medium

.
.

	1	2	3	4	5	6	7	DMFHEI
113	high	high	high	low	low	low	low	medium
114	high	high	high	low	low	low	high	>=medium
115	high	high	high	low	low	high	low	>=medium
116	high	high	high	low	low	high	high	>=medium
117	high	high	high	low	high	low	low	high
118	high	high	high	low	high	low	high	high
119	high	high	high	low	high	high	low	high
120	high	high	high	low	high	high	high	high
121	high	high	high	high	low	low	low	>=medium
122	high	high	high	high	low	low	high	>=medium
123	high	high	high	high	low	high	low	>=medium
124	high	high	high	high	low	high	high	>=medium
125	high	high	high	high	high	low	low	high
126	high	high	high	high	high	low	high	high
127	high	high	high	high	high	high	low	high
128	high	high	high	high	high	high	high	high

Figure 4. Decision rules for DMFHEI (domain level)

3. Once a hierarchical model has been created, and after the rules of decision are defined, the final step is evaluation of alternatives. Once the alternatives are evaluated, mutual comparison determines which is the best. The input values of the alternatives by individual criteria are determined by discretization of the continuous value space. This process can be done in following ways:

a. The first approach that can be used is the threshold. Values above the threshold assign the best qualitative value to criterion scale (high). The interval below the threshold is divided into several equal intervals (depending on the scale criteria) that frame the scale's criterion values. The threshold is often defined in a way that 1% or 10% of the best alternatives meet the highest criterion value. This is done for each criterion separately.
b. The second approach of discretization is based on the calculation of the percentile, and the values belonging to the 25 – 75 percentiles are classified as "middle" on the criterion scale.

The third step of the DEX method is presented in Table 2 in two examples of HEI.

Table 2. Evaluation of alternatives

Options	HEI1	HEI2
DMFHEI	medium	medium
Leadership, planning and management	high	high
Quality assurance	high	high
Scientific-research work	low	high
Technology transfer and service to society	high	high
Learning and teaching	low	high
ICT culture	low	low
ICT resources and infrastructure	high	high

Table 2 contains examples of two HEIs and their values on domain level. They are related to two of 128 decision rules from Figure 4, and in both cases the total DML is *medium*.

4 Two-component Measure of Digital Maturity Level of HEI

In the process of designing the framework and instrument for determining the digital maturity level (DML) of HEIs, it was decided that the digital maturity level would incorporate two components:

- The first component is quantitative, and it is a result of application of the ANP method.
- The second component is qualitative, and it is a result of application of the DEX method.

There are several reasons for DML to be a two-component measure:

- Some aspects of the digital maturity framework are qualitative, and some aspects are quantitative.
- Applying two methods acts as a sort of control, or at least a comparison mechanism, in determining the DML of HEI. For example, the ANP can result in a high quantitative value of DML. Then, if DEX offered a low qualitative value, further analysis would be mandatory.
- The two methods, ANP and DEX, have different aggregation mechanisms, and it is possible that when certain HEIs have a *very low* value on some element and others are *high*, then (1) quantitative DML values obtained by ANP will be *just a bit lower than the high* value, but (2) qualitative DML values obtained by DEX can be *low* because starting *very low* value on some element can overcome through hierarchy.
- ANP and DEX complement each other.

5 Meta model of ANP-DEX Integration in Case of Two-component Measure of Digital Maturity Level of HEI

The modelling paradigm is one of the most important concepts for realizing the enterprise-wide integration. The model is a simplification of the reality—a blueprint of a system. As the result of an abstraction process, the model reflects the general, essential and permanent features from the modelling target's view, and it serves as a formal specification to describe the functionality, structure, and/or behaviour of the system.

A good model includes elements that have broad effects and omits minor elements irrelevant to the given level of abstraction. As the reality is very complex, it may be described from different aspects—what we call "model views"—being semantically closed abstractions of a system. The highest level of the abstraction is the metamodeling level (Raffai, 2008).

Most generally, metamodeling is the analysis, construction and development of the frames, rules, constraints, models and theories applicable and useful for the modelling in a predefined class of problems. This concept is composed with the notions of the terms meta and modelling. Thus, metamodeling is the construction of a collection of concepts within a certain domain, a precise definition of the constructs and rules needed for creating semantic models. As a model is an abstraction of real world phenomena, a metamodeling is yet another abstraction, highlighting properties of the model itself in the form of an abstract language for defining different kinds of metadata.

Authors (Vangheluwe & de Lara, 2002) describes modelling as a complex systems of difficult task, with components and aspects whose structure as well as behaviour cannot be described in a single comprehensive formalism. The term metamodel, actually means "model modeling language". The "meta" prefix indicates again that it is a concept at a higher abstraction level than the modeling language itself. Metamodel can provide ways to describe

abstract syntax, specific syntax or semantics of a language.

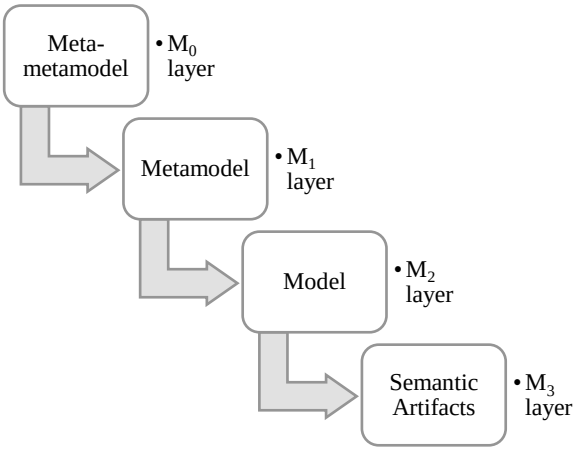


***Figure 5**: The four-layer metamodeling architecture*

The Object Management Group (OMG) is an independent organization that focuses on issuing standards specifically related to modeling of programs, business processes, information systems, etc. Their most famous and best-used specification is the Unified Modeling Language (UML) specification. In its concept of "Model-driven Architecture" (MDA) - one of the foundations and the very language of the UML - the OMG group proposes a four-grade meta-architecture (Figure 5) that can accommodate languages according to their own characteristics, provide these other languages with which they may be linked in some way (Karsai, Nordstrom, Ledeczi, & Sztipanovits, 2000), (J. Sprinkle, 2004), (Jonathan Sprinkle, Rumpe, Vangheluwe, & Karsai, 2010).

Level M_0 is the data itself. These can be objects in the program language, ranks in tables in the database, etc. Level M_1 contains a "model" of data at level M_0. In the case of object, programming languages at M_1 level are templates of objects, ie classes. If the database management system at level M_1 is a table definition, the data is stored (eg. SQL DDL commands). At this level, you can also find the entity-connection pattern of a system. At M_2 level, there are metamodels, ie languages that provide the model syntax. Finally, the M_3 level is the meta-metamodel level. What is important to note is that elements of higher metalevels provide building blocks for the definition of lower-level elements. When choosing a way to model a system, the most common choice is to select a language at M_2 level; this choice dictates how the system's systems look like which elements to contain, what limitations will be available and what purpose the model ultimately has. By selecting a language at level M_2, it can be started with M_1 modeling and final implementation at level M_0. OMG defined M_3 level language and called it Meta-Object Facility (MOF). The language is recursively descriptive; it can itself be described with the help of the elements it defines, thus solving the problem of the existence of higher metalevels. The MOF language represents the generic starting point for building blocks that can be used to define M_2-level languages.

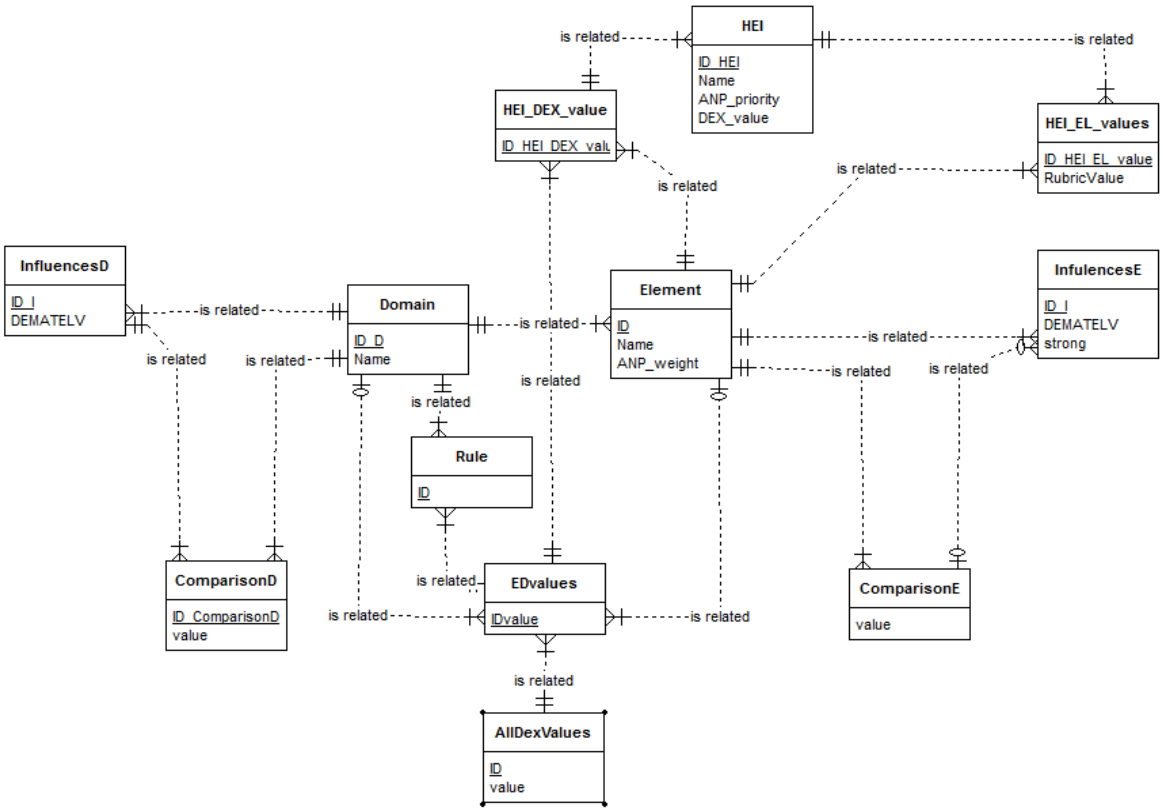


Figure 6. ERA model of ANP-DEX integration in case of two-component measure of DML of HEI

As a part of this paper, we prepared an ERA model of an ANP-DEX integration in a case of two-component measure of the digital maturity level of HEIs. The ERA model contains data about entities (tables), relationships between entities and attributes of entities.

The model is presented in Figure 6. The model consists of several entities:

1. Domain – contains data about domains from DMFHEI and their descriptions
2. Element – contains data about elements of all domains. After the ANP is applied, finale element weights will be written into table
3. Rule – contains all rules defined in the DEX method on the domain and root levels
4. EDvalues – contains possible values that can be achieved in certain element, domain and on the root level
5. AllDexValues – contains list of all possible values (elements, domains and root)
6. InfluencesE – contains data about the influences between criteria identified by using the DEMATEL scale
7. ComparisonE – contains data about the pairwise comparisons between elements with respect to other elements
8. ComparisonD – contains data about the influences between domains identified by using the DEMATEL scale
9. InfluencesD – contains data about the pairwise comparisons between domains with respect to other domains
10. HEI – contains data about HEIs and their priorities obtained by applying the ANP method, as well as results obtained by applying the DEX method (two-component result of DML)
11. HEI_Dex_value – contains data about values achieved by HEIs in terms of each element of the DEX hierarchy
12. HEI_El_value – contains data about all values that are achieved by HEI in terms of each ANP element. In this situation, the rubric will be used as a data collecting method. The rubric consists of five values per element described through statements previously evaluated by experts.

Entities 1-5 and 11 are related to the DEX. Entities 6-9 are related to the ANP. Entity 10 is related to the ANP-DEX integration. The entities are connected according to the relationships shown in Figure 6. In the phase of creating the software that will support ANP-DEX integration, it is possible that some changes in ERA have to be implemented. Besides possible changes, functions that support data collecting and ANP and DEX application have to be implemented as well.

6 Conclusion

This paper proposes a design for the two-component digital maturity level of certain HEIs. This is a different approach than currently known methods used in developing different frameworks and instruments related to concepts of *readiness* and *maturity*. This approach has some advantages, as described in the paper. Two-component design is modelled using the ERA model.

In future research, it is planned that experts from the field of digital maturity of HEIs and members of HEI management will give their inputs related to weights of influences between elements and domains, pairwise comparisons of elements and domains, elements and domains DEX values and design of decision-making rules in the DEX method. After that, element weights will be calculated.

In the evaluation phase, the IADMHEI will be applied at several HEIs in Croatia, and results will be compared to digital maturity of HEIs obtained by qualitative analysis. Ultimately, it will be possible to determine the two-component DML of Croatian HEIs.

Besides in the HEI digital maturity level area, this two-component approach can be applied in other contexts that are related to the investigating the readiness or maturity. Additionally, this approach can be generally applied in multi-criteria decision making.

Acknowledgments

The Croatian Science Foundation supported this work under project IP-2014-09-7854.

References

Begičević, N. (2008). *Višekriterijski modeli odlučivanja u strateškom plniranju uvođenja e-učenja*. *Doktorski rad*. Varaždin: University of Zagreb, Faculty of Organization and Informatics.

Bohanec, M., Žnidaršič, M., Rajkovič, V., Bratko, I., & Zupan, B. (2013). DEX Methodology: Three Decades of Qualitative Multi-Attribute Modeling. *Informatica*, *37*, 49–54. Retrieved from http://kt.ijs.si/

Divjak, B., & Rođep, N. B. (2015). Strategic Decision Making Cycle in Higher Eeducation: Case Study in E-learing. *International Conference e-Learning 2015 STRATEGIC*. Retrieved from https://files.eric.ed.gov/fulltext/ED562492.pdf

Đurek, V., Begičević Rođep, N., & Divjak, B. (2017). Digital Maturity Framework for Higher Education Institutions. In V. Strahonja & V. Kirinić (Eds.). *Conference Proceedings Central European Conference on Information and Intelligent Systems 2017* (pp. 99–106). Varaždin:

University of Zagreb, Faculty of Organization and Informatics.

Đurek, V., Kadoić, N., & Begičević Ređep, N. (2018). Assessing the Digital Maturity Level of Higher Education Institutions. *Proceedings of 41st Jubilee International Convention, MIPRO 2018*, 747–752. Varaždin: University of Zagreb, Faculty of Organization and Informatics.

Figueira, J., Greco, S., & Ehrgott, M. (2005). Multiple Criteria Decision Analysis: State of the Art Surveys. *International Series in Operations Research & Management Science*, *25*(3), 627–649.

Hines, T. (2000). An evaluation of two qualitative methods (focus group interviews and cognitive maps) for conducting research into entrepreneurial decision making. *Qualitative Market Research: An International Journal*, *3*(1), 7–16.

Kadoić, N., Begičević Ređep, N., & Divjak, B. (2016). E-learning decision making: methods and methodologies. In *Re-Imagining Learning Scenarios* (pp. 24). Budapest, Hungary: European Distance and E-Learning Network.

Kadoić, N., Begičević Ređep, N., & Divjak, B. (2017). Decision Making with the Analytic Network Process. In M. Kljajić Borštnar, L. Zadnik Stirn, J. Žerovnik, & S. Drobne (Eds.). *SOR 17 Proceedings* (pp. 180–186). Bled, Ljubljana: Slovenia Society Informatika – Section for Operational Research.

Kampylis, P., Punie, Y., & Devine, J. (2015). *Promoting effective digital-age learning : a European framework for digitally-competent educational organisations.* European Union, 2015: Joint Research Centre.

Lawshe, C. H. (1975). A Quantitative Approach to Content Validity. *Personnel Psychology*, *28*(4), 563–575.

Mihelčić, M., & Bohanec, M. (2017). Approximating incompletely defined utility functions of qualitative multi-criteria modeling method DEX. *Central European Journal of Operations Research*, *25*(3), 627–649. https://doi.org/10.1007/s10100-016-0451-x

Raffai, M. (2008). Model Oriented Enterprise Integration: Metamodel for Realizing the Integration. In *Research and Practical Issues of Enterprise Information Systems II* (pp. 807–816). Boston, MA: Springer US. https://doi.org/10.1007/978-0-387-76312-5_1

Saaty, T. L. (1999). Fundamentals of the analytic network process — Dependence and feedback in decision-making with a single network. *Journal of Systems Science and Systems Engineering*, *13*(2), 129–157.

Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, *1*(1), 83–98.

Saaty, T. L., & Cillo, B. (2008). *A Dictionary of Complex Decision Using the Analytic Network Process, The Encyclicon, Volume 2* (2nd ed.). Pittsburgh: RWS Publications.

Saaty, T. L., & Vargas, L. G. (2006). *Decision Making with the Analytic Network Process: Economic, Political, Social and Technological Applications with Benefits, Opportunities, Costs and Risks* (pp. 278). New York: Springer US.

SCALE CCR. (2012). Up-scaling creative classrooms in Europe. Retrieved from http://is.jrc.ec.europa.eu/pages/EAP/SCALECCR.html

Karsai, G., Nordstrom, G., Ledeczi, A., & Sztipanovits, J. (2000). Specifying graphical modeling systems using constraint-based meta models. U *CACSD. Conference Proceedings. IEEE International Symposium on Computer-Aided Control System Design (Cat. No.00TH8537)* (str. 89–94). IEEE. http://doi.org/10.1109/CACSD.2000.900192

Sprinkle, J. (2004). Model-integrated computing. *IEEE Potentials*, *23*(1), 28–30. http://doi.org/10.1109/MP.2004.1266937

Sprinkle, J., Rumpe, B., Vangheluwe, H., & Karsai, G. (2010). Metamodelling. U *Metamodelling. Lecture Notes in Computer Science* (str. 57–76). http://doi.org/10.1007/978-3-642-16277-0_3

Vangheluwe, H., & de Lara, J. (2002). Meta-Models are models too. U *Proceedings of the Winter Simulation Conference* (Sv. 1, str. 597–605). IEEE. http://doi.org/10.1109/WSC.2002.1172936

Yin, S.-H. (2012). Application of DEMATEL, ISM, and ANP for key success factor (KSF) complexity analysis in R&D alliance. *Scientific Research and Essays*, *7*(19), 1872–1890.

Watts, S., & Stenner, P. (2005). Doing Q methodology: theory, method and interpretation. *Qualitative Research in Psychology*, *2*, 67–91. https://doi.org/10.1191/1478088705qp022oa

Educational Data Driven Decision Making: Early Identification of Students at Risk by Means of Machine Learning

Romano Kovač
Poslovna inteligencija d.o.o.
Krste Pavletića 1, 10000 Zagreb
romano.kovac@outlook.com

Dijana Oreški
University of Zagreb, Faculty of Organization and Informatics
Department of Information Systems Development
Pavlinska 2, 42000 Varaždin
dijana.oreski@foi.hr

Abstract. *In the last few years there has been a notable increase in the data mining usage for educational purposes. Educational data mining is emerging field of research which has the aim of analysing data about students` activities. Prediction of student achievements is among the fastest growing research in this domain. Main goal of this paper is to provide useful knowledge to faculties and their management using data about students` activity at the LMS Moodle and comparing different machine learning techniques in order to analyse this data. In this paper we have evaluated four machine learning algorithms: neural networks, decision tree, support vector machines and logistic regression. Decision tree shown to be most accurate predictive model. Results indicated lecture and seminar attendance as significant predictors of academic success.*

Keywords. Data-driven educational decision making, decision support system, machine learning, academic performance.

1 Introduction

As education increasingly relies on technology, huge amount of data about students` activity is available. Records with student activities, ratings, interactions with teachers and other students are now collected through learning management systems (LMS) such as Edmodo or Moodle. Within all levels of education there is a need to use this data in order to develop systems that help to increase the likelihood of students' success. Data mining is one of the most popular tools for student performance analysis and it has been widely applied in the educational area leading to the development of special subfield called educational data mining. Purpose of educational data mining is to transform raw data into useful knowledge. Educational data mining enables data-driven educational decision making at all levels (Van Barneveld et al., 2012). Main direction of the research in educational data mining is analysing data from online courses. This research aims to analyze data from students` activity at the one Moodle course which was carried out at the University of Zagreb, Faculty of Organization and Informatics.

Paper is organized as follows. Second section gives brief overview of previous research in this domain. Framework for research is set up in the third section, following by modelling description and research results in fourth section. At the end, conslucion is given and directions for further research are presented.

2 Related work

This chapter will highlight some related scientific research on the issue of predictive models development for intervention system and student performance measurements by the means of data mining methods.
Al-Barrak and Al-Razgan (2015) collected data from 170 students enrolled ate the *Data Structure* course. *Data Structures* course has high failure among IT students and goal of the research was to identify how to predict students´ failure? Data mining methods are applied in the study of student success in this course. There were 158 students included in the research. Each student record has the following features: student ID, student name, test 1 grade, test 2 grade, test 3 grade, inter-exam 1, intermediate level 2, project, tutorials, final exam, and total number of points earned. The score for the course was 60 points per year and 40 points for the final exam. The student must have at least 60 out of 100 points to complete the course. Result of the research was predictive model with the precision of 91 %.
Daud, Aljohani, Abbas, Lytras, Abbas and Alowibdi (2017) performed research with the purpose to predict students' performance: whether they would successfully finish their study or they would fail. Data were collected by graduate and undergraduate students at various universities in Pakistan during the period of 2004 to 2011. Number of 776 student records were

collected, of which 690 students who completed their study and 86 students who dropout. Authors applied Support Vector Machines (SVM), C4.5, Classification and Regression trees (CART), Bayesian nets (BN), and Naive Bayes (NB) in their comparative analysis. Precision, recall and F1 were used as performance measures. SVM works the best on data set with a F1 result of 0.867, which is 13% better compared to the second best method. BN and NB classifiers gave better results compared to C4.5 and CART. Student academic achievements are based on various factors such as the environment, personality, social, psychological and other variables.

Bhardwaj and Pal (2011) use Naive Bayes classification algorithm. They found that student performance was highly dependent on their grade obtained from high school examinations. Furthermore, other important variable for the student success prediction was student's location of residence. In, general, their study shows that the academic outcomes are not always dependent on their own efforts but also other factors have a significant impact on student performance.

Romero, Ventura and Garcia (2007) investigated Moodle as source of a large amount of data that is very useful for student behavioral analysis and could create a real gold mine for educational data mining. Moodle records all the student activities involved, such as reading, writing, testing, performing different tasks, and even communicating with peers. Moodle also provides personal user information (profile), academic results, and user interaction data.

Corsate and Walker (2015) performed comparison of machine learning algorithms application to the Moodle data set. The results of the modelling were analyzed with the aim of transforming available information into structured intelligent system. This study used unsupervised learning method, K-means clustering, method that divides data into clusters. Students with the highest level of similarity are grouped together in the same cluster.

Keshtkar, Cowart and Crutcher (2018) used data set that contains metadata for interactivity of students and professors on LMS Moodle through 11 programming courses over two semesters at the State University South East Missouri. The final grade was output variable and consisted of five values: A, B, C, D and F, with A to C being a passing grade. They analyzed the success and failure of the students. Number of included students was 195, with 157 passes and 37 drops. Data set contained the following features: average and total number of interactions per session, whether the interaction is performed within a campus or outside the campus, the result of the first exam and the final grade of the course. The results show that Logistics Tree Model can serve as tool for prediction of dropout. In this way, it is possible to find out which students are at risk of failure.

Mayilvaganan and Kalpanadevi (2014) collected data about 197 students for the purpose of data analysis by means of data mining techniques. Student data includes the following attributes: specialty, previous grades, additional knowledge or skills, resources, class attendance, time spent learning, grade exam, seminar achievement, laboratory work, tests results and online assignments results. In this study, the discussion focuses on three classification techniques: decision tree, Naive Bayes, and k-nearest neighbor. The result concluded that the k-nearest neighbor had the best accuracy in the classification compared to other techniques due to the significance of the test results.

Saa (2016) also performed a study with objective to detect the relationship of personal and social factors of students on their educational performance by using data mining methodology. The data set used in this study was collected by a survey distributed to students in their daily classes and also as an online survey. The initial set of data set consisted of 270 records. Four decision tree algorithms have been implemented: Naive Bayes algorithm, ID3, C4.5 and CART. Interesting results from the classification models were extracted. It has been discovered that the performance of students is not entirely dependent on their academic efforts and there are many other factors that are equally influential.

Shahiria, Husain and Rashida (2015) claim that final grades are based on the structure of the course, assessment methods, results of the final exam and extracurricular activities. They applied several techniques for evaluating students' success. The results indicated neural networks as the method with the highest accuracy of predictions (98%), followed by decision trees (91%). Support vector machines and KNN gave the same accuracy of 83%, and Naive Bayes shown to be the method with the lower predictive accuracy (76%).

Based on the literature review, we have identified mostly used machine learning algorithms which will be used in our research.

3 Research framework

Graduation rates are important parameters and management of educational institutions are looking for new ways of predicting success and failure early enough in student education to achieve effective interventions as well as identifying the effectiveness of different interventions. The need to analyze the large amounts of data generated from the educational ecosystems urged development of educational data mining. Al-Barrak and Al-Razgan (2015) define educational data mining as a process of applying tools and techniques for data analysis in educational setting. The application of data mining techniques to educational data will help the education sector to improve its learning process. This is main motive of this paper.

There are two main elements in educational data mining: data set consisting of features which are

describing student activity and methods for analysis of data set.

Several methods are developed for predictive modelling. Depending on the task employed, methods are categorized into classification or regression methods. The most popular task of student performance prediction is mostly categorized as classification task. There are several algorithms in the classification task that are applied to predict student success. Literature review revealed decision tree, artificial neural network, logistic regression, and Support Vector Machine as mostly used methods.
An example of data stored in databases of educational institutions includes information on students enrollment on courses, their scores, activities, teacher notes, socio demographic characteristics. Once this data is properly analyzed, it will help improve knowledge in the education sector.
EDM can help universities to better plan the foreseen number of students enrolling in specific programs, predicting a dropout ratio, identification of students at risk of dropout, and better exploiting the available resources.
This will help the educational institutions to evaluate, plan and decide on their educational programs. It is expected that this new knowledge will reveal hidden forms that will help academic programs to use their resources more efficiently.

The system design in this research was carried out through four steps:
(i) data collection: students use LMS system and their interaction data is stored in the database. In this paper we have used data from students enrolled at the Moodle course.
(ii) data preprocessing: includes data cleaning and data transformation to the appropriate mining format. To pre-process Moodle data, we have used pre-processing possibilities of Azure tool.
(iii) modelilng: Four data mining algorithms were applied in order to develop a model that find patterns and summarizes knowledge to the teachers, administartors and students about students` behaviour.
(iv) system evaluation and interpretation: evaluate and implement results. The results obtained by model are interpreted and used by instructors for further procedures. The instructor can use disclosed information to make decisions about students' activities and the Moodle program to enhance student learning.
Next section explains implementations of this steps as well as model design and evaluation.

4 Model design and evaluation

This section explains empirical research based on the methodology described previously. Microsoft Azure Machine Learning platform for machine leraning is applied with the aim to develop predictive models. Four machine learning algorithms were employed to do so: logistic regression, neural networks, support vector machines and decision tree. First we will explain data set used for training and testing the model.

4.1 Data description

Data set used in this research was collected from the one course taught at the University of Zagreb, Faculty of Organization and Informatics. Data set consists of 235 records about students activities. Data were collected from LMS system Moodle. Range of transformations was made to the data set including feature selection and transformation of categorical features into numerical. Dimensionality of data set was 13, but only 11 variables was included in the data analysis since some of them were derived one from another. Description of the features is given in the Table A1. Furthermore, descriptive statistics for output variable is presented at Fig 1.

Statistics

Mean	1.4979
Median	1
Min	1
Max	5
Standard Deviation	0.993
Unique Values	5
Missing Values	0
Feature Type	Numeric Feature

Figure 1. Descriptive statistics

Output variable is binary and indicates pass or fail: fail = 0 and 1 = pass. Thus, we are dealing here with the classification since the mai aim is the prediction of binary target. Task of predicting a continuous target is referred to as a regression task (Kelleher, Mac Namee, D'Arcy, 2015).
Output variable is constructed based on the grades. Distribution of grades is given at Fig 2.

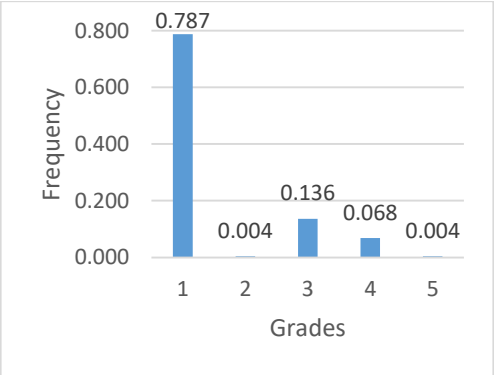


Figure 2. Distribution of grades

Following data preprocessing and data exploration phase, modelling was performed and implemented in Microsoft Azure ML cloud platform. Data set was split into trening (165 instances) and test data (70 instances). Results of the data modelling are presented in the next section.

4.2 Model description

Microsoft Azure Machine Learning served as platform for data model flow (Figure 3). First, data set was introduced. Then features were selected and input and output features were defined.

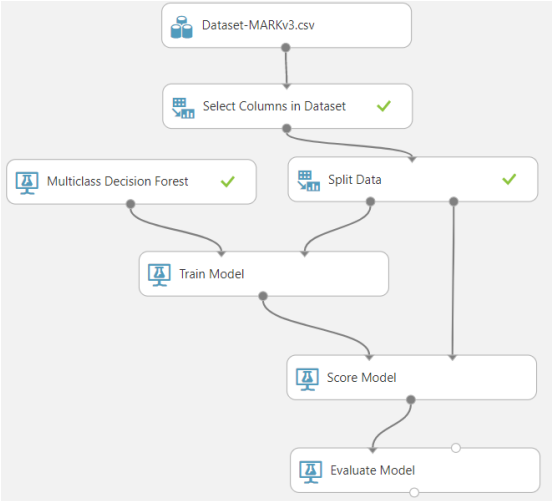


Figure 3. Data modeling flow

Data set is divided into the training set and the test set in a ratio of 70 % : 30 %. Training set is used to train the model while the test set is used to identify how precise a model is in generalizing over a new data set. Evaluation of the model completes the whole process by calculating accuracy, precision and recall measure. The results of the model training by using logistic regression are shown at figure 4. Logistic regression algorithm achieved a score of 92.85% accurate predictions.

Table 1. Performance scores of classification algorithms

Metrics/ Algorithm	*LR*	*NN*	*DT*	*SVM*
Overall accuracy	0,9286	0,9429	0,9714	0,9571
Average accuracy	0,9286	0,9429	0,9714	0,9571
Micro-averaged precision	0,9286	0,9429	0,9714	0,9571
Macro-averaged precision	0,8837	0,8839	0,9833	1
Micro-averaged recall	0,9286	0,9429	0,9714	0,8622
Macro-averaged recall	0,8577	0,9325	0,9167	0,7501

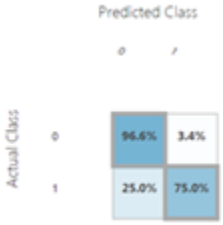


Figure 4. Confusion matrix of logistic regression model

Performance of the model developed by neural networks algorithm is shown at Figure 5. The neural network algorithm achieved a score of 94.28% accurate predictions.

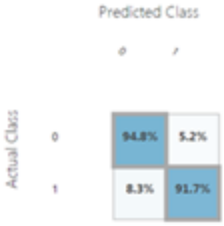


Figure 5. Confusion matrix of neural network model

Decision tree prediction model shown to be the most accurate model achieving 97.14% accurate predictions.

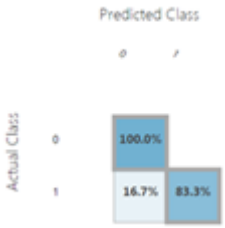


Figure 6. Confusion matrix of decision tree model

The results of the modelling by using support vector machine algorithm, are shown in Figure A1, attached to this paper. The support vector machine algorithm proved to be a very accurate model since achieved a score of 95.7% of accurate predictions.

4.3 Model interpretation

Sensitivity analysis is also performed in the research in order to identify predictors of academic performance. Seminar attendance and lecture attendance shown to be features with the highest impact on the passing or

failing the course in all machine learning models. Student attendance is a consistent source of discussion for researchers. Positive effect of attendance commonly reported in literature is noteable here. Our results are in line with previous work of Paisey and Paisey (2004). They identified strong positive relationship between attendance at classes and academic performance. However, Chen and Lin (2015) are sceptical and they consider that positive effects of an attendance considered in prior literature must be reassessed. Andrietti (2014) did find a positive and significant effect of attendance on academic performance. Furthermore, she used proxy variables regressions to capture the effect of unobservable student traits possibly correlated with attendance. As suggested in the research of Mearman et.al. (2014) wide range of factors affect attendance, such as the quality of teaching sessions or students aspirations. This must be investigated in future research. Our study highlights various interesting findings, but also opens questions that require further investigation.

Results regarding machine learning algorithms comparison emphasized decision trees as most accurate model. Decision tree has valuable characteristic of considering only those attributes that are helpful to the classification.

5 Conclusion

These paper deals with machine learning algorithms applications in educational domain with the aim to develop decision support system for identification of students at risk.
Through this paper, the basic concepts of four learning algorithms were presented: logistic regression, neural networks, support vector machines and decision tree are applied on real data set. Machine learning was applied in educational domain. Special emphasis was on model evaluation in order to compare different approaches to machine learning. Main research question was: in which extent the total data on students, generated by LMS Moodle, can be a good basis for predictive modelling. Such predictive models of student success based on Moodle activity serve as input into decision support system used in detecting behavior of future generations of students. Decision support system provides valuable tool for educational institutions in achievieng one of their main goals: increasing quality of the study. Management of educational institutions are potential users of such systems. At the beginning of the academic year, groups where students of similar profiles could be formed. Also, the system would briefly inform the relevant body about the student's performance during the key moments of the semester and highlight those students with whom a personalized program was needed for the purpose of studying. Results of this research yielded highly accurate models and provided basis for educational data driven decision making. However, there are several limitations of this research. First, models are developed using students` data from only course. Second, specific group of students was included, just informatics students. Thus, it is hard to generalize results. In the future research we will increase our sample and include students from different fields of study. Furthermore, other approaches to machine learning will be applied.

References

Andrietti, V. (2014). Does lecture attendance affect academic performance? Panel data evidence for introductory macroeconomics. International Review of Economics Education, 15, 1-16.

Al-Barrak, M. A., Al-Razgan, M. S. (2015) *Predicting Students' Performance Through Classification: A Case Study, Journal of Theoretical and Applied Information Technology*

Bhardwaj, B.K., Pal, S. (2011) Data Mining: A prediction for performance improvement using classification, (IJCSIS) International Journal of Computer Science and Information Security, Vol. 9, No. 4

Chen, J., & Lin, T. F. (2015). Effect of peer attendance on college students' learning outcomes in a microeconomics course. The Journal of Economic Education, 46(4), 350-359.

Corsatea, B., Walker, S. (2015) Opportunities for Moodle data and learning intelligence in virtual environments, School of Computer Science and Electronic Engineering, University of Essex, UK

Daud, A., Aljohani, N. R., Abbasi, R. A., Lytras, M. D., Abbas, F., Alowibdi, J. S. (2017) Predicting Student Performance using Advanced Learning Analytics, International World Wide Web Conference Committee (IW3C2)

Kelleher, J. D., Mac Namee, B., & D'Arcy, A. (2015). Fundamentals of machine learning for predictive data analytics: algorithms, worked examples, and case studies. MIT Press.

Keshtkar, F., Cowart, J., Crutcher, A. (2018) Predicting Risk of Failure in Online Learning Platforms Using Machine Learning Algorithms for Modeling Students' Academic Performance, available at: http://medianetlab.ee.ucla.edu/papers/ICMLWS1.pdf

Mayilvaganan, M., Kalpanadevi D. (2014) Comparison of Classification Techniques for predicting the performance of Students Academic Environment, International Conference on

Communication and Network Technologies (ICCNT)

Mearman, A., Pacheco, G., Webber, D., Ivlevs, A., & Rahman, T. (2014). Understanding student attendance in business schools: An exploratory study. International Review of Economics Education, 17, 120-136.

Paisey, C., & Paisey, N. J. (2004). Student attendance in an accounting module–reasons for non-attendance and the effect on academic performance at a Scottish University. Accounting education, 13(sup1), 39-53.

Romero, C., Ventura, S., Garcia E. (2007) Data mining in course management systems: Moodle case study and tutorial, Department of Computer Sciences and Numerical Analisys, University of Cordoba, 14071 Cordoba, Spain

Saa, A. A. (2016) Educational Data Mining & Students' Performance Prediction, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 5

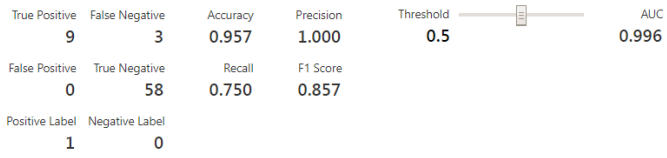
Shahiria, A. M., Husaina, W., Rashida, N. A. (2015) A Review on Predicting Student's Performance using Data Mining Techniques, School of Computer Sciences Universiti Sains Malayisa 11800 USM, Penang, Malaysia

Van Barneveld, A., Arnold, K. E., & Campbell, J. P. (2012). Analytics in higher education: Establishing a common language. Educause Learning Initiative, 1, 1–11. ELI Paper.

APPENDIX

Table A1. Description of features

Feature	Description	Values
Gender	Students gender	1 – female 2 – male
Status	Student status	1 – part time student 2 – full time student
Lecture_attendance	Frequency of lecture attendance	Numeric; scale: 0 – 6
Seminar_attendance	Frequency of seminar attendance	Numeric; scale: 0 – 4
Activity	Students activity at the seminars and lectures	Numeric; scale: 0 – 6
Mark_plan	Marketing plan as students assignment in the project	Numeric; scale: 0 – 25
Presentation	Presentation of the marketing plan	Numeric; scale: 0 – 10
Exam_1	First written exam	Numeric; scale: 0 – 25
Exam_2	Second written exam	Numeric; scale: 0 – 25
Assignemnt effort	Assignment for adittional points	Numeric; scale: 0 – 3
In class activity	Additional points for a course assignment	Numeric; scale: 0 – 3
Pass/Fail	Confirmation that the student has passed or the course	0 – Fail 1 – Pass



Score Bin	Positive Examples	Negative Examples	Fraction Above Threshold	Accuracy	F1 Score	Precision	Recall	Negative Precision	Negative Recall	Cumulative AUC
(0.900,1.000]	0	0	0.000	0.829	0.000	1.000	0.000	0.829	1.000	0.000
(0.800,0.900]	1	0	0.014	0.843	0.154	1.000	0.083	0.841	1.000	0.000
(0.700,0.800]	1	0	0.029	0.857	0.286	1.000	0.167	0.853	1.000	0.000
(0.600,0.700]	6	0	0.114	0.943	0.800	1.000	0.667	0.935	1.000	0.000
(0.500,0.600]	1	0	0.129	0.957	0.857	1.000	0.750	0.951	1.000	0.000
(0.400,0.500]	2	3	0.200	0.943	0.846	0.786	0.917	0.982	0.948	0.047
(0.300,0.400]	1	3	0.257	0.914	0.800	0.667	1.000	1.000	0.897	0.099
(0.200,0.300]	0	6	0.343	0.829	0.667	0.500	1.000	1.000	0.793	0.203
(0.100,0.200]	0	17	0.586	0.586	0.453	0.293	1.000	1.000	0.500	0.496
(0.000,0.100]	0	29	1.000	0.171	0.293	0.171	1.000	1.000	0.000	0.996

Figure A1. Evaluation of support vector machines model

Author Guidelines

Microsoft Word Template for the Central European Conference on Information and Intelligent Systems (CECIIS)*

Author 1, Author 2
Author's Common Affiliation
Department/Institute
Full Address
{author1, author2}@domain.com

Author(s) Name(s)
Author Affiliation(s)
Department/Institute
Full Address(es)
E-mail(s)

Abstract. *The abstract is to be in fully-justified italicized text, at the top of the left-hand column as presented here, below the author information. Use the word "Abstract." in 10-point Times, boldface type, left positioned, initially capitalized, followed by the abstract in 10-point, single-spaced type, up to 100 words long.*
Leave one blank line after the abstract, and then begin the keywords. Use the word "Keywords." in 10-point Times, boldface type, left positioned, initially capitalized, followed by up to ten keywords in 10-point, separated by comma, as below.
Leave two blank lines after the keywords, and then begin the main text.

Keywords. CECIIS, conference paper, template

1 Introduction

This is an example of a Central European Conference on Information and Intelligent Systems (CECIIS) submission. These guidelines include complete description of the paper style including formatting, fonts, spacing, and related information for producing your proceedings manuscripts.

This template can be used to create your own paper but if you have any troubles or if in doubt do not hesitate to contact us via e-mail: ceciisoo@foi.hr or via the conference web site at http://www.ceciis.foi.hr.

2 Structure and Requirements for Research Paper (Examples)

Central European Conference on Information and Intelligent Systems (CECIIS) welcomes research papers on all topics of interest (Data and Knowledge Bases, Economics and Information Systems, Education for Information Society, Information and Communication Technologies, Information Systems Security, Intelligent Information Systems, Intelligent Transport Systems and Autonomous Vehicles, Software Engineering, Quality of Software and Services) as well as on the Conference theme/special emphasis announced on yearly base, addressing both, research theory and practice.

The following paper structure examples regarding the research type could be used for preparing a CECIIS paper:

1. Quantitative questionnaire-based: background research (state of the art); research problem; hypothesis or research question; description of instrument used; sampling; statistical methods used for analysis; discussion of results with limitations of research.
2. Qualitative research: background research (state of the art); research problem; research question; qualitative method used (case study research, interviews, content analysis etc.); discussion of results with limitations of research.
3. Algorithms and theoretical research: theory description with references; original theoretical contribution (theorem, algorithm etc.) with proofs or confirmations; future research and conclusion
4. Systematic literature review with meta-analysis: scope of research, research questions or research purpose; sources of literature with period covered (data bases, journals etc.), methods of analysis, conclusions with future research

Research paper should be theoretically and evidence based, clearly and consistently as well as ethically prepared and written. Central European Conference on Information and Intelligent Systems (CECIIS) has zero-tolerance policy on any kind of plagiarism. Paper containing text copied and used from another author(s) work without permission or from authors previously published own work, without referencing, will be retracted.

* The responsibility to apply the prescribed formatting of the CECIIS papers rests entirely with the authors

3 Paper Organization and Formatting

All papers should be written in English, up to 8 pages long (approximately 20 000 characters), and arranged in the following order:

- Main title
- Author(s), affiliation(s), full and e-mail addresses
- Abstract
- Keywords
- Body text (Main text)
- Footnotes
- Acknowledgements
- References

All printed material, including text and figures, must be kept within a print area of 16 cm (6.3") wide by 24.7 (9.7") high. Do not write or print anything outside the print area. Paper size is A4: 21x29.7 cm (8.3x11.7"). Margins (top, bottom, left, right) are 2.5 cm (1"). All *text* must be in a two-column format. Columns are to be 7.7 cm (3") wide, with a 0.6 cm (0.24") space between them. Text must be fully justified. Indent each paragraph by 0.5 cm (0.2").

The final submission has to be submitted in a single PDF file with all fonts that are used embedded, and also the source file (Microsoft Word, LaTeX - zip original files and images).

4 Main Title

The main title (on the first page) should begin 3,5 cm from the top edge of the page, centred, and in Times 18-point, bold face. Capitalize the first letter of nouns, pronouns, verbs, adjectives, and adverbs; do not capitalize articles, coordinate conjunctions, or prepositions (unless the title begins with such a word). Leave a blank line after the title.

5 Author Name(s) and Affiliation(s)

Author names and affiliations are to be centred beneath the title and printed in Times 10-point type. Author names should be in bold face. Multiple authors shall be grouped by affiliation as shown in the title above. Include also e-mail addresses in 9-point `Courier/Typewriter` face. Follow the author information by two blank lines before main text.

6 Type-style and Fonts

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your word processor, please use the font closest in appearance to Times that you have access to. Please avoid using bit-mapped fonts if possible. True-Type 1 fonts are preferred.

Also make sure that all fonts in the final PDF file are embedded. Thus we strongly encourage the use of Latex especially if you use non-standard fonts for formulas for your examples. You can make sure that your fonts are embedded by opening the PDF file in some PDF viewing program and taking a look at the properties of the document.

7 First-order Headings (May Extend to the Next Line)

For example, “1 Introduction”, should be Times 14-point boldface, initially capitalized, flush left, with two blank lines before, and one blank line after.

Don't use periods (“.”) after the heading number, use them only in lower order headings to separate them from higher order headings. For *long* headings use *a hanging indent* aligning the text to the right of the heading number as shown above.

All paragraphs in the text should be indented except for the first paragraph in a section.

7.1 Second-order Headings (May Extend to the Next Line)

As in this heading, they should be Times 12-point boldface, initially capitalized, flush left, with one blank line before, and one after. Use a hanging indent for long headings.

7.1.1 Third-order Headings (May Extend to the Next Line)

Third-order headings, as in this paragraph, are discouraged. However, if you must use them, use 10-point Times, boldface, initially capitalized, flush left, with one blank line before, and one after. Use a hanging indent for long headings.

8 Main Text

Type your main text in 10-point Times, single-spaced. Do not use double-spacing. All paragraphs should be indented 0.5 cm (0.2") except for the first paragraph in a section.

Be sure your text is fully justified—that is, flush left and flush right. Please do not place any additional blank lines between paragraphs.

8.1 Figures

All figures are to be included within the text. **Figure captions** are to be *below* the figures, in 10-point

Times (or a similar serif font), normal face. Initially capitalize only the first word of each figure caption.

Figures are to be numbered consecutively with Arabic numerals throughout the paper, for example: "Figure 1. Database contexts", and are referred to in the text as Fig. 1, Fig. 2, etc.

8.1.1 Illustrations, Photographs and Graphs

Illustrations, photographs and graphs are considered as figures. All graphics should be centred. Your artwork must be in place in the article (preferably printed as part of the text rather than pasted up). If you are using photographs and are able to have halftones made at a print shop, use a 100- or 110-line screen. Supply the best quality photographs and illustrations possible.

Pencilled lines and very fine lines do not reproduce well. Remember, the quality of the book cannot be better than the originals provided. The conference proceedings will be printed in grey scale, but you can provide images in colour since the papers will also be published on the web site of the conference.

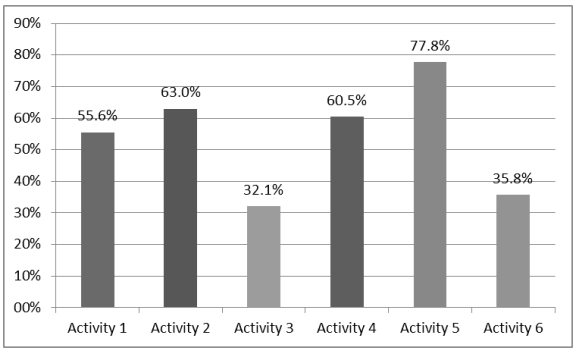


Figure 1. Times, 10 points, normal face

8.2 Tables

All tables are to be included within the text. **Table titles** are to be *above* the tables, in 10-point Times (or a similar serif font), normal face. Initially capitalize only the first word of each table title. Tables are to be numbered consecutively with Arabic numerals throughout the paper, for example: "Table 1. Input data", and are referred to in the text as Table 1, Table 2, etc. Column headings should be as brief as possible.

Table 1. Times, 10 points, normal face

	Column 1	Column 2	Column 3
Row 1	1	2	3
Row 2	4	5	6
Row 3	3	4	7
Row 4	2	6	4
Row 5	2	2	4

8.3 Equations

Equations should be numbered serially on the right-hand side by Arabic numerals in parentheses, and referred to in the text by eq. 1, eq. 2 etc.

$$3x^2 + 3x - 23 = 0 \quad (1)$$

8.4 Footnotes

Use footnotes sparingly (or not at all!) and place them at the bottom of the column on the page on which they are referenced.[1]

9 In-text Citations and Referencing

In-text citation must be used to denote all text used from another author(s) work or from authors previously published own work. When referenced in the text, enclose the author surname(s) and the publication year in brackets, for example (Surname, year) or (Surname, year, p. x) where x is the page number of the cited reference. When referenced paper has more than three authors, cite the paper as (Surname et al., year).

List and number all bibliographical references, alphabetically sorted, in 10-point Times, single-spaced, with a hanging indent, at the end of your paper. Use 6-point paragraph spacing after each reference. References should be distributed evenly in both columns as much as possible, so use a column break where appropriate.

In the References section below you will find examples of common types of bibliographical items: book (Mayer, 2009, p. 52), journal paper (Bule & Peer, 2014), journal paper with more than three authors (Shukor et al., 2014), conference proceedings paper (Steingartner & Novitzká, 2015), report or standard (*ISO/IEC 25010: Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models*, 2011) and web page ("Web Content Accessibility Guidelines (WCAG) 2.0," 2008). For other types of bibliographical items refer the **APA style guidelines**. It should be ensured that every reference cited in the text is also listed in the References section (and vice versa).

Acknowledgments

Acknowledgments, if necessary, should appear in a separate paragraph preceding the references.

[1] Use Times 8-point type, single-spaced. Avoid using footnotes altogether and include necessary peripheral observations in the text (within parentheses, if you prefer, as in this sentence).

References

Bule, J., & Peer, P. (2014). Technical, Legal, Economic and Social Aspects of Biometrics for Cloud Computing. Journal of Information and Organizational Sciences, 38(2), 83–95.

ISO/IEC 25010: Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models. (2011).

Mayer, R. E. (2009). Multimedia Learning (Second Edi.). New York: Cambridge University Press.

Shukor, N. A., Tasir, Z., Van der Meijden, H., & Harun, J. (2014). A Predictive Model to Evaluate Students' Cognitive Engagement in Online Learning. Procedia - Social and Behavioral Sciences, 116, 4844–4853. doi:10.1016/j.sbspro.2014.01.1036

Steingartner, W., & Novitzká, V. (2015). A new approach to operational semantics by categories. In T. Hunjak, V. Kirinić, & M. Konecki (Eds.), Proceedings of the 26th Central European Conference on Information and Intelligent System (CECIIS 2015) (pp. 247–254). University of Zagreb, Faculty of Organization and Informatics Varaždin.

Web Content Accessibility Guidelines (WCAG) 2.0. (2008). Retrieved from https://www.w3.org/TR/WCAG20/

CECIIS 2018 Sponsors

INFODOM


ORACLE®

CECIIS 2018 Financial Support

Ministarstvo znanosti i obrazovanja RH

Varaždinska županija

CECIIS 2018 Program Committee

Vjeran Strahonja, Program Committee Chair, University of Zagreb, Croatia

Nina Begičević Ređep, University of Zagreb, Croatia
Leo Budin, University of Zagreb, Croatia
Tonči Carić, University of Zagreb, Croatia
Ivica Crnković, Mälardalen University, Sweden
Blaženka Divjak, University of Zagreb, Croatia
Jasminka Dobša, University of Zagreb, Croatia
Matjaž Gams, Jožef Stefan Institute, Slovenia
Gordan Gledec, University of Zagreb, Croatia
Andrina Granić, University of Split, Croatia
Valentina Kirinić, University of Zagreb, Croatia
Melita Kozina, University of Zagreb, Croatia
Marjan Krašna, University of Maribor, Slovenia
Alen Lovrenčić, University of Zagreb, Croatia
Sandra Lovrenčić, University of Zagreb, Croatia
Ivan Luković, University of Novi Sad, Serbia
Aleksandar Marković, University of Zagreb, Croatia
Dunja Mladenič, Jožef Stefan Institute, Slovenia
Oliver Moravčik, Slovak University of Technology, Slovakia
Jan Paralič, Technical University of Košice, Slovakia
Elisabeth Pergler, Evolaris next level GmbH, Austria
Wolf Rauch, University of Graz, Austria
Sonja Ristić, University of Novi Sad, Serbia
William Steingartner, Technical University of Košice, Slovakia
Violeta Vidaček Hainš, University of Zagreb, Croatia
Mladen Vouk, North Carolina State University, USA
Neven Vrček, University of Zagreb, Croatia
Ksenija Vuković, University of Zagreb, Croatia

Honorary Members - Former Program Committee Chairs

Boris Aurer, University of Zagreb, Croatia
Tihomir Hunjak, University of Zagreb, Croatia
Mirko Maleković, University of Zagreb, Croatia
Miroslav Žugaj, University of Zagreb, Croatia

CECIIS 2018 Organizing Committee

Valentina Kirinić, Organizing Committee Chair, University of Zagreb, Croatia

Ivana Dvorski Lacković, University of Zagreb, Croatia
Matija Kaniški, University of Zagreb, Croatia
Izabela Oletić Tušek, University of Zagreb, Croatia
Nela Kivač, University of Zagreb, Croatia

CECIIS 2018 Research Tracks' Chairs

Computer Games: Mario Konecki and Markus Schatten, University of Zagreb, Croatia
Data and Knowledge Bases: Alen Lovrenčić, Mirko Maleković and Kornelije Rabuzin, University of Zagreb, Croatia
Education and Learning Analytics : Blaženka Divjak and Igor Balaban, University of Zagreb, Croatia
Emerging trends in ICT: Neven Vrček and Sandra Lovrenčić, University of Zagreb, Croatia
ICT Entrepreneurship and Innovation: Ksenija Vuković and Marina Klačmer Čalopa, University of Zagreb, Croatia
Intelligent Information Systems: Jasminka Dobša and Ivan Magdalenić, University of Zagreb, Croatia
Quality of Software and Services: Valentina Kirinić and Melita Kozina, University of Zagreb, Croatia
Software Engineering: Vjeran Strahonja and Zlatko Stapić, University of Zagreb, Croatia
Strategic Planning and Decision Making: Nina Begičević Ređep and Tihomir Hunjak, University of Zagreb, Croatia

CECIIS 2018 Professional Tracks' Chairs

Digital Transformation of Educational Institutions (Digitalna transformacija obrazovnih institucija): Nina Begičević Ređep and Katarina Tomičić-Pupek, University of Zagreb, Croatia
Computer Games - Professional Development (Računalne igre - stručna postignuća): Mario Konecki, Alen Lovrenčić and Mladen Konecki, University of Zagreb, Croatia

CECIIS 2018 Student Poster and Presentation Section Chair

Violeta Vidaček Hainš, University of Zagreb, Croatia

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