PROGRESS

Journal for political theory and practice

Publisher
Foundation “For Serbian People and State”, Belgrade
Co-publisher
Institute for Political Studies, Belgrade

For the publisher
Tatjana Vukić, Director
For the co-publisher
Miša Stojadinović, PhD, Principal research fellow, Director

The journal is published three times a year.
The first edition of the journal “Napredak” was published on Vidovdan, on 28th June 2020

Print run: 500 copies • Printed by: Birograf, Belgrade
Address of the Editorial Board: Palmira Toljatija 5, Belgrade

http://fondacijasnd.rs/casopis-napredak/

Papers are sent electronically

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

32(497.11)

PROGRESS : journal for political theory and practice / editor-in-chief Zoran Jevtović. - [English ed.]. - Vol. 5, no. 1 (2024)- . - Belgrade : Foundation “For Serbian People and State”, 2024- (Belgrade : Birograf). - 24 cm

Tri puta godišnje. - Povremeno sa tematskim br. - Preuzima numeraciju srpskog izd. - Drugo izdanje na drugom edijumu: Напредак (Београд. Online) = ISSN 2683-6114. - Prevod dela: Напредак (Београд) = ISSN 2683-6106
ISSN 3042-0261 = Progress (Belgrade)

COBISS.SR-ID 143948297

Vol. V / No. 2
2024.

Thematic issue

Artificial intelligence – an open gate to the future

Content

Articles

Editorial board

UDC 004.8(497.11)
Editorial
Received: 14.08.2024.
Accepted: 19.08.2024.
doi: 10.5937/napredak5-52765

Artificial Intelligence in Serbia – An Open Gate to the Future

The researchers of social studies have for centuries searched for an answer that discovers the secret of power functioning: who holds it; how it works; where it is hidden, what its meaning is and why there is not sufficient power for all of us. The problem of information management has always been part of social challenges – the space of political and legal systems has hidden a labyrinth of "deep secrets" on which the state functions, while the imbalance of economic, military and security power has enabled the international order to establish itself. Detailed supervision of *others*, along with the collection and processing of large amounts of data, has become one of the most important forms of power, which has been dedicated the attention of scientists worldwide in the past eight decades. The struggle for technological supremacy has also overflowed into the field of national security in which artificial intelligence systems make significant infrastructure of future confrontations at the ideological or military scenes. Extreme future awaiting us will be full of changes, challenges and risks, which means that our ways of thinking must also be adjusted to the modified vision of the computing environment. The combination of nanotechnology, biotechnology, information technology and neurotechnology will shape knowledge and innovation in new markets which will be exposed to permanent pressures of megacorporate capital.

With the beginning of the 21st century, the rise of generative artificial intelligence placed new landmarks in the view of global society because digital technology changed strategically man's environment, from economy, energy and politics, via healthcare and ecology, music and art, to science, education and religion. In the academic field, possibilities have become inexhaustible because idea generation and data processing simplify most complex scientific ideas and operations. The nations that fail to notice the depth and direction of such changes in a timely manner will have huge troubles because their workforce will no longer

be competitive, while the existence of millions of people will be threatened. The map of the future is being drawn at the moment, in the context of the digital environment and geopolitical remodelling, virtual reality and ever-present networks which affect our lives, changing us without our noticing it whatsoever.

Hence the need for an academically precise and scientifically competent response to the importance, characteristics and role of the application of artificial intelligence in Serbian society. The states that decide against the use of the benefits of artificial intelligence will remain trapped in the past because they unconsciously renounce economic progress and numerous strategically relevant information, from climate change to medical operations and treatment. Strategic investments in the research and development of artificial intelligence are one of the priorities of the Serbian government and President Aleksandar Vučić, which is illustrated by the economic indicators: in 2012, we had ICT sector exports in the amount of US$375 million, while one decade later it was between US$3.6 and US$3.7 billion! Our country is the technological hub of the Western Balkans, the regional centre for American IT companies and the leader in the development of software, artificial intelligence and smart IT solutions, which are applied in all sectors of our environment.

When it comes to the question as to what happens behind the doors of large technological companies and who teaches artificial intelligence about ethics and aesthetic values, it is impossible to get a precise answer in this edition of our journal. The fact that private companies control the majority of development processes in Europe and the USA, which means that their focus is on further acquisition of capital. In China and Russia, such research is financed by the respective governments, with the strong coordination or private and public investments. At the beginning of 2024, the EU Parliament adopted the Artificial Intelligence Act (AI Act), which as of the beginning of the next year will impose substantial punishments for the failure to observe the regulations about the use of prohibited AI systems. Researchers observe that, in the absence of the global management of artificial intelligence development, the companies doing it nowadays have power that used to be characteristic of the developed national states. Serbia rapidly understood and began applying the potential of creative industries, becoming the member of the Global Partnership for Artificial Intelligence (GPAI), equally with France, the USA, the UK, India, Japan, Germany Canada and twenty other countries. As an acknowledgement of the proper strategic orientation and a specific international recognition, Serbia has been appointed to chair the work of this organization in the next three years, which strives to establish global standards and rules for the development of artificial intelligence.

Reading thematically different texts in the new edition of *Napredak* (*Progress*) by the authors who analyze this technological-social phenomenon, you will see how the practicality of new technologies on the international scene increases efficiency of the communities using them. To become the leader of the changes in Southeast Europe, the Government of Serbia has done many things, from adopting the National Strategy in 2019 and founding the Institute for Artificial Intelligence two years later, via implementing new technological knowledge

in the educational system, to the formation of the research infrastructure and active changes in the international environment. Generative artificial intelligence can fundamentally change the labour market, increase the global gross domestic product and improve the public and private sectors, which is proved by our authors who particularly point to the current trends and structural changes in the field of economy. We also offer a comprehensive overview of the references so that you can join us in the analysis of theoretical concepts, current AI research outcomes and possible applications of Memristor technology, with ethical considerations and the analysis of regulatory frameworks. Our authors indicate that the development of Memristor technology may place our country in the leading position in relation to innovation in AI hardware and attract international investors, thus encouraging technological progress of the entire economy. The main purpose of our thematic edition is to motivate all interested parties, the state leaders, scientific and expert public, lawyers, innovators, as well as the increasing number of users to consider the challenges deriving from the application of this new technology, while emphasizing a broad range of risks constitutes a step ahead in the identification and understanding of the problems we need to encounter readily, while duly avoiding and preventing other potential problems.

In the past two decades, Serbia has been intensively trying to improve the educational technology by using multimedia, hypermedia, virtual and expanded reality, as well as educational software, which contributes to the improved quality of learning, increased dynamics and obviousness of teaching, as well as encouraging students' motivation. In the past few years, we have particularly explored the sphere of personalized learning, more dynamic teaching, as well as complex evaluation of students' work. The new paradigm of education we propose in one of the texts also stipulates more substantial use of robots in the teaching process with the aim of encouraging creativity and critical thought, as well as to develop problem-solving skills by developing algorithmic thought. In the past year, the application of ChatGPT has intensified, thus making it necessary to define rules and procedures with which this technology would contribute to teaching and learning. In any case, robot autonomy sounds like science fiction even today, although autonomous robots have already become a reality. Our authors show that robot autonomy is not an unreasonable idea, but quite a realistic one, which can be elaborated in detail and achieved by constructing information systems which can act autonomously, although due to potential abuse it is necessary to think seriously about their application. Through an interdisciplinary approach that combines art, technology and social sciences, in this edition we also give insight into the potential and challenges brought by the integration of artificial intelligence into art. Particularly in the context of creative Serbia.

Thanks to technology democratization and good-quality infrastructure which has been constructed in the meantime, artificial intelligence has recently become available to the broadest population. The main domains of AI technology application in healthcare can be seen on a daily basis in clinical practice and clinical research, but also in the production of new medications, personalized medicine, public health and medical administration.

More specifically, the researchers from the Institute for Artificial Intelligence of Serbia (IVI, 2024) are implementing a number of projects in the field of healthcare, including the diagnostics of cancer and rare diseases. The aim is to ensure faster and more simple carcinoma detection, to reduce the complexity of diagnosing rare diseases and shorten the time until giving an accurate diagnosis, to alleviate the pressure on the healthcare system and, consequently, redirect medical treatment costs towards an adequate innovative therapy, as well as accelerate the process of finding and developing new medications.

The fear of artificial intelligence initiated by religious feelings is one of the perspectives which we approach, aware that every new technology also raises the question of the morality of its application. Believers' opinions are divided, which is not surprising since we know that the omnipresence of the phenomenon of fear in Christianity has historically produced certain resistance and doubts. The most frequent reason for it is insufficient knowledge of what artificial intelligence is and what its place is in modern society. As for religious communities, huge responsibility in the attitude towards artificial intelligence is actually on the clergy and its willingness to speak about it to people who have been entrusted to them for religious guidance.

Ny choosing to dedicate our thematical edition to artificial intelligence, we have also presented the attitude of our editorial board. In the past fifty years, the world has seen more changes than in the previous fifty thousand years. This acceleration is difficult to understand, the changes are radical and extreme, and there is no room for all of us on the planet... In the next two years (until the end of 2026), Serbia will invest about a hundred million Euros in the development of artificial intelligence, out of which thirty million will be invested in the supercomputer that will be free of charge for researchers and startups, while the same amount of money will be allocated for the development of artificial intelligence and its broader application. The government's strategy and orientation are clear, the gates to the extreme future have been opened and it is up to us only to take a firm step forward!

Articles

Stefan R. Badža[1]
Council for Artificial Intelligence
Government of the Republic of Serbia
Adviser to the President of the National Assembly
Belgrade (Serbia)

UDC 004.8(497.11)
005.21:004.8(497.11)
Review scientific article
Received: 01.08.2024.
Accepted: 09.08.2024.
doi: 10.5937/napredak5-52537

Development of Artificial Intelligence in Serbia. Serbia as the Regional Leader

Abstract: Artificial intelligence (AI) is transforming industries and social spheres, and Serbia has come forward as the leader in Southeast Europe. By adopting the National Strategy for Artificial Intelligence in 2019 and founding the Institute for Artificial Intelligence in 2021, Serbia laid strong foundations for the development of this technology. The introduction of AI in the educational system, the development of the research infrastructure and active participation in international initiatives have contributed to important results. Ethnical guidelines adopted in 2023 constitute the basis for accountable AI application. Being elected Chair of the Global Partnership on Artificial Intelligence for 2024 further confirms Serbia's position. Continued support of the Government, educational institutions and the economic sector is crucial for the future development of AI technologies.

Keywords: artificial intelligence, Serbia, National Strategy, Institute for Artificial Intelligence, Ethical Guidelines, international initiatives

Introduction

Artificial intelligence (AI) is one of the most important technologies today, which transforms industries and social spheres. AI covers a wide range of technologies, including machine learning, natural language processing, computing vision and robot systems, thus enabling machines to perform tasks which traditionally required human intelligence. This technology has the potential to improve efficiency and productivity, to improve decision-making and create new economic opportunities. Many countries have recognized the importance of AI and invest in its research and development, while Serbia has come forward as one of the leaders in this sphere in Southeast Europe.

According to the definition used by the European Commission, AI refers to systems which display sensible, intelligent behaviour based on the analysis of their environment and take actions with

[1] stefan.badza@parlament.rs

some degree of autonomy to achieve specific goals. AI-based systems can be purely software-based and act in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones etc.) (European Commission, 2018).

Methodology

The analysis in this article is based on official documents, government reports, scientific papers and statements of experts in the field of artificial intelligence. The data have been collected from national and international sources, including the publications of OECD and the Global Partnership on Artificial Intelligence (GPAI). The primary sources have been used from Serbia's archives, as well as secondary sources from scientific papers and books.

Results

This strategy, covering the period 2020–2025, laid strong foundations for the development of AI through defining key goals and priorities. The goals of the Strategy include the development of research infrastructure, encouragement of innovation, AI integration in different sectors, as well as the development of human resources through education and training (Strategy for the Development of Artificial Intelligence in the Republic of Serbia for the period 2020–2025, 2019).

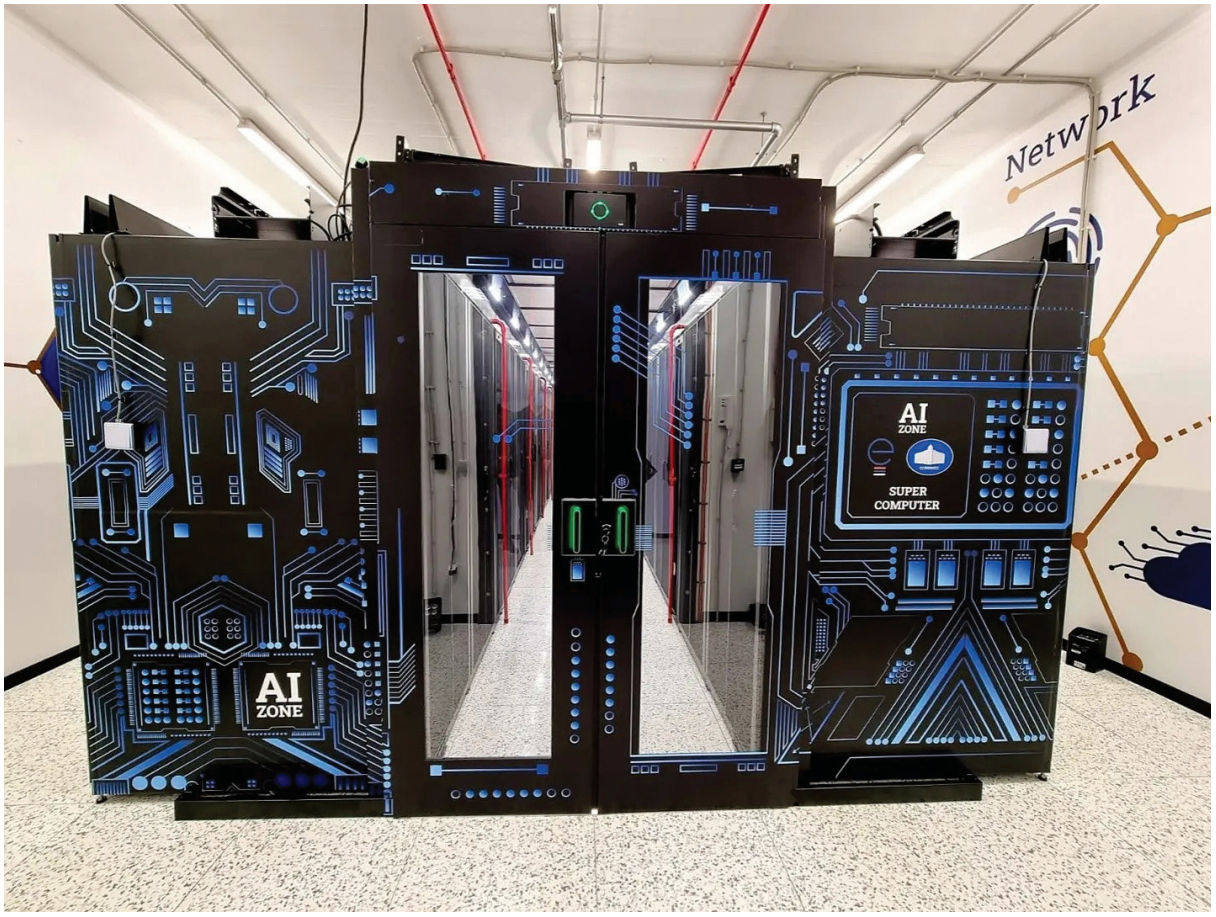
In 2021, Serbia made an important step in the educational system by introducing artificial intelligence in the curriculum of primary and secondary schools. By introducing AI as part of two subjects in primary schools and an optional subject in secondary schools, Serbia made it possible for young generations to become familiar with the rudiments of this technology and its applications from the earliest age. At the end of 2021, Serbia was one of 11 countries in the world which, according to the UNESCO analysis, introduced AI in primary and secondary education at the state level (UNESCO, 2021). Moreover, it was in the same year that the first Institute for Artificial Intelligence in Southeast Europe was founded, which plays the key role in the research and development of AI technologies. Today this Institute has more than 50 scientists involved in the following areas – language processing, computer vision, generative artificial intelligence etc.

In 2019, Serbia became the first country in Southeast Europe to adopt the National Strategy for the Development of Artificial Intelligence in the Republic of Serbia.

In 2022, Serbia continued strengthening its positions on the global scene through seven new master programs in the field of artificial intelligence, initiated at six faculties and four universities. These programs provide students with an opportunity to acquire advanced skills and knowledge which are crucial for further AI development and application. In 2022, Serbia also became part of the Global Partnership on Artificial Intelligence (GPAI), which enabled it to participate in global initiatives and exchange experiences with the leading experts and

institutions. Serbia thus became part of the elite number of only 29 countries within the GPAI, which was founded at the initiative of the French President Macron and the Canadian Prime Minister Trudeau.

Serbia's national AI platform, the supercomputer which was made available free of charge to universities, faculties, institutes and domestic start-up companies, was rated by OECD as one of the most innovative projects in the public sector in 2023 (OECD, 2023). This platform was ranked among best 10 innovative projects, out of as many as 1,048 projects from 94 countries. At the Global Summit in Dubai, the Government of Serbia was recognized as one of the most innovative governments in the world, while the initiative to enable the work of startups on the National AI platform was rated as



The National Artificial Intelligence Platform (supercomputer) in the State Data Centre in Kragujevac.
Photo: Office for Information Technologies and Electronic Administration

one of the most innovative ones (Global Partnership on Artificial Intelligence (GPAI), 2024).

In the course of 2023 Serbia adopted the Ethical Guidelines for Development and Implementation of Artificial Intelligence, ensuring that AI technologies are used in an accountable and ethically acceptable manner. These Guidelines laid the foundation for further regulation of this field and promoting principles of transparency, accountability and protection of privacy (Ethical Guidelines for Development, Implementation and Use of Robust and Accountable AI, 2023).

The year of 2024 has brought another great recognition for Serbia because it will chair the Global Partnership on Artificial Intelligence (GPAI). This organization, which has been integrated in OECD and increased its membership to include 45 countries, will hold a large conference in Serbia, with the participation of several thousand experts from the AI field. According to the plan, the conference will be attended by the representatives of more than 50 countries at the ministerial level, which will position Serbia as the central venue for discussion about the future of AI.

Discussion

Serbia has succeeded in positioning itself as the leader in the field of artificial intelligence in Southeast Europe through strategic initiatives and investments. Continued investments in education, research and development, as well as the integration in international initiatives, have enabled Serbia to achieve important results in a relatively short period. The key factor of success lies in the synergy between government initiatives, educational institutions and the private sector.

The applications of good practice include the organization of the Simulation Centre at the Faculty of Medicine, the University of Belgrade, which uses virtual and expanded reality for training medical staff. GovTech Program gathers the representatives of the public sector and economy for the purpose of identifying problems which can be resolved by the application of AI. The examples of AI application include "read to me" function on the website of the Government of Serbia, the digital avatar at the Prokop railway station for helping persons with impaired hearing, and the "Hawk Eye" system for parking surveillance in Belgrade.

Serbia has also taken first steps regarding the possibilities of applying artificial intelligence in the field of autonomous vehicles. The Law on Bases of Traffic Safety on Roads was amended by introducing the concept of an autonomous vehicle (Law on Bases of Traffic Safety on Roads, 2009), the Rulebook on the Conditions for Conducting Autonomous Driving was enacted (Rulebook on the Conditions for Conducting Autonomous Driving, 2023) and the Rulebook of Technical Inspection of Vehicles was amended (Rulebook of Technical Inspection of Vehicles, 2012), which enabled the examination and issuance of licenses for testing these vehicles up to Level four in real traffic.

In the previous period, Serbia achieved significant progress in the field of education regarding AI. Two optional subjects were introduced in primary schools, while there are three optional subjects available in secondary schools. At a number of faculties there are study programs which fully or partly deal with AI. In addition, the Centre for Robotics and Artificial intelligence in Education was founded at the University of Belgrade. The National Academy for Public Administration organizes training for all public servants dedicated to AI-related topics.

Serbia has also recognized the importance of developing infrastructure necessary for AI. A supercomputer platform for artificial intelligence has been established in the State Data Centre, which is used for the development and application of AI technologies. In 2023, this platform was rated by the OECD as one of best 10 innovative projects in the public sector (OECD, 2023).

Furthermore, the improvement of infrastructure, as well as data availability, are crucial for the development of AI. Currently in Serbia there are more than 2,500 data sets available at the Open Data Portal. The adoption of the Electronic Administration Act in 2018 made the public entitled to use data once again for commercial or non-commercial purposes (Electronic Administration Act, 2018).

The legal framework for AI development in Serbia includes the adoption of ethical guidelines prepared in line with the recommendations of the UNESCO and the European Union. Serbia actively participates in the work of international bodies, such as the Global Partnership on Artificial Intelligence (GPAI). The preparation of the legal and institutional frameworks is an important condition for further development and application of artificial intelligence. It is a mechanism which will organize the activities related to artificial intelligence and at the same time define the framework enabling Serbia to use its specific features while also being in compliance with international rules and principles.

Some of the key challenges are the proper development and application of this technology, which includes safe, secure and reliable AI in all life cycle stages. Furthermore, it is necessary to inform and educate all members of society about the possibilities and risks brought about by artificial intelligence. In line with the above-mentioned need and the fact that the nature, conditions, manner of development, possibilities and manner of application of artificial intelligence should be properly presented to all members of society, there is an idea that it is necessary to further improve promotion and education. Apart from advantages, for the purpose of proper use and protection of all subjects, promotional campaigns should also point out and inform about potential challenges and risks, as well as manners of overcoming those risks.

In this stage of the development of artificial intelligence, it is necessary to include to a larger extent business enterprises, as well as professionals from the non-engineering fields. Artificial intelligence requires a comprehensive approach and it is not necessary to develop only engineering capacities, but also the regulatory and usable ones so as to enable the proper use of what has been developed.

It is particularly relevant that we develop language technologies in order to protect our language and culture. In that respect, attention is dedicated to finding solutions for collecting, systematizing, making available for use and further development of resources for the use of the Serbian language, both written and spoken, and potentially of video content. It seems quite important to include in the process all creative industries and cultural and national institutions so that we can together preserve the Serbian language, Serbian culture and Serbian worldview.

Concluding considerations

Serbia has achieved significant progress in the development of artificial intelligence through strategic initiatives, investments in education and research,

as well as active participation in international initiatives. Continued support of the government, educational institutions and the economic sector is crucial for further development and application of AI technologies in Serbia. The challenges include the need for further regulation, education and development of infrastructure to ensure safe and ethically acceptable application of artificial intelligence. With further investments and cooperation, Serbia has the potential to become the leader in the field artificial intelligence in the region and on a larger scale.

In the future, the focus should be on further integration of AI in the educational system, larger investments in research infrastructure and strengthening international cooperation. Moreover, it is necessary to develop ethical and legal frameworks which will ensure accountable application of AI, taking into account potential differences and challenges.

Documents and Legal Regulations

Electronic Administration Act. 2018. Official Gazette of RS – No 25/2018.

Ethical Guidelines for Development, Implementation and Use of Robust and Accountable AI. Official Gazette of RS – No 23/2023

Law on Bases of Traffic Safety on Roads 2009. (2009). Official Gazette of the RS, No 50/1988-1337, 63/1988-1604, 80/1989-1987, 29/1990-1014, 11/1991-181, 34/1992-645, 13/1993-284 16/1993-333, 31/1993-685, 41/1993-857, 50/1993-1020, 24/1994-29, 41/1994-573 ,28/1996-5 (3/2002-44, 101/2005-28, 41/2009-3

Rulebook of Technical Inspection of Vehicles (2012.). Official Gazette of RS - No 8/2012-123, 13/2013-51, 31/2013-76, 114/2013-291, 40/2014-5, 140/2014-38, 18/2015-96, 82/2015-21, 88/2016-55, 108/2016-48, 129/2021-139, 83/2023-21, 7/2024-32, 55/2024-14

Rulebook on the Conditions for Conducting Autonomous Driving (2023). Official Gazette of the RS, No. 104/23.

Strategy for the Development of Artificial Intelligence in the Republic of Serbia for the period 2020-2025. (2019). Official Gazette of the RS – No. 96/2019-5

Sources

Global Partnership on Artificial Intelligence (GPAI). 6th meeting of the GPAI Ministerial Council – New Delhi. (2024). Available at: https://pib.gov.in/PressReleasePage.aspx?PRID=2030534

Official Page of European Commission (2018). A definition of AI: Main capabilities and scientific disciplines, Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission. Available at: https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

Official Page of OECD (2023). OECD Science, Technology and Innovation Outlook 2023. Available at: https://www.oecd.org/en/topics/policy-issues/government-innovation.html

Official Page of UNESCO (2021). Global education monitoring report, 2021, Central and Eastern Europe, Caucasus and Central Asia: inclusion and education: all means all: summary Available at: https://unesdoc.unesco.org/ark:/48223/pf0000377713_srp

Ivana M. Luknar[1]
Institute for Political Studies
Belgrade (Serbia)

UDC 004.8
Review scientific article
Received: 05.08.2024.
Accepted: 19.08.2024.
doi:10.5937/napredak5-52577

Artificial intelligence as a challenge

Abstract: This paper addresses the technology of artificial intelligence as a modern challenge. The main purpose of this article is to motivate all interested parties, the government leaders, scientific and expert public, lawyers, innovators, as well as end users to consider the challenges deriving from the application of this technology. Interpreting the technology of artificial intelligence from the perspective of different disciplines and practices contributes to its better understanding, while emphasizing a broad range of risks is a step forward in identifying and understanding the problems which we must readily encounter. If possible, some of the problems should be avoided and prevented in a timely manner. The main goal of the paper is to point to the challenges deriving from the application of this modern technology and raise awareness of the importance of organizing production and application of artificial intelligence technology. The paper starts from the hypothesis that knowing and resolving the broadest possible range of challenges is a prerequisite for regulating this modern technology for the common good. The article provides the analysis of content and a detailed review of literature dealing with artificial intelligence from the perspective of different scientific disciplines. The importance of the study is seen in its being a useful stimulus for initiating a debate about this topic. The social contribution of this study is huge because it puts forward the question of the challenges of artificial intelligence and leads to the strengthening of the awareness of the need for identifying and considering all potential challenges for the sake of their resolution. It is in the interest of whole humanity to focus on regulating artificial intelligence, with the primary focus on the achievement of general social welfare, and not on mere acquisition of profits or some other interest advantages. Only in this way can we protect ourselves from potential harmful consequences of this modern technology.

Keywords: artificial intelligence, challenges

Introduction

The question of challenges deriving from the production and application of artificial intelligence is one of today's most important problems. Combining physical space with cyberspace creates new virtual space for mutual cooperation of people with the aid of artificial intelligence systems, but also

[1] ivanaluknar@gmail.com ; https://orcid.org/0000-0003-4867-2263

for forming a specific relationship between people and AI systems. Such complex environment, made by man and modern technologies, creates new forms of connection and appears as a source of new complex risks and "cyber threats" (Luknar & Jovanović, 2024). Artificial intelligence technology has found us relatively unprepared having in mind that it has already been applied to a certain extent in different branches of industry throughout the world, while there is no aligned legislative framework regulating this technology. Technological development easily stays away from legislative regulation due to insufficient flexibility of the legal system and the faults arising because of the need for harmonizing and fitting certain regulations with the entire legislative system of a country. Moreover, procedures take time to be established in practice. Due to the absence of adequate legal regulations, ethical guidelines have been proposed which refer to the production and application of this modern technology (Luknar & Jovanović, 2023).

Developed technologies change human life in a revolutionary rapid manner, becoming part of our everyday life and affecting the way in which we perform everyday activities, and that is why there is a discrepancy between technological development and its legal regulation. This situation requires urgent action and imposes the need for regulating the further course of the development and application of artificial intelligence which can cover the broadest range of risks. The ruling structures, scientific and academic workers, lawyers, experts, innovators and all other interested parties need to provide an overview of this problem because only in this way is it possible to reach a receptive and efficient comprehensive framework for regulating artificial intelligence. First it is necessary to strive for a general model of regulation which primarily refers to the production and application of artificial intelligence and is in line with main social goals, whose usefulness and efficiency are not achieved at the expense of human rights, freedoms and security. Furthermore, this model needs to be aligned with specific requirements because artificial intelligence technology is applied in different forms for performing specialized tasks in different branches of industry. In this manner, with the deductive method, it is possible to further regulate the specialized production and application of artificial intelligence and to provide protection from its abuse in the given specific environment. The first step in adequate regulation of artificial intelligence is as detailed as possible identification of challenges and facing them in the most efficient manner. Having in mind the complexity of this technology and the various applications of artificial intelligence systems, a multidisciplinary approach is necessary in providing an answer to the researched question.

Artificial intelligence technology

The first step in regulating a certain social phenomenon is its understanding, or defining. Artificial intelligence technology is a complex system of different technological tools with a certain specialized application for which the system has been designed. The most developed artificial intelligence systems try to behave like a live organism, i.e., to emulate cognitive functions possessed by man. Namely, this technology offers a broad range of functions and possible applications, which may cause confusion

in its theoretical determination. Concurrently with the development of artificial intelligence, its different definitions were created. In order to understand this technology fully, it is necessary to take into account its basic characteristics, i.e., its technological performances. Hence it can be defined from the perspective of computer sciences or other disciplines, depending on the primary focus, i.e., on whether we interpret it mainly from the perspective of the degree of its developed intellectuality or the field of practical application, such as support to the functioning of crucial state infrastructure, the work of public administration, for communication, in the field of culture, entertainment etc. In addition, we must not neglect the fact that this technology is developing and becoming more complex with incredible speed, which brings further ambiguities.

It is frequently wrongly made equal to machine learning, which constitutes only one of its segments. Artificial intelligence includes different types of technological systems which function within previously set logic and regularity; systems working by the model of perception among people and select data from the base, handle a large amount of data, discover regularities and make conclusions or certain kinds of predictions based on identified regularities. All the above-listed is included in artificial intelligence technology. This technology was primarily made with the aim of substituting man in all those activities which can be performed by a machine more precisely, more accurately, faster, more efficiently and less expensively, for example, different mathematical and statistical operations used in electronic banking or trade, data collection and weather forecast, complication and simplification of data about soil composition for the needs of smart farming, faster and more simple identification of anomalies and changes on human skin, and many other activities.

In modern literature, different types of artificial intelligence (AI) are mentioned, which can to a certain degree help in understanding the scale of the potentials of this technology. The lowest level of artificial intelligence which has developed ever since its emergence is *Artificial Narrow Intelligence*, also known as "weak artificial intelligence", designed to perform the simplest tasks. The next type is *Artificial General Intelligence*, also known as "hypothetical AI", and it can perform any intellectual task like man, for example, learning, understanding and application of certain knowledge. This level of artificial intelligence is still being developed and perhaps it will not be possible to achieve its equal development in all domains. Hypothetical AI is the level of development given as an antipode to "weak AI". The next type of AI is *General Purpose AI*,[2] which is mentioned in the legally binding Artificial Intelligence Act of the European Union (EU Artificial Intelligence Act, 2024). General Purpose AI refers to machines which are generally able to perform a broad range of intellectual tasks by emulating man's cognitive abilities. Unlike specialized AI, which is designed to perform specific tasks, this level of intelligence tries to emulate human intelligence, i.e., it is simi-

[2] In speaking practice, namely in the Serbian language, there is no established translation of this term. For the purpose of this research, the author uses the translation which most closely reflects the analyzed concept. For this reason, the original names have been supplied in brackets in the Serbian version of the paper.

lar to human intelligence in performing activities (Pinto, 2024). This type of intelligence includes *Open AI GPT*, *Dolly AI*, *Claude AI by Anthropic*, *Meta's Liama* and other similar AI systems. Artificial intelligence systems which were presented at the beginning of the 2020s are a step ahead of "weak AI", but they have still not reached the level of "strong AI", so that they are between these two levels of the development of artificial intelligence (Figure 1). The last development level of artificial intelligence *Artificial Super Intelligence*, which not only emulates and replicates general (hypothetical) AI, but is also placed in the category of science fiction (Jones, 2024). Bostrom has written in further detail about this type of AI, the so-called "super intelligence" (Bostrom, 2014).

Mass digitization has helped the expansion of this technology in everyday life through smart devices, virtual assistants, software that enables service personalization, personalized search, photograph analysis, face and speech recognition, and many other functions. Machine intelligence comes as a result of complex processes such as data collection, handling available information and data bases ac-

Figure 1. Development stages of artificial intelligence

cording to the previously programmed goals/logic which serve for resolving certain tasks. That this is a complex technology is also supported by the fact that the data collection process may proceed in several different ways, either via sensors, by direct entering of instructions, or from the communication with other machines. The authors point to the characteristics which a device must have in order to be considered smart:

"• possibility of processing natural language for performing successful communication in English;

• presenting knowledge based on stored information;

• automated reasoning and using kept information when answering posed questions and making new conclusions;

• ability of machine learning to adjust to new circumstances and reveal and extrapolate patterns" (Russel & Norvig, 2009, p. 2).

This modern technology has the potential to substantially supplement and facilitate the performance of different human activities. However, the problem is that the existing designed systems might be insufficiently flexible in the recognition of ethical issues and problems, while at the same time unsafe because of the absence of adequate legal regulations referring to the design and application of artificial intelligence technology.

ChatGPT
The photo was generated by AI images of the generator on Freepik Picasso platform.

Challenges deriving from the development and application of artificial intelligence

Technology constitutes an important lever of power in today's society. Artificial intelligence is understood as a technological imperative in the most developed countries of the world. At the international level, there is an ongoing raced in technological development and innovation which give a competitive advantage, but at the same time confirm the position of a country in the system of power. Artificial intelligence takes up a central position in the arena of balance of international powers and conflict of interests. The countries worldwide (the USA, Russia, United Arab Emirates, India, China, Serbia and many others) have included the development of artificial intelligence technology in their development strategic plans as one of the primary goals. State industrial and trade policies have a direct impact on the development, production and application of the basic equipment which is necessary for the production and application of artificial intelligence. international competition regarding artificial intelligence is manifested in practice mostly through competition in the supply chain of the products which ensure greater computing strength and through (non)availability of tools and performances which provide necessary support in the functioning of this technology.

Apart from achieving competitiveness at the international level, we must not overlook the market structure and competition among companies at the national level, which encourages technological development, primarily in a competitive, profit-oriented spirit. This situation may push to the background the essence of the technology, which should serve a general interest and not mere acquisition of profits or achieving other benefits deriving from technological development. There is a risk, due to the rivalry, of the worst scenario, i.e., that artificial intelligence might continue developing autonomously, without human surveillance. People have already invented the technology (simulation learning, independent play and meta-learning) which could turn the development of artificial intelligence into an undesirable direction, thus making it substantially harder for people to monitor the development and control AI. Although the primary purpose of the self-

learning method is to increase the efficiency of machine learning, there is a danger of gradual reduction of the importance of human participation in the creation and work of smart machines. Seen in the long run, due to the regular application of AI systems, trust in the work of this technology is being built. There is a risk of beginning to take AI "for granted", i.e., that the correctness of its outputs will not be considered at all with the passage of time, but that it will be unquestioningly accepted as truthful, while the correctness of its content will not be disputed whatsoever. We can say that this modern technology is like a double-edged sword.

The technology is not constant and/or unchangeable, and that is exactly why an active approach is required to its interpretation and regulation. We must be aware of all the challenges deriving from the production and application of this modern technology because only in this way will we learn where to direct attention during its regulation. The key challenges being related to artificial intelligence technology are:

1. production and design – the key challenge is how to make its usefulness larger than its harmfulness;
2. regulations – the key challenge is how to find best solutions, through law, policy, procedures and regulations, to the regulation of the fields of the production, design, application and development of artificial intelligence technology;
3. safety and correctness of the content of data and bases over which AI systems operate;
4. the matter of ownership over systems and data bases, determining responsibilities and sanctions for certain violations and abuses;
5. observance and implementation of general ethical guidelines in design and manner of functioning of AI systems, as well as human rights and freedoms, and anti-discrimination principles.

Thinkers throughout the world warn about the broad range of risks being related to artificial intelligence technology. No technology invented by man has ever caused such concern for the existential survival of humanity such as artificial intelligence has. Even the best-known computer engineers have expressed their concern. Geoffrey Hinton, who worked on the development of modern technologies and artificial intelligence, resigned from Google Company because he found it necessary to point to the danger and risks of this technology for society. Hinton states that machines will soon be able to think better than us and decide to take the helm over our society, which should particularly concern us, and that is why we need to devise how to prevent that from happening (Allyn, 2023). Elon Musk also points to the harmful effects of this modern technology and considers artificial intelligence "the most destructive force in history" (Times Now, 2023). Opinions about this technology vary drastically, causing both fear and ultimate rejection, pointing to its harmful consequences, as well as to its full acceptance. World leaders share opinions about artificial intelligence as an unavoidable technology of the future. However, serious challenges derive from the automation of weapons and the application of this technology in different military operations (Kissinger, Schmidt & Huttenlocher, 2023).

The prerequisites for the development of a country in modern era are its flexibility and willingness to accept innovation. Apart from indisputable advantages deriving from the application of artificial intelligence technology, numerous questions arise as well. Artificial intelligence technology in its work

uses information available to it. That is why the questions referring to data are the first to arise, such as issues of data privacy and protection, data quality and completeness, data availability and security, fairness, transparency etc. (Donald et al., 2023).

Artificial intelligence technology may reproduce the existing socio-economic inequality, affect capital flows, but also contribute to the deepening of the gap between the rich and the poor. The poor remain in the enchanted circle of poverty because of insufficient technological development and not possessing necessary technological tools and expertise in the field of computer sciences, and that is why they are forced to perform poorly paid jobs. Moreover, artificial intelligence can be a cause of numerous changes in the labour market and affect market volatility in a negative manner. Automation and robotization can lead to the disappearance of certain professions. AI has the potential to affect significantly supply and demand of certain types of professions (Goos, Manning & Salomons, 2014). In addition, other observed risks include the violation of privacy and algorithmic partiality due to inaccurate and insensitive data regarding different forms of discrimination. The application of artificial intelligence technology in education may serve as significant support in learning, but only if this technology is applied "fairly and effectively" (Mandić, Miščević & Bujišić, 2024) in education. Furthermore, there is concern about the development of this technology's performances and its increasing use might minimalize the human aspect and push to the background the human element in creating modern technologies. As one of the challenges, the question arises about how to stop the development of uncontrollable and self-conscious artificial intelligence. Moreover, another challenge is excessive reliance on artificial intelligence systems, which may lead to inertia and excessive dependence of people on AI services even in the performance of the simplest tasks, while easy availability of digital assistants may lead to the uncritical reliance on the work of AI tools. It is well known that intelligence in humans develops through their everyday encounters with certain problems/tasks, or decision-making. Blind trust in the validity of AI functioning, seen in the long run, may lead to the inability of new generations to solve the simplest tasks on their own because they are used to relying on AI tools. Apart from the above-listed challenges, there are also problems of consumption and demand of certain materials used in the production of AI tools. The challenge to be resolved is also what to do with the technological waste of non-biological nature which, together with the already-known climate change and problems, does not guarantee a bright future, unless artificial intelligence tools have been applied for finding the most efficient ecological solution.

The risks deriving from artificial intelligence technology can be summed up into certain risk groups. The first group is made of risks deriving from the lack of transparency and clarity. Artificial intelligence systems can be difficult to understand even to those who use this technology, while data over which AI performs operations are unclear and insufficiently transparent. The second group of risks includes economic risks and those referring to the marked due to the process of automation and robotization. The third group consists of social risks, which refer to the manipulation, threatening people's privacy and safety (*deepfakes*), political manipulation, cultural pressures, weaking of the human element etc.

Conclusion

Artificial intelligence is considered a vital strategic technology. In modern social-political conditions, this technology takes up the central place in the international balance of power and conflict of interests. However, the current regulation of modern technologies is already lagging behind worldwide. The first law which specifically refers to the technology of artificial intelligence was adopted by the European Union at the end of 2023. This law was preceded by numerous directives and ethical guidelines which were not legally binding. First the international community needs to harmonize the general legal framework for the primary regulation of production, application and distribution of this modern technology, which should comprehensively cover a broad range of challenges and problems presented in this paper.

References

Allyn, B. (2023). *'The godfather of AI' sounds alarm about potential dangers of AI*, NPR. Available at: https://www.npr.org/2023/05/28/1178673070/the-godfather-of-ai-sounds-alarm-about-potential-dangers-of-ai

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. UK: Oxford University Press.

Donald, D., Aditya S. S., T., Rekha, K., Anjali, D., Dwaraka S. (2023). The Data Revolution: A Comprehensive Survey on Datafication. *International Journal of Advanced Research in Science, Communication and Technology* III (1), 41–51. DOI:10.48175/IJARSCT-9007

EU Artificial Intelligence Act (2024). Up to date developments and analysis of the EU AI Act. Available at: https://artificialintelligenceact.eu/

Goos M, Manning, A., Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, CIV (8), 2509–2526. DOI: 10.1257/aer.104.8.2509

Jones, B. (2024). *AI literacy fundamentals*. Washington: Data Literacy Press.

Kissinger, H., Schmidt, E., Huttenlocher, D. (2023). *The Age of Artificial Intelligence and Our Human Future*. Beograd: Club Plus. [In Serbian]

Luknar, I. & Jovanović, F. (2024). Various types of cyber threats. *Srpska politička misao*, LXXXIII (1), 161–177. DOI: 10.5937/spm83-46059 [In Serbian]

Luknar, I. & Jovanović, F. (2023). An Ethics for Emerging Technologies. In: Đ. Jovanović et al (eds) *2nd International Conference Proceedings "Conference on Advances in Science and Technology COAST 2023"* (1046-1056). Montenegro, Herceg Novi: Faculty of Management.

Mandić, D., Miščević, G., Bujišić, Lj. (2024). Evaluating the quality of responses generated by ChatGPT. *Methodology Theory and* Practice, XXVII (1), 5-19. doi:10.5937/metpra27-51446

Pinto, T. (2024). General-purpose AI under the AI Act. Available at: https://artificialintelligenceact.com/general-purpose-ai-under-the-ai-act/

Russell, S. & Norvig, P. (2009). *Artificial Intelligence: A Modern Approach* (3rd Edition). UK: Pearson.

Times Now (2023). Elon Musk raises alarm bells: AI – 'The most destructive force in history!' Available at: https://www.timesnownews.com/technology-science/elon-musk-says-ai-the-most-destructive-force-in-history-article-104942193

Dragan V. Vukmirović[1]
University of Belgrade,
Faculty of Organizational Sciences
Belgrade (Serbia)

Dejana S. Kresović[2]
Securities Commission
Belgrade (Serbia)

UDC 330.341:004.8
004.8
Original scientific article
Received: 09.07.2024.
Accepted: 22.07.2024.
doi: 10.5937/napredak5-52069

The transformative potential of generative artificial intelligence

Abstract: This paper analyses the transformative potential of generative artificial intelligence at macro, meso, and micro levels of social and economic structures. The aim is to determine the impact of these technologies on various aspects of society and economy, including business operations and the labour market. The potential of new technologies to increase productivity, transform business models, and create new professional roles has been examined through a comprehensive analysis of data and studies. It has been concluded that generative artificial intelligence can fundamentally change the labour market, globally increase gross domestic product, and improve both the public and private sectors. The paper provides insights into future trends and regulatory and structural changes that are necessary for optimising the application of generative AI.

Keywords: generative artificial intelligence (GenAI), transformative potential - macro, meso, and micro

Introduction

Society is globally witnessing accelerated development in the field of artificial intelligence (AI). This progress includes the implementation of technologies that provide machines with the ability to learn and perform cognitive tasks traditionally a matter referring solely to humans. This technological progress could have significant consequences on society and culture. Since AI is a cognitive technology, its impact pervades the crucial domains such as education, science, culture, and communication. AI-based systems are increasingly advising medical doctors, scientists, and judges (UNESCO, 2019).

Starting from the 1950s, there has been a multitude of papers that deal with the technical aspects of AI and AI applications in general. For example, a paper by Talib et al. (2020), provides a systematic overview of references that deal with hardware implementation of AI algorithms and machine

[1] dragan.vukmirovic@fon.bg.ac.rs ; https://orcid.org/0000-0003-0248-016X
[2] dejana.kresovic@sec.gov.rs ; https://orcid.org/0009-0008-9798-0733

learning (ML). The authors analysed 169 research papers published during the period between 2009 and 2019, whose focus was the implementation of neural networks as a tool for detecting objects in various applications (Talib et. al., 2020).

The research into the above-mentioned references also indicates that the published papers pay relatively little attention to a comprehensive consideration of socio-economic consequences that underlie AI. There is an obvious research gap in terms of understanding the impact of AI on social dynamics and economic structures, especially in the context of labour, education system, and policies globally. Stahl et al. (2023) note that their paper was the first to present a systematic overview of references that deal with AI impact assessment. The authors analysed 38 documents that address impact assessment so as to understand the purpose, extent, organisational context, expected issues, timeframes, processes and methods, transparency, and challenges. They concluded that there was a certain degree of convergence between various AI impact assessments, but that there was still no full consensus in terms of content, structure, and implementation. Impact assessments can be best understood as a means of encouraging thinking and discussion on social and ethical consequences of the AI ecosystem. The authors do admit that there are certain limitations of their study, including difficulty in defining AI and defining the extent of impact assessments and they also emphasise a need for additional research in order to better understand their impact and roles in a broader AI ecosystem. This paper provides an important contribution to understanding impact assessments and their potential role in AI management. Its findings are relevant to researchers, policy makers, and organisations in the field of AI.

The main feature of GenAI is that it has found its place in the area of intellectual, creative and generally better paid jobs, which stands for a true socio-economic revolution, or, to say the least, evolution.

To date, every industrial revolution has been marked by technology dominance in manufacturing, it mainly being replaced by manual labour, be it the steam engine or industrial robots (Fahle, Prinz & Kuhlenkötter, 2020). What makes GenAI special is its potential to affect jobs that require intellectual and creative skills. Not only does this significant change bring about technological innovation, but it also has a deep impact on the socio-economic context, which makes this technology extremely disruptive.

This is confirmed by a growing number of scientific and specialised papers that investigate GenAI as a disruptive technology. These papers cover a range of aspects of the application of GenAI across industries as well as its impact on business operations and social structures.

For example, the study published in the Nature Reviews Urology journal discusses how GenAI can transform science publishing, by asking the question whether this technology is disruptive or destructive for the existing work methods in this sphere. The study shows that GenAI can expedite the process of innovation and improve the quality of papers, but at the same time it also brings challenges regarding work authenticity and integrity (Bertolo & Antonelli, 2023).

Similarly, the MIT Technology Review reports on the impact of GenAI on business functions such

as product innovation, supply logistics, and customer experience. This paper points out that not only does GenAI change the manner in which certain jobs are performed, but also the approach to innovation and competitiveness taken by organisations in their respective industries. The report shows that the majority of managers expect a substantial change to occur in the next five years, but at the same time it notes that there are technological and organisational obstacles that may interfere with a successful implementation of GenAI (MIT Technology Review Insights, 2024).

The document *The GenAI is out of the bottle: generative artificial intelligence in various domains*, which was published in Springer, analyses the application of GenAI across industries and its potential to cause significant change in business operations and society. Through six underlying hypotheses, the authors consider how GenAI democratizes the approach to knowledge, combines factographic knowledge and creative thinking, and changes the skill set necessary for creating content. Three industries are analysed - software engineering, healthcare and financial services - in order to demonstrate potential changes in the business models. This document underlines the rapid growth and potentials of GenAI, but also notes challenges such as ethical issues and the need for it to be regulated. In its conclusion, this paper contributes to understanding how companies can use GenAI for innovation and improvement of their business models (Kanbach, Heiduk, Blueher, Schreiter & Lahmann, 2024).

The article *How Generative AI Will Transform Knowledge Work*, published in the Harvard Business Review, deals with how GenAI can transform knowledge-based jobs and what challenges there are in GenAI application. GenAI has the potential to automate certain knowledge-based tasks, but this does not entail the replacement of all workers. On the contrary, GenAI can allow the workers to have more time that they could use for performing more important tasks, thus improving their performance and productivity. The text explains in detail the ways in which GenAI can support knowledge-based work, including reduced cognitive burden through the automation of structured tasks, improvement of cognitive abilities for non-structured tasks, as well as the streamlining of the workplace learning process. The authors also deliver recommendations to managers on how to assist their workers in terms of the optimal use of GenAI (Alavi & Westerman, 2023).

The authors of the article *How to Capitalize on Generative AI* (McAfee, Rock & Brynjolfsson, 2023) discuss the importance and risks of GenAI versus its potential benefits. The example of a large software company demonstrates how GenAI can improve efficiency and customer satisfaction, especially through assistance to customer support agents. GenAI should be viewed as a general-purpose technology, akin to electricity or the internet, which will have a rapid impact on economy, thanks to the already existing infrastructure. People can communicate with these systems easily, which lowers the barriers for their use. The authors recommend that business leaders should take steps such as the inventorisation of knowledge work roles, benefit assessment, and prioritisation of GenAI efforts with the highest cost-benefit ratio, along with the need to experiment with rapid iteration so as to maximise the benefits of GenAI.

A comprehensive overview of references that deal with GenAI is still to be presented since the currently existing studies are in the early stages and mainly address technical matters. Future research should offer a deeper insight into long-term effects of GenAI on society, the labour force, and economic policies. This research will be crucial for understanding and efficiently managing transformations brought about by this revolutionary technology.

The first part of this paper presents the results of published studies and an overview of the existing GenAI references. The second part of this paper analyses the transformative levels of GenAI impact, which are divided into three segments: macro, meso, and micro levels. The third part of this paper shows research results and discussion, in which relative findings and their implications on the social and economic structures are presented. Finally, the fourth part of this paper summarises key conclusions and offered recommendations for future research and the application of GenAI in practice.

Methodology

The research presented in this paper deals with an extensive analysis of the impact of GenAI on the social and economic systems. The objective is two-fold:

1. To identify the ways in which GenAI is integrated into different sectors:
    - How does GenAI promote innovation?
    - What are the new professional possibilities created by it?
2. To investigate the challenges that come with GenAI:
    - How does it affect the existing practices and paradigms?
    - What are the potential dangers?

With this comprehensive research, we aim to provide a clear picture of the possibilities and challenges that GenAI poses for modern society.

Transformative potential of generative artificial intelligence on macro, meso, and micro levels

GenAI stands for a major milestone in the evolution of technological development, bringing with it extraordinary potentials for a fundamental transformation both of individual lives and of broader social and economic structures. Given the wide range of its impact, from individual to global levels, GenAI opens new perspectives in terms of efficiency, innovativeness, and adaptability. This technology pushes the boundaries of traditional methodologies, thus allowing for revolutionary changes in business modalities, education, and communication.

Wide availability of these tools has aroused growing curiosity of the general public. ChatGPT stands out as one of the most striking examples having become one of the fastest-growing consumer application ever. After it was first launched, in just two months ChatGPT attracted 100 million users (Baum et al., 2023), thus breaking all previous records in terms of the speed of user base growth, compared to any other digital service. Not only does this remarkable success illustrate the growing allure of GenAI technologies, but it also under-

lines the potentials that these innovations have in the transformation of both business and personal aspects of our lives._

To be able to utilise all options offered by Gen AI to a maximum extent, it is imperative to have in-depth understanding of its implications at the various levels of transformation - **macro, meso, and micro levels**. The complexity of this technology requires a comprehensive approach which integrates interdisciplinary knowledge and strategic planning so as to identify and utilise the potentials of GenAI, along with a concurrent consideration of potential risks and ethical challenges. This approach allows for the implementation of strategic paradigms brought by GenAI to various spheres of society and economy.

At the macroeconomic level, GenAI holds the capacity for a thorough transformation of the economic and social infrastructures. This technology has a crucial role in the shaping of global trends and directing economic policies, with a potential to redefine the labour market. The implementation of GenAI may result in substantial change in the structure of labour, whereby certain occupations could discontinue, whereas other, new occupations could emerge. Such dynamics calls for an all-encompassing review of educational programmes and strategies for the training of the labour force, in order to meet the latest demands of the market.

Economic output of GenAI is potentially revolutionary. As Paul Krugman, winner of the Nobel Prize in Economic Sciences notes, productivity plays a key role in the long-term economic growth (Colford, 2016). Predictions indicate that an extensive use of GenAI in the next decade could raise the annual work productivity rate by approximately 1.5% in the United States of America, as well as in other developed countries. It is expected that such a trend will largely affect gross domestic product (GDP) globally, whose projected annual growth is up to 7%. The analyses conducted by various financial institutions, such as Goldman Sachs, suggest that the value of indices, such as S&P 500, could grow by up to 9% thanks to the improvements to productivity and corporate performances (Goldman Sachs, 2023).

Every year we can see dynamic growth of the AI market. The size of the global AI market in 2023 was approximately USD 208 billion, with a prediction that by 2030 it will grow to approximately USD 1.85 trillion. These figures not only illustrate the speed of AI technology development, but also its ever stronger integration into various segments of economy and society (Duarte, 2024).

For the purpose of a continued consideration of the impact of artificial intelligence on the social and cultural spheres, it is necessary to look into its economic implications. According to the assessment made in a 2023 report created by McKinsey Company, AI could contribute to global economy by adding the impressive USD 25.6 trillion. This number is indicative of the potentials AI holds to transform economic structures worldwide (Chui et al., 2023).

GenAI has the potential to significantly prompt the growth of gross domestic product (GDP) through several key mechanisms:

- **Raising labour productivity**: The automation of repetitive and time-consuming tasks allows workers to focus on more complex and creative activities, which directly

contributes to economic growth. Natural language processing and data analysis tools, as well as the automation of business processes, can transform the companies' operating models, reduce operating costs, and increase capacities.

- **Innovation and the development of new industries**: The use of GenAI makes it possible to create new products, services, and industries, which range from healthcare technologies to personalised digital assistants. Both the creation of new markets and the expansion of the existing ones significantly contribute to economic growth.
- **Creating new jobs**: In addition to the optimisation of existing tasks, GenAI creates a need for new types of occupations. The need arises for new jobs which require specialised skills in AI systems management, monitoring, and improvement, including AI ethics experts, AI security engineers as well as data analysts.
- **Supply chain and logistics optimisation:** GenAI can significantly improve supply chain efficiency through analyses and predictions by reducing losses and cutting operating costs, which contributes to overall economic well-being.
- **Improving public services and administration**: AI-assisted data automation and analysis can enhance the efficiency of public services, optimise resource allocation, and cut costs, which has a positive effect on economic growth.
- **Attracting foreign investments**: Efficient and innovative economic systems that use GenAI have become increasingly attracting of foreign investments, which contributes to the growth of foreign direct investments.

At the meso level of transformation, the focus of GenAI is aimed at specific industries and sectors, with an emphasis on how technology can prompt adaptation and innovation within the business environments. Some industries, such as finance, healthcare and the legal industry, are already using the possibilities GenAI offers quite noticeably for the purpose of reshaping their operations. This technology allows companies to automate routine tasks, create new products and services, and also alter their business models and operational strategies. Furthermore, GenAI plays a key role in the transformation of corporate culture and the development of new skills within teams.

AI also offers the possibility to automate a wide range of jobs, which affects the labour force in different ways. According to a 2023 report, approximately two thirds of jobs have the potential to be partially automated through the application of AI technologies. However, this does not necessarily mean the replacement of human labour, but complementing it instead, thus opening new possibilities for those workers who perform more complex tasks (Goldman Sachs, 2023).

GenAI can have a significant impact on the careers of professionals with high education, who work in the economically valuable sectors. Rakesh Kocchar notes that "workers who are more exposed to AI see more help than harm" (Kochhar, 2023). Such changes can rapidly and extensively affect the labour market by driving significant adaptation in the majority of professional fields (Schulz, 2023).

According to an analysis by the McKinsey Global Institute, a major part of work tasks in the U.S. are facing the challenge of automation. Specifically, between 20% and 50% of jobs in the U.S. could be exposed to various levels of automation by 2030 (Kenan Institute of Private Enterprise, 2023). This projection suggests that almost a quarter of all work activities could be transferred to AI-based technologies, which implies substantial changes in the structure of the labour force and required skills. Such a transition calls for extensive adjustments both at an individual level, as well as at the organisational one, which encourages education and the development of new skills for the purpose of efficiently responding to such changed market requirements.

The challenges of work task automation are particularly noticeable among women, who are more significantly affected than men are. According to reports, women face the risk of job automation by 21% more than men. This phenomenon can be explained in part by the fact that men make up the majority in the labour market. However, women mainly have jobs in the sectors which are most susceptible to automation, such as admin support, client services, and the food industry (Kenan Institute of Private Enterprise, 2023).

A 2023 survey conducted among professionals in the U.S. showed that 37% of workers in the advertising and marketing sectors used AI as support in their business activities. By contrast, the healthcare

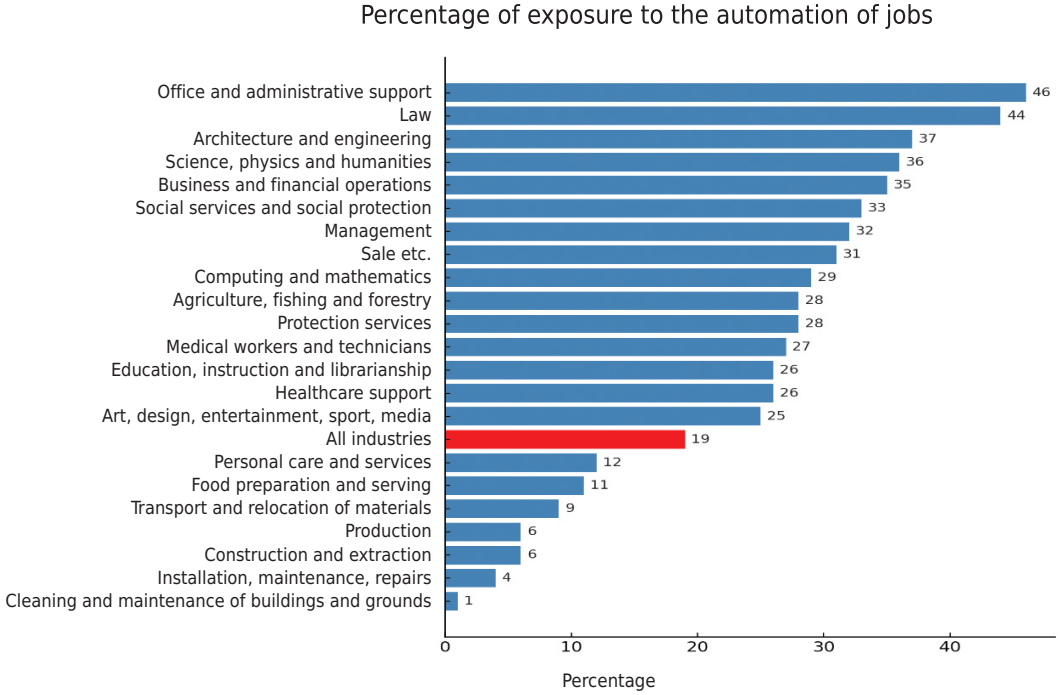


Figure 1. Percentage of exposure to the automation of jobs which could be automated by AI (Goldman Sachs, 2023).

sector showed the lowest degree of the use of AI, with as few as 15% respondents who reported using AI in the work environment. The high acceptance level of AI technologies in marketing and advertising can be explained by the nature of this industry, which strongly integrates creative processes in the media, thus making room for a streamlined use of AI tools for optimisation and innovation purposes (Statista, 2024).

Specialised versions of GenAI, such as ChatGPT Enterprise, provide companies with advanced tools which are adapted to their business needs. This platform provides enhanced security and privacy, faster access to the models such as GPT-4 and expanded functions for the processing of complex tasks. Such platforms allow the automation of chatbots for user support; they facilitate creative processes such as content writing and marketing, and they also encourage integration with other platforms. These innovations are crucial for the improvement of operational performance and company competitiveness (Open AI), through:

- **Improving business processes**: GenAI allows the optimisation of business processes through automation and data analytics, thus increasing efficiency, cutting costs, and shortening task performance time. This ensures a more a prompter response to market requirements and redirecting resources to strategic initiatives.
- **Product and service innovation**: GenAI encourages the development of new products and services as well as the enhancement of existing ones. By analysing large quantities of data, companies can identify new trends and customer needs, which allows the creation of targeted and innovative solutions, this being applicable to anything from personalised financial advice to advanced diagnostic tools in medicine.

At the micro level, GenAI research is directed towards individual users, small teams and specific uses in daily lives and the work environment. This technology allows individuals to enhance their productivity via professional assistants that provide support for various activities, including writing, research, and other creative tasks. Additionally, GenAI offers customised user experience in many applications, be it e-commerce or educational tools, thus significantly improving the quality of customer experience.

It is worth noting that GenAI transforms the access to information and resources, which allows the users to utilise their time and resources more efficiently. For example, platforms such as ChatGPT, can increase the workers' productivity in writing, in terms of generating ideas, text editing, or creating full textual content. This technology also contributes to reducing workplace inequality by allowing the lesser skilled workers to perform tasks equally well as their skilled peers. This in particular refers to improving grammar, spelling, and thought organisation, thus providing the workers who are less skilled at writing with an opportunity to create quality blogs, articles, e-books, and other content that is published online (Vukmirović, 2024).

In the context of scientific research, AI is now assuming a key role in the processes of data analysis and interpretation. Furthermore, a gradual replacement of traditional human labour by innovative technologies calls for the development

of new capacities for work process resilience and adaptability. Even some prominent thinkers, such as Stephen Hawking, have voiced their concerns that AI could bring an existential threat to humankind, because of its potential to take control of many aspects on our daily lives and societal organisation (UNESCO, 2013).

The results of a 2023 study conducted by the authors from the National Bureau of Economic Research (NBER), whose focus was on quantifying the efficiency of resolving business tasks per hour, show that the workers who used GenAI tools had productivity which was by 14% higher than that of their peers who did not use those tools. This improved productivity was extremely perceptible in newly-hired workers as well as in lesser-qualified workers. On the other hand, in highly qualified professionals, the increase in productivity was less noticeable. These results demonstrate the potentials of GenAI tools to raise workplace efficiency levels, especially by encouraging productivity in less experienced team members, whereas in the already established professionals, GenAI is applied in more specific and more complex contexts (Brynjolfsson, Li & Raymond, 2023).

- **The impact on personalised productivity**: GenAI allows the users to customise the tools to their specific needs, which increases efficiency and satisfaction. The users can utilise GenAI to organise activities,

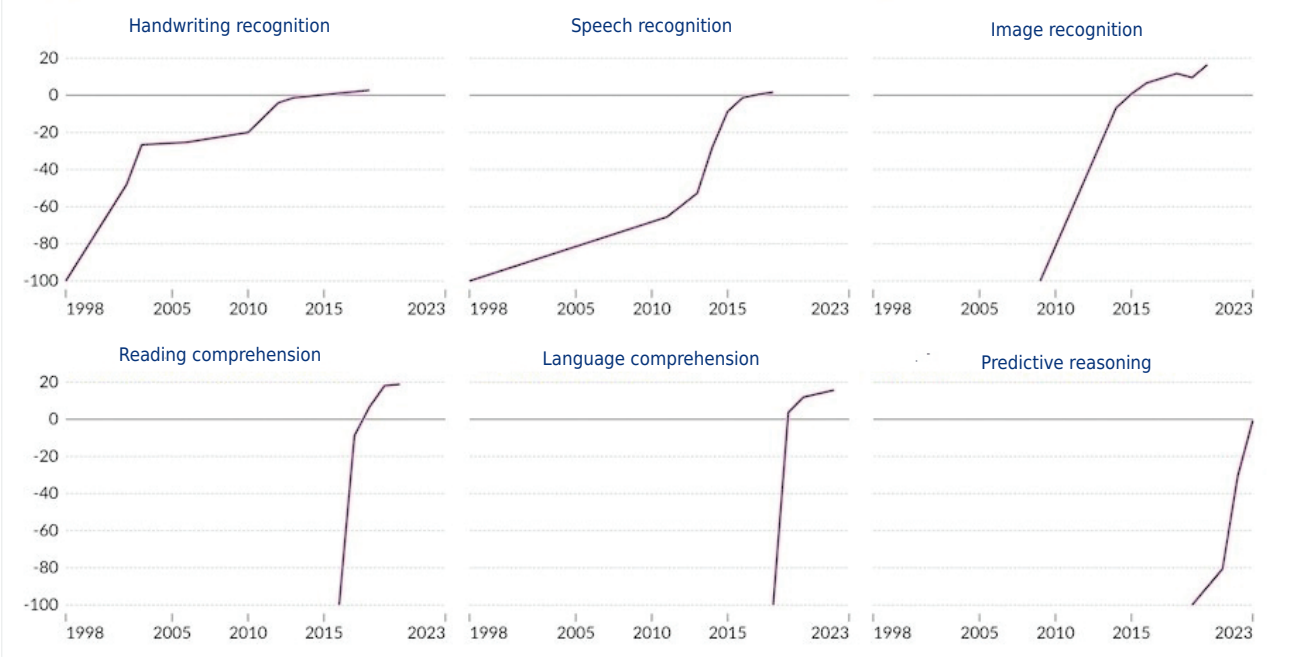


Figure 2. For each of the areas, AI initial performance was set at 100, while human performance was used as the basis, i.e., zero. As AI performance increases and exceeds the basis, it achieves better results than humans (Kiela et al., 2023).

manage finances, and for medical advice. Personalisation enhances productivity and engagement, but it also requires to be legislatively regulated for the purpose of protecting privacy and reducing bias. The micro level of transformation via GenAI represents significant advancement in the manner in which individuals and teams use technology because it shows how technological advancement can be customised to various aspects of human activity.

In 2023 research conducted by Our World in Data (Kiela et al., 2023), AI performances were tested and then compared with humans' performances in the same fields. Task performance was tested, and it ranged from image creation, design, to reading comprehension. The results show that AI surpassed human performance in understanding linguistic and text documents, but also in subject recognition in photographs. Also, AI performance is equal with human performance in terms of speech recognition, handwriting recognition, and predictive judgement.

GenAI: a step ahead

Although they are widely known as "neural networks", AI tools do not work in a manner that is similar to the human brain. AI machines rely mainly on advanced statistical methods for data processing and response generation. Basically, these machines carry out complex statistical analyses by using the algorithms that are based on word frequency and patterns in large data sets. Their ability to grasp abstraction, as well as the ability of deep reasoning, are limited because they do not have internal world models which are a feature of human cognition (Goldman Sachs, 2023).

The learning performed by these machines is based on the analysis of enormous amounts of data, whereby patterns and correlations are identified, which allows the generation of relative responses to specific queries. Nevertheless, unlike humans, the machines cannot understand abstract concepts or complex interactions that are characteristic of human perception and thought. Hence, their intelligence is limited to reflexive and statistical tasks, without being able to truly understand and have awareness.

The "study of artificial intelligence" is planned "to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it" (UNESCO, 2019). This presumption represents the fundamental principle based on which modern AI systems are developed, in reliance of statistical and mathematical models which make attempts at simulating human cognitive processes.

Conclusion

This research has shown that generative artificial intelligence has the capacity for a deep and extensive transformation at macro, meso, and micro levels of social and economic structures. At the macroeconomic level, the implementation of generative artificial intelligence technologies can significantly raise the global labour productivity, encourage innovation, and create new professional opportunities, which ultimately leads to the growth

of gross domestic product. The application of these technologies at the meso level allows the industries such as finance, healthcare and the legislative sector to redefine their business models and operational strategies, thereby additionally increasing their efficiency and competitiveness. At the micro level, the generative artificial intelligence tools offer individuals and small teams the possibility of enhancing personal productivity and improving customer experience by using personalised assistants and customised applications.

Nevertheless, while the potentials of generative artificial intelligence, as well as the technology itself, are no doubt immense, it should be pointed out that this transformation comes with substantial challenges and risks. Ethical issues, including data privacy, algorithm bias, and potential replacement of the human labour force, must be carefully considered and regulated. In order to fully utilise all benefits offered by generative artificial intelligence, it is necessary to develop interdisciplinary approaches and strategies that integrate technological development with ethical and social aspects. Only in this way will it be possible to maximise the positive effects of this revolutionary technology and minimise potential adverse consequences.

References

Alavi, M., & Westerman, G. (2023). *How generative AI will transform knowledge work*. Harvard Business Review. Available at: https://hbr.org/2023/11/how-generative-ai-will-transform-knowledge-work

Baum, A., et al. (2023, September). *Unleashing AI: The AI arms race*. City GPS: Global Perspectives & Solutions. Available at: https://ir.citi.com/gps/JdGH5yFEQ39Tk0vlBoktHGiHWIZj%2B%2FakJbUuxN2mF%2FE2AlVFsFFhvg8Eq8rUb9WUQElCCL-GRYOc0xwXu%2B6CEQQ%3D%3D

Bertolo, R., & Antonelli, A. (2023, November 15). Generative AI in scientific publishing: disruptive or destructive?. *Nature Reviews Urology*, XXI (1–2) (2024). DOI: https://doi.org/10.1038/s41585-023-00836-w

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative AI at work*. National Bureau of Economic Research. Available at: https://www.nber.org/system/files/working_papers/w31161/w31161.pdf

Chui, M., et al. (2023, June 14). *The economic potential of generative AI: The next productivity frontier*. McKinsey & Company. Available at: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#business-value

Colford, C. (2016, November 15). *Productivity for prosperity: In the long run, it is almost everything*. World Bank Blog. Available at: https://blogs.worldbank.org/en/psd/productivity-prosperity-long-run-it-almost-everything

Duarte, F. (2024, January 29). *AI market size statistics (2024)*. Exploding Topics. Available at: https://explodingtopics.com/blog/ai-market-size-stats

Fahle, S., Prinz, C., & Kuhlenkötter, B. (2020). Systematic review on machine learning (ML) methods for manufacturing processes – Identifying artificial intelligence (AI) methods for field application. *Procedia CIRP*, XCIII, 413–418. DOI: https://doi.org/10.1016/j.procir.2020.04.109

Goldman Sachs. (2023, April 05). *Generative AI could raise global GDP by 7%.* Available at: https://www.goldmansachs.com/intelligence/pages/generative-ai-could-raise-global-gdp-by-7-percent.html

Goldman Sachs. (2023, July 10). *Generative AI: Hype, or truly transformative?* Available at: https://www.goldmansachs.com/intelligence/pages/generative-ai-hype-or-truly-transformative.html

Kanbach, D. K., Heiduk, L., Blueher, G., Schreiter, M., & Lahmann, A. (2024). The GenAI is out of the bottle: generative artificial intelligence from a business model innovation perspective. *Review of Managerial Science*, XVIII (4), 1189–1220. DOI: https://doi.org/10.1007/s11846-023-00696-z

Kenan Institute of Private Enterprise. (2023, April 18). *Will generative AI disproportionately affect the jobs of women?* Available at: https://kenaninstitute.unc.edu/kenan-insight/will-generative-ai-disproportionately-affect-the-jobs-of-women/

Kiela et al. (2023, Last updated 2024, April 2). *Test scores of the AI relative to human performance*. Our world in data. Available at: https://ourworldindata.org/artificial-intelligence

Kochhar, R. (2023). *Which U.S. Workers Are More Exposed to AI on Their Jobs?* Pew Research Center. Available at: https://www.pewresearch.org/social-trends/wp-content/uploads/sites/3/2023/07/st_2023.07.26_ai-and-jobs.pdf

McAfee, A., Rock, D., & Brynjolfsson, E. (2023, November 1). *How to capitalize on generative AI*. Harvard Business Review. Available at: https://hbr.org/2023/11/how-to-capitalize-on-generative-ai

MIT Technology Review Insights. (2024, February 29). *Generative AI: Differentiating disruptors from the disrupted*. Telstra. Available at: https://www.telstra.com.sg/en/campaigns/generative-ai-differentiating-disruptors-from-the-disrupted

Open AI. *What is ChatGPT Enterprise?* Available at: https://help.openai.com/en/articles/8265053-what-is-chatgpt-enterprise

Schulz, B. (2023, August 4). *What jobs are most exposed to AI? Pew research reveals tasks more likely to be replaced*. USA Today. Available at: https://www.usatoday.com/story/money/2023/08/04/jobs-most-exposed-ai/70523057007/

Stahl, B. C., et al. (2023). A systematic review of artificial intelligence impact assessments. *Artificial Intelligence Review*, LVI (11), 12799-12831. DOI: 10.1007/s10462-023-10420-8

Statista. (2024, May 10). *Rate of generative AI adoption in the workplace in the United States 2023, by industry*. Available at: https://www.statista.com/statistics/1361251/generative-ai-adoption-rate-at-work-by-industry-us/

Statista. (2024, May 28). *Market size and revenue comparison for artificial intelligence worldwide from 2020 to 2030*. Available at: https://www.statista.com/statistics/941835/artificial-intelligence-market-size-revenue-comparisons/

Talib, M. A., Majzoub, S., Nasir, Q., & Jamal, D. (2020). A systematic literature review on hardware implementation of artificial intelligence algorithms. *The Journal of Supercomputing*, LXXVII, 1897–1938. DOI: https://doi.org/10.1007/s11227-020-03325-8

UNESCO Paris (2019, February). *Preliminary study on the ethics of artificial intelligence*. World Commission on the Ethics of Scientific Knowledge. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000367823

Vukmirović, D. (2024). *The Transformational Potential of Generative Artificial Intelligence in Marketing and Communications*. DOI:10.13140/RG.2.2.31292.65927

Ljubiša M. Bojić[1]
Institute for Artificial Intelligence Research and Development of Serbia
Novi Sad (Serbia)

Milan M. Čabarkapa[2]
University of Kragujevac,
Faculty of Engineering Sciences,
Department of Electrical Engineering and Computing
Kragujevac (Serbia)

UDC 004.8(497.11)
001.895
Original scientific article
Received: 20.06.2024.
Accepted: 09.07.2024.
doi: 10.5937/napredak5-51738

Strategic Investment in the Research and Development of Memristor Technology in the Republic of Serbia

Abstract: The rapid advancement of Artificial Intelligence (AI) has significantly impacted both high technology development and economic and social progress. The Republic of Serbia has been strategically supporting research and development of in the field of AI. Given the dramatic dynamic development of AI, the aim of this paper is to identify and describe memristor technology as currently very relevant and attractive, in order to achieve technological innovation, socio-economic benefits, and potentially global breakthroughs. The paper presents an overview of literature to analyze theoretical concepts, current research outcomes in AI, and possible applications of memristors. The analyses indicate that adoption and development of memristor technology in Serbia can position the country as a leader in AI hardware innovation, attracting international partners and fostering a technologically advanced industrial system. Therefore, this paper suggests that future research should focus on overcoming practical challenges in the production of memristors, developing hybrid architectures, and formulating advanced neuromorphic algorithms.

Keywords: memristors, neuromorphic computing, technological innovation, strategic positioning, global breakthrough

Introduction

AI has rapidly transformed from a narrow field of computing sciences to an important factor which shapes economy and society worldwide. Since we are at the threshold of a new technological revolution, it is necessary to identify a specific field within AI in order to direct resources into investment and development, with the aim of achieving technological progress and socio-economic benefits. Some of the more relevant steps taken by the Republic of Serbia in order to support development in this field are

[1] ljubisa.bojic@ivi.ac.rs ; https://orcid.org/0000-0002-5371-7975
[2] mcabarkapa@kg.ac.rs ; https://orcid.org/0000-0002-2094-9649

the adoption of the Strategy for the Development of Artificial Intelligence for the period 2020-2025, the foundation of the institute dedicated to AI research and development, the adoption of the guidelines for ethical development of AI, and the establishment of a legal framework for autonomous driving (Strategy, 2020; IVI, 2022; Etika, 2023; Auto, 2023). Regarding this field, the Republic of Serbia has the advantage over many countries in the world thanks to its geopolitical position, strong academic institutions and the scientific- technological system.

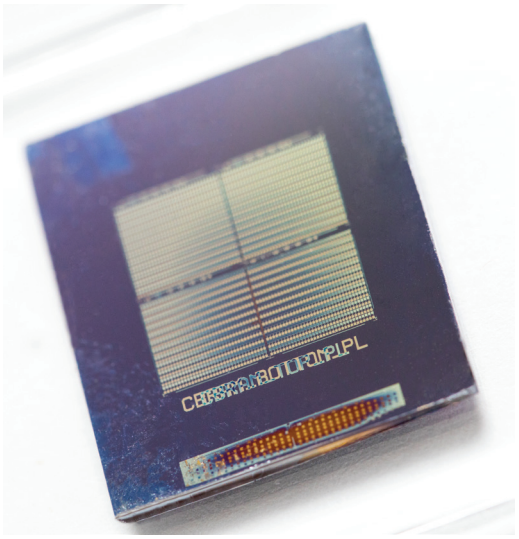
The main question to be considered in this paper is the manner in which the strategic investment of the Republic of Serbia in the field of AI may lead to technological innovation. With the dynamic development of AI, apart from the areas already identified in the valid Strategy for the Development of AI, as well as the favourable geopolitical position for establishing a forum which would work on global solutions for regulation and ethics of AI (Bojic, 2024; Bojic, 2022), memristor technology appears as interesting because of the expectations that it would solve a problem of vital importance, namely the high consumption of electric power of AI systems, such as large language models (ChatGPT) (Bojic et al., 2024, Talanov et al., 2024). This is of particular importance because big technological companies develop general-type artificial intelligence which will be applied through different solutions and services in all aspects of society. Therefore, this paper primarily considers the possibility of the development of memristors, proposing that the Republic of Serbia should begin investing in this type of technology, which is expected to enable the development of a new generation of neural networks and lead to a revolution in AI.

Memristors and neural networks of the third generation

The paper by Leon Chua from 1971 presents the theoretical basis about memristors, where they are described as an element of electric circuits with the commonly known elements – resistors, capacitators and inductors. Chua explains that memristors can store and process information similarly to the activities of brain synapses. This technology connects in a unique manner electric voltage and magnetic flow, thus enabling a non-linear connection between voltage and electricity. This connection enables memristors to keep memory without constant power supply, which is a characteristic that reminds of the functioning of brain synapses.

> Memristors are the third generation of neuron networks which provide greater efficiency and the possibility of energy saving in AI computing processes. Because of the ability of memristors to emulate the functions of brain synapses, memristor technology is suitable for the application in artificial intelligence.

Memristor technology was a mere theoretical construct until as late as 2008, when it was constructed with the aid of titanium- dioxide nanometre coating (Strukov et al., 2008). This invention confirmed



Memristor, developed by the University of Illinois and National Energy Technology Laboratory of the USA, invented by Leon Chua in 1971.
Photo: Wikipedia

Leon Chua's assumptions and opened up the road for further research functionality of memristor technology derives from its characteristic of changing resistance based on the history of voltage. This property is known as hysteresis and it emulates the synaptic plasticity of biological neuron networks in which synaptic power varies depending on neuron activity.

Since nanometre coatings are used in the construction of memristor technology, it is possible to substantially reduce energy consumption as compared to traditional silicon resistors. Since they enable fast and efficient information processing, the devices in which memristors are used can create more capable AI systems, and also be applied in different fields, from robotics to data analysis in real time. Recent research points to the application of neuron networks based on memristors in solving tasks of pattern recognition (Prezioso et al., 2015), which corroborates the fact that memristors technology systems can autonomously learn and solve problems in a similar way to biological systems.

Technological achievements and challenges in the construction of memristor technology

The most relevant field of research related to memristor technology refers to the materials used for the construction of memristors. As it has already been mentioned, titanium-dioxide was first used in its production. Later on, other materials were used,

such as hafnium-oxide, graphene-oxide and organic compounds. Each of these materials has unique properties which affect the performances and production costs of memristor technology (Adhikari et al., 2012). Graphene-oxide is a suitable material because of its electric conductivity, mechanic stability and flexibility. Research shows that memristor technology based on graphene-oxide is characterized by durability and a great switching speed function, which is quite important in the application in AI. However, there are still problems in relation to the achievement of uniformity in production. Variability of the properties of materials leads to uneven performances of memristor technology. That is why new techniques of its production, such as atomic deposition and chemical deposition of coating layers from the steam phase, the subject of research, with the aim of improving reproductivity of devices which use memristor technologies (Chang, Jo, Lu, 2011).

The application of memristors in traditional computer systems requires innovative approaches in the design of electric circuits. Unlike conventional transistors, non-linear functioning of memristor technology requires new architectures of electric circuits. Hybrid circuits in which memristors supplement traditional transistors can contribute to the efficiency of computing processes (Wang et al., 2017). An important step forward in the development are crossbar arrays in which memristors are distributed on the network structure. Crossbar arrays ensure great density of memory storage and parallel computing procedures. Research shows that it is possible to apply crossbar arrays in neuromorphic systems with the ability of pattern recognition and autonomous learning (Prezioso et al., 2015). It should be noted that designing efficient mechanisms for reading and writing circuits which are based on memristors is necessary for the application of these circuits. That is why reading and writing algorithms are explored, as well as error correction techniques in order to improve the reliability of the memory systems which use memristor technologies.

By emulating the plasticity of the synapses, memristors can facilitate synaptic weighting and adaptation (Yang et al., 2013). Memristor-based accelerators go beyond traditional systems based on graphics processing units (GPUs). With the expansion of the *Internet of Things*, there is an increasing need for efficient edge computing, in which data are processed locally instead of relying on remote servers. Memristor-based systems are suitable for application in edge computing, as well as in edge AI (Ambrogio et al., 2018). These systems can be applied in smart networks and autonomous vehicles, for example, which are only some of the numerous possibilities of the memristor technology application.

Applications of memristor technology in different industries

The applications of memristor technology are diverse and promise revolutionary changes in different industries, from healthcare and agriculture to energy and automobile industry. Thanks to its ability to increase energy efficiency and enable advanced computing processes, memristor technology is considered the key bearer of innovation

and technological progress. In this chapter we will analyze the most relevant applications of memristors in different fields, as well as their advantages and challenges.

Memristor technology can substantially improve medical diagnostics and therapy. Neuromorphic systems based on memristors ensure advanced methos of medical image analysis and pattern recognition in complex biomedical data. These systems can help doctors in making more precise diagnoses in a shorter period of time.

One of the innovative applications of memristor technology is in the development of implantable medical devices, such as neurostimulators and pacemakers. Memristors can share energy requirements of these devices, thus reducing the need for frequent replacement of batteries or for charging these devices. Moreover, these devices might be able to adapt to individual patients' needs thanks to their ability of self-learning.

Agriculture is facing huge challenges in terms of sustainability and efficiency. Memristor technology can contribute to the resolution of these problems through the development of smart mechanisms for irrigation and resource management. Memristor-based systems may analyze a huge amount of data in real time, including soil moisture, weather conditions and the condition of crops, in order to optimize the use of water and other resources.

Moreover, memristor technology enables the development of autonomous robots which can perform agricultural activities such as sowing, harvesting and pest control. These robots might use advanced algorithms for autonomous learning and adaptation to different conditions, thus substantially increasing the efficiency of agricultural processes.

Energy industry can also benefit largely from memristor technology. One of the most promising applications is in the development of smart grids. These grids can use memristors for storing energy and managing its flow in an efficient manner. Memristors can ensure fast data processing and decision-making in real time, which is crucial for the optimization of electricity consumption and integration of renewable energy sources, e.g., solar and wind-electric systems.

In addition, memristor technology can play an important role in the development of energy-efficient devices and systems for energy storage. Great density of memory storage and small energy consumption make memristors ideal for the application in new-generation batteries and other systems for storing energy.

Automobile industry is preparing for a revolution with the development of autonomous vehicles. Memristor technology can significantly improve the possibilities of these vehicles through advanced systems of data processing and autonomous learning. Autonomous vehicles require fast and accurate processing of large amounts of data coming from sensors and cameras, and memristors could ensure substantially greater efficiency in processing these data as compared to traditional computer systems

Memristors can also contribute to the development of energy-efficient systems for storing energy in electric cars. Reduced energy consumption and increased density of memory storage may improve battery performances and extend the time of vehicle functioning between the charging periods.

The Internet of Things (IoT) is increasingly becoming part of everyday life, while memristor technology can ensure substantial improvements in this field. Memristor-based devices can process data directly at the source (edge computing), which reduces dependence on central servers and clouds and ensures faster and more efficient decision-making. The application of memristors in IoT devices can include smart sensors for monitoring the environment, energy management in smart homes and the development of autonomous robots and drones. Memristor technology is ideal for this application thanks to its small energy consumption and the abilities of autonomous learning and adaptation.

Although memristor technology offers a large number of advantages, there are also challenges in relation to ethical and security aspects of its application. The development of the systems which may autonomously learn and make decisions raises questions of responsibility and transparency. There is a need for the development of ethical standards and regulatory frameworks which will ensure that the use of memristor technology is in line with the principle of social responsibility and privacy protection.

Security challenges also include protection of memristor-based systems from malicious attacks and manipulations. Securing data and integrity of the systems must be a priority in the development and application of memristor technology.

The applications of memristor technology in various industries have the potential of bringing important innovation and improving the efficiency and sustainability of different systems. No matter whether it is healthcare, agriculture, energy, automobile industry or the Internet of Things, memristors provide the base for the development of advanced and efficient technological solutions.

However, successful application of memristors also depends resolving challenges in relation to their reliability, production costs and ethical aspects. Cooperation of the academic community, economy and the state can encourage accelerated development and broad acceptance of memristor technology, which would position Serbia as the leader in the field of technological innovation.

Conclusion

Investing in AI research and development can position Serbia at the very top of technological innovation and economic growth. By directing research towards memristor technology, as the third generation of neural networks, the Republic of Serbia may use its unique geopolitical and economic advantages to become relevant at the global level of AI research.

Investing in AI and memristor technologies should contribute to economic growth, as well as to the welfare of society on the whole. Memristor technology should lead to a revolution in AI technology and make computing technology more efficient, more energy-efficient and usable. By emulating synaptic functions of biological neural networks, memristors ensure neuromorphic computing and advanced application of AI. The applications of memristor technology in neuromorphic systems and edge computing enable efficient resolution of social problems and improvement of the quality of life. According to the conclusions listed

in this paper, future research should be directed towards advanced materials, hybrid architectures, neuromorphic algorithms and resolution of problems of changeability and reliability of memristor technology in order to realize all the potentials of memristor technology.

Cooperation between the academic community, economy and the state can encourage the development and application of memristor technology. Its development and application in AI systems in healthcare, agriculture and edge computing, for example, apart from contributing to the welfare of society, should further direct research and development. By strategically investing in the research and development of memristor technology, Serbia can use the transformative potential of this technology.

References

Adhikari, S. P. et al. (2012). Memristor bridge synapse-based neural network and its learning. *IEEE Transactions on Neural Networks and Learning Systems*, XXIII (9), 1426–1435. https://doi.org/10.1109/TNNLS.2012.2204770

Ambrogio, S. et al. (2018). Equivalent-accuracy accelerated neural-network training using analogue memory. *Nature*, DLVIII (7708), 60–67. https://doi.org/10.1038/s41586-018-0180-5

Auto (2023). *New steps in the development of the legal framework for autonomous driving*. Available at: http://www.ai.gov.rs/vest/sr/771/novi-koraci-u-razvoju-pravnog-okvira-za-autonomnu-voznju.php

Bojic, L. (2022). Metaverse through the prism of power and addiction: What will happen when the virtual world becomes more attractive than reality? *European Journal of Futures Research*, 10 (1), 22. https://doi.org/10.1186/s40309-022-00208-4

Bojic, L. (2024). AI alignment: Assessing the global impact of recommender systems. *Futures*, CLX, 103383. https://doi.org/10.1016/j.futures.2024.103383

Bojic, L. et al. (2024a). AI and Energy Consumption: Social Aspects, 1-4. DOI: 10.23919/SpliTech61897.2024.10612493.

Chang, T., Jo, S.-H. & Lu, W. (2011). Short-term memory to long-term memory transition in a nanoscale memristor. *ACS Nano*, V (9), 7669–7676. https://doi.org/10.1021/nn202983n

Chua, L. (1971). Memristor-The missing circuit element. *IEEE Transactions on Circuit Theory*, XVIII (5), 507–519. https://doi.org/10.1109/TCT.1971.1083337

Etika (2023). *Adopted ethical guidelines for the development and use of artificial intelligence*. Available at: https://www.srbija.gov.rs/vest/692988/usvojene-eticke-smernice-za-razvoj-i-upotrebu-vestacke-inteligencije.php

IVI (2022). *Artificial Intelligence Research and Development Institutes*. Available at: https://ivi.ac.rs/

Prezioso, M. et al. (2015). Training and operation of an integrated neuromorphic network based on metal-oxide memristors. *Nature*, DXXI (7550), 61–64. https://doi.org/10.1038/nature14441

Strategy (2020). *Strategy for the Development of Artificial Intelligence*. Available at: https://www.srbija.gov.rs/tekst/437277/strategija-razvoja-vestacke-inteligencije.php

Strukov, D. B. et al. (2008). The missing memristor found. *Nature*, CDLIII (7191), 80–83. https://doi.org/10.1038/nature06932

Talanov, M., Vallverdu, J., Bojic, L. (2024b). Neuropunk revolution: memristive spinal CPG learning approach. *9th International Conference on Smart and Sustainable Technologies (SPLITECH)*, June 25–28, 2024, Split, Croatia.

Wang, Z. et al. (2017). Memristors with diffusive dynamics as synaptic emulators for neuromorphic computing. *Nature Materials*, XVI (1), 101–108. https://doi.org/10.1038/nmat4756

Yang, J. J., Strukov, D. B., Stewart, D. R. (2013). Memristive devices for computing. *Nature Nanotechnology*, VIII (1), 13–24, https://doi.org/10.1038/nnano.2012.240

Nenad D. Filipović[1]
University of Kragujevac,
Faculty of Engineering
Kragujevac (Serbia)
BioIRC Bioengineering Research
and Development Centre
Kragujevac (Serbia)

UDC 61:004.8
Original scientific article
Received: 08.08.2024.
Accepted: 23.08.2024.
doi: 10.5937/napredak5-52622

Computer modelling and artificial intelligence with big data for better diagnostics and therapy of cardiovascular disease

Abstract: In silico clinical trials are the future of medicine and virtual testing and simulation are the future of medical engineering. The use of a computational platform can reduce costs and time required for developing new models of medical devices and drugs. The computational platform in different projects, such as SILICOFCM, was developed using state-of-the-art finite element modelling for macro simulation of fluid-structure interaction with micro modelling at the molecular level for drug interaction with the cardiac cells. SILICOFCM platform is used for risk prediction and optimal drug therapy of familial cardiomyopathy in a specific patient.
STRATIFYHF project is to develop and clinically validate a truly innovative AI-based Decision Support System for predicting the risk of heart failure, facilitating its early diagnosis and progression prediction that will radically change how heart failure is managed in both primary and secondary care. This rapid expansion in computer modelling, image modalities and data collection, leads to a generation of so-called “Big Data” which are time-consuming to be analyzed by medical experts.
In order to obtain 3D image reconstruction, the U-net architecture was used to determine geometric parameters for the left ventricle which were extracted from the echocardiographic apical and M-mode views. A micro-mechanics cellular model which includes three kinetic processes of sarcomeric proteins interactions was developed. It allows simulation of the drugs which are divided into three major groups defined by the principal action of each drug.
The presented results were obtained with the parametric model of the left ventricle, where pressure-volume (PV) diagrams depend on the change of Ca2+. It directly affects the ejection fraction. The presented approach with the variation of the left ventricle (LV) geometry and simulations which include the influence of different parameters on the PV diagrams are directly interlinked with drug effects on the heart function. It includes different drugs such as Entresto and Digoxin that directly affect the cardiac PV diagrams and ejection fraction.
Computational platforms such as the SILICOFCM and STRATIFYHF platforms are novel tools for risk prediction of cardiac disease in a specific patient that will certainly open a new avenue for in silico clinical trials in the future.

Keywords: heart modelling, fluid-structure interaction, machine learning, big data, drug modelling

[1] fica@kg.ac.rs ; https://orcid.org/0000-0001-9964-5615

1. Introduction

Cellular and molecular biology have a very strong influence on our understanding of the structure and function of the heart at the microscopic level. At the macroscopic level, the heart functions as a pump that continuously pumps blood throughout the human body. It is necessary to apply an interdisciplinary approach in order to understand the integrated function of the heart, which includes electricity, physical chemistry, solid mechanics, and fluid dynamics (multiphysics simulation). To better understand different events that occur during a cardiac cycle, both microscopic and macroscopic mechanisms should be taken into account in the development of an integrated model.

Familial cardiomyopathies (FCM) are most commonly diagnosed, or progress of the disease is monitored through in vivo imaging, with either echocardiography or, increasingly, cardiac magnetic resonance imaging (MRI). The treatment of symptoms of FCM by established therapies could only in part improve the outcome, but novel therapies need to be developed to more fundamentally affect the disease process and time course.

It is very important to use a detailed, complex, and anatomically accurate model of the whole heart electrical activity which requires extensive computation times, dedicated software, and even the use of supercomputers (Gibbons et al., 2006, Pullan et al., 2005). We have recently developed a methodology for a real 3D heart model by using the linear elastic and orthotropic material model based on Holzapfel experiments. Using this methodology, we can accurately predict the transport of electrical signals and displacement field within heart tissue (Kojic et al., 2019). Muscles in the body except the heart muscle are activated by electrical signals, transmitted from the nervous system to muscle cells, affecting the change of the cell membranes potentials. Additionally, calcium current and concentration inside muscle cells are the main cause of generating active stress within muscle fibers. Clinical validation in humans is very limited since simultaneous whole heart electrical distribution recordings are inaccessible for both practical and ethical reasons (Trudel et al., 2004). The rapid development of information technologies, simulation software packages and medical devices in recent years provides the opportunity for collecting a large amount of clinical information. Creating comprehensive and detailed computational tools has become essential to process specific information from the abundance of available data. From the point of view of physicians, it becomes of paramount importance to distinct "normal" phenotypes from the appearance of the phenotype in a specific patient in order to estimate disease progression, therapeutic responses and future risks. Recently developed computational models have significantly improved integrative understanding of the heart muscle behaviour in HCM and DCM cardiomyopathies. The development of novel integrative modelling approaches could be an effective tool in distinguishing the type and severity of symptoms in, for example, multi-genic disorder patients, and assess the degree to which normal physical activity is impaired.

Some of the main problems in developing fast and accurate algorithms for automatic LV segmentation in apical images are the presence of speckle, low signal to noise ratio, weak echoes etc., which commonly occur in ultrasound images. Additional-

ly, there is no simple connection between the pixel intensity values in images and physical characteristics of the tissue of interest, which makes thresholding algorithms impossible to use in segmentation in ultrasound images (Moradi et al., 2019). As a result, many authors tried to address the problem of segmentation using different approaches, including active shape, active contours, appearance methods, as well as machine learning-based methods (Noble & Boukerroui, 2006). Their main focus is the endocardial border detection on one echocardiography image frame. The literature shows that the level-set approaches are not that sensitive to initial conditions, but instead their main limitations are the imaging conditions. In contrast, deformable templates are robust to imaging conditions, however they are very sensitive to the initialization conditions (Bosch et al., 2002).

As a result, Big Data technologies contain new frameworks for processing medical data playing an important role in data management, organizing, and analysis through the use of machine learning and deep learning approaches (Kouanou et al., 2018). It also enables fast data access via the NoSQL database (Kouanou et al., 2018). In the area of medical image analysis, due to significant improvement in image collecting equipment, the data is relatively huge (going to Big Data), which makes image analysis challenging (Razzak, Naz & Zaib, 2018). It is said that due to digitalization of medical repositories in hospitals, as well as the use of medical images, digital medical archives size is growing at exponential rate (Ashraf et al., 2020). According to the McKinsey Global Institute, if US healthcare uses Big Data creatively and efficiently, the sector could generate more than $300 billion in value per year. Two-thirds of the value would be realized through lowering US healthcare spending (Belle et al., 2015). This fast expansion in medical imagery and modalities necessitates considerable and time-consuming efforts by medical experts, who are subjective, prone to human error, and there are also interpersonal differences. Using machine learning techniques to automate the diagnosis process is an alternative response to aforementioned challenges; however, typical machine learning methods are unable to cope with complex problems (Razzak, Naz & Zaib, 2018). The successful combination of high-speed computers with machine learning promises the ability to cope with large amounts of medical image data for accurate and fast diagnosis (Razzak, Naz & Zaib, 2018). In recent years, machine learning (ML) and artificial intelligence (AI) have advanced quickly, finding their role in medical image processing, computer-aided diagnosis, image fusion, registration, image segmentation, as well as image-guided treatment. ML techniques extract information (called features) from images and effectively perform decision making (Razzak, Naz & Zaib, 2018).

The main focus of the SILICOFCM project (www.silicofcm.eu) has been on multiscale modelling of familial cardiomyopathy, taking into consideration a comprehensive list of patient-specific features such as genetic, biological, pharmacologic, clinical, imaging and cellular aspects. The main result of the project is the in silico clinical platform with biomechanics of the heart as its main part. The platform is developed using state-of-the-art finite element modelling for macro simulation of fluid-structure interaction with micro modelling at the molecular level for drug interaction with

the cardiac cells. The platform can be used for risk prediction and optimal drug therapy of familial cardiomyopathy in a specific patient. The overarching aim of STRATIFYHF project is to develop and clinically validate a truly innovative AI-based Decision Support System (DSS) for predicting risk of heart failure (HF), facilitating its early diagnosis and progression prediction that will radically change how HF is managed in both primary and secondary care. The DSS integrates patient-centred data obtained using existing and novel technologies, a digital patient library and AI-based algorithms and computational modelling.

2. Method

2.1 Image reconstruction from echocardiography

The proposed methodology for echocardiography image reconstruction is divided into two sections: the first section includes the methods used to analyze apical view, while the second one includes the methods used to analyze M-mode view. A detailed description is provided in Fig. 1. DICOM image format is used as the input to the system. The end user (expert) selects which view is best represented by the image and feeds it to the algorithm.

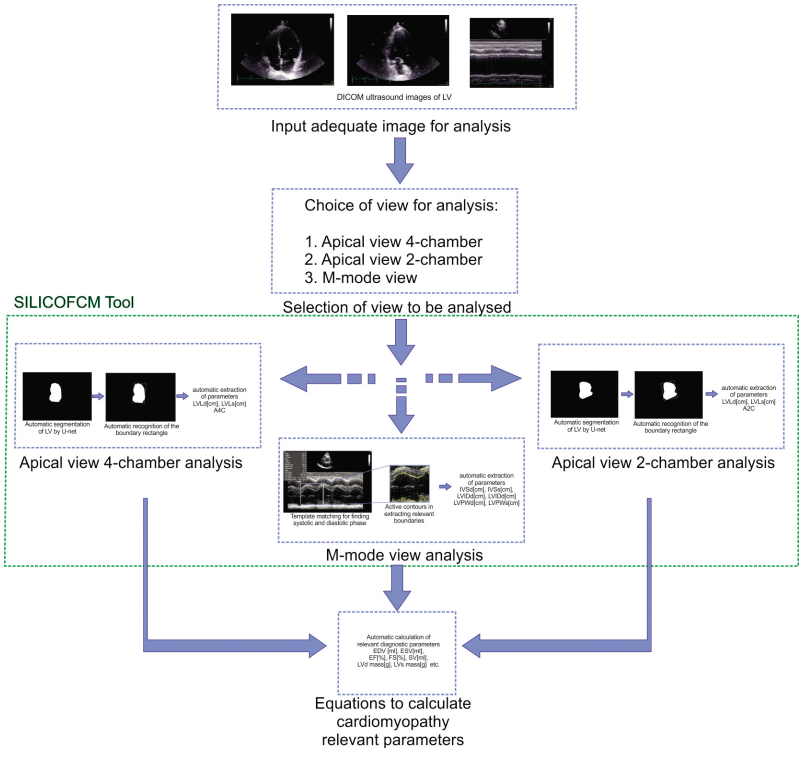


Fig. 1 Description of the proposed methodology for the automatic heart ultrasound segmentation and geometric parameter extraction.

Three alternatives are provided by the SILICOFCM tool: 4-chamber, 2-chamber, or M-mode. This tool will further analyze the images depending on the view mode:

1. **Apical 4-chamber view analysis** includes segmentation of the LV using the U-net previously trained and calculating the bordering rectangle as shown in Fig. 1 (left side), based on which parameters LVLd [cm] and LVLs [cm] A4C will be calculated. The user should define if the view represents the systolic or diastolic phase.
2. **Apical 2-chamber view analysis** includes segmentation of the LV using the U-net previously trained and calculating the bordering rectangle as shown in Fig. 1 (right side), based on which parameters LVLd [cm] and LVLs [cm] A2C will be calculated. The user should define if the view represents the systolic or diastolic phase.
3. **M-mode view analysis** includes bordering of the characteristic areas of the LV – septum in diastole, diameter in diastole, LV wall in diastole, septum in systole, diameter in systole and LV wall in systole (Fig. 1 - middle). Based on these areas, parameters IVSd [cm], IVSs [cm], LVIDd [cm], LVIDs [cm], LVPWd [cm], LVPWs [cm] will be calculated. The user should define that the view is M-mode.

If the user has all three views in systolic and diastolic phase available (which should be the case when imaging the patient), then all relevant parameters are calculated from these three views and automatic calculation of relevant cardiomyopathy diagnostic parameters can be further performed (i.e. – EF [%], ES [%], SV [ml], LVd mass [g], LVs mass [g], etc.).

2.2 Big Data Technologies for Medical Image Processing

Parallel computing is detected as critical infrastructure for managing Big Data. It can perform analysis on a cluster of devices or supercomputers at the same time. Big Data technology with Artificial Intelligence (AI) and massively parallel computing can be used for a revolutionary way of prediction and personalized medicine (Dilsizian & Siegel, 2014). Novel parallel computing models, such as Google's MapReduce (Dean & Ghemawat, MapReduce: simplified data processing on large clusters, 2008), have been proposed in recent years for a new large data infrastructure. Apache has launched Hadoop (White, 2015), an open-source MapReduce software for distributed data management. Concurrent data access to clustered servers is supported via the Hadoop Distributed File System (HDFS). Hadoop-based services may also be thought of as cloud computing platforms, allowing for centralized data storage as well as remote access through the Internet. As such, cloud computing is a revolutionary concept for distributing customizable computational resources across a network (Armbrust, Fox & Griffith, 2010), and it may function as an infrastructure, platform, and/or software to provide an integrated solution. Furthermore, cloud computing may increase system speed, agility, and flexibility by eliminating the need to maintain hardware or software capacity and necessitating less resources

for system maintenance, such as installation, setup, and testing. Cloud technologies are at the heart of many emerging Big Data applications (Luo, Wu, Gopukumar, & Zhao, 2016). Additionally, Hadoop and Spark frameworks have been identified as optimal and efficient architecture for biomedical image analysis (Kouanou et al., 2018).

In addition, High Performance Computing (HPC) uses parallel processing and advanced programs, or software packages speed up massive calculations. In that sense, Finite Element Method (FEM), which represents a continuum method for very powerful scientific computation analysis, strongly relies on advanced computer technology and HPC. Traditional database and software techniques cannot be used for these large-scale computations (Demchenko, Grosso, De Laat, & Membrey, 2013). High Performance Computing (HPC) can be used in medicine contained in Big Data (Lavignon et al., 2013). Massive multiscale computation with multiscale material models, or finite element computation with adaptive mesh refinement can be run only on supercomputers with Big Data on parallel disk systems (Parashar, 2014). A detailed, complex, and anatomically accurate model of the whole heart electrical activity which requires extensive computation times, and the use of supercomputers are already established in the literature (Gibbons Kroeker, Adeeb, Tyberg, & Shrive, 2006; Kojic et al., 2019). The authors of

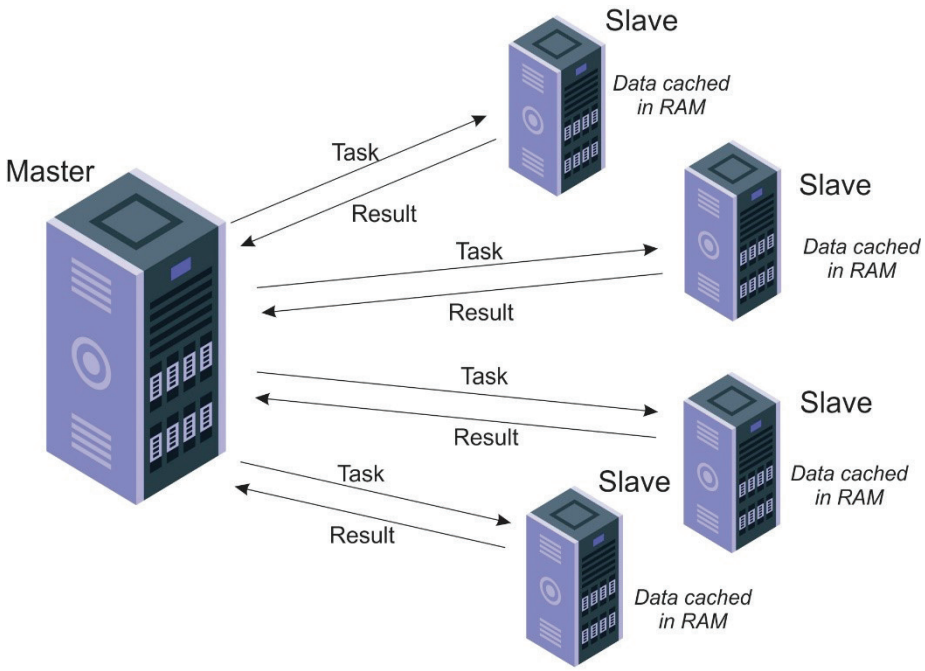


Fig. 2. Job execution using Spark technologies – one master cluster and four slaves

this paper have recently developed a methodology for a real 3D heart model, by using the linear elastic and orthotropic material model based on Holzapfel experiments. Using this methodology, the transport of electrical signals and displacement field within heart tissue can be accurately predicted (Filipovic et al., 2022). Clinical validation in humans is very limited since simultaneous whole heart electrical distribution recordings are inaccessible for both practical and ethical reasons (Filipovic et al., 2022).

On the other hand, Apache Spark is a distributed computing platform that has become one of the most powerful frameworks in the Big Data situation. Spark provides a consistent and comprehensive framework for managing the needs for Big Data processing using a range of datasets (graph data, image/video data, text data, and so on) from various sources (batch, real-time streaming) (Tchito Tchapga et al., 2021). According to its designers, the Spark framework was intended to address the shortcomings of the Hadoop framework. In some cases, the Spark framework has shown to be quicker than Hadoop (more than 100 times in memory). Performance can be quicker than other Big Data technologies with advantages such as in-memory data storage and near real-time processing (Tchito Tchapga et al., 2021). The Spark framework can prepare data for iteration, query it frequently, and load it into memory. The main program (driver) in the Spark framework supervises many slaves (workers) and collects their results, whilst slaves' nodes read data partitions (blocks) from a distributed file system, run various computations, and write the results to disk (Fig. 2). This means that the master controls and assigns jobs to slaves.

Spark, like Hadoop, is built on parallel processing MapReduce, which seeks to process data in a simple and transparent manner across a cluster of computers. Spark enables SQL queries, streaming data, machine learning, and graph processing data in addition to Map and Reduce operations (Kouanou, et al., 2018). In Spark, program can occasionally run the algorithm on several clusters at the same time. Although the number of slaves can be increased due to dataset size, the increase in the number of slaves results in an increase in processing time.

2.3 Cellular model (Mijailovich-Prodanovic MP surrogate and drug model)

The finite element (FE) solvers require calculation of active tension and variable muscle stiffness in each element integration point over all finite elements. Moreover, a relatively fine FE mesh and, therefore, a large number of finite elements are required to precisely calculate the change of heart geometry during a heartbeat. On the other hand, calculation of instantaneous active tension and muscle stiffness by, for example, sliding filament cross-bridge models, requires a solution of partial differential equations (PDE) or Monte Carlo approaches (Mijailovich et al., 2019). Furthermore, coupling of FE solvers to simulate muscle function at the organ level (Mijailovich et al., 2021) with even simpler models, involving the solution of PDEs by the method of characteristics, requires extremely large computational memory and a prolonged time for the execution of simulations even if the simulations are limited to coarse FE meshes.

Parameters for the MP surrogate model were obtained through an automated process of parameter fitting based on a genetic algorithm. The goal was to minimize the root mean square error (RMSE) for obtaining the muscle prediction that would be the closest fit to the one provided by MUSICO Fiber (Mijailovich et al., 2021).

Since the relaxation period is generally harder to fit, RMSE was weighted, giving the fitness of the second part of force development a greater impact on the resulting error.

2.4 Drug testing workflow

The drug actions are different for treating a variety of symptoms associated with cardiomyopathies. In particular, drugs simulated using MUSICO (Mijailovich et al., 2021) are divided into three major groups or pathways defined by the principal action of each drug, such as the effect on modulating calcium transients or changing kinetics of contractile proteins. Each group consists of two subgroups based on a type of cardiomyopathy:

I Modulation of [Ca2+] transients

HCM – Disopyramide, which lowers peak and baseline levels of [Ca^{2+}] transient during twitch contractions,

DCM – Digoxin, which increases peak of [Ca^{2+}] transient during twitch contractions, but does not change time to peak and relaxation time,

The workflow for testing these types of drugs is shown in Fig. 3. The experimental observations in action potentials and changes in ionic currents are simulated using O'Hara-Rudy electro-physiological model that produces intracellular calcium transients as an input for MUSICO and MP surrogate models.

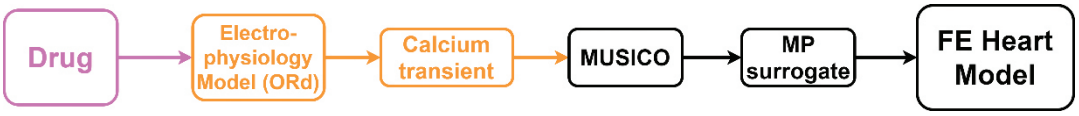


Fig. 3 Pathway 1: Drug action via modulation of calcium transients through changes in ionic currents or membrane properties.

II Changes in kinetic parameters

HCM – Mavacamten, which is associated with the regulation of kinetics rates of transition between disordered myosin detached states and ordered SRX state,

DCM – dATP, which modulates cross-bridge cycle rates,

The workflow for testing these types of drugs is shown in Fig. 4. The experimental observations in the experiments in vitro that quantify the effect of specific drug (dose) are used for the estimation of parameters for MUSICO and MP surrogate models.

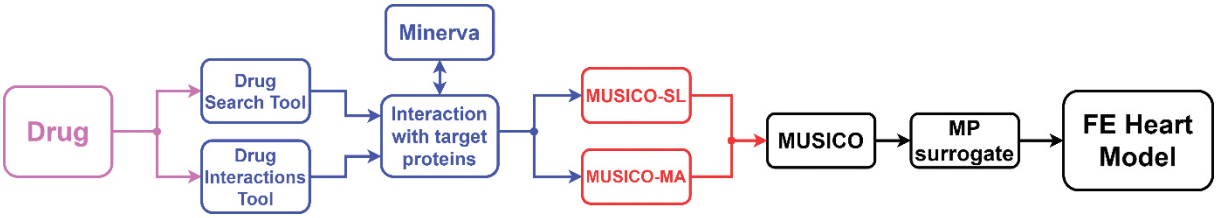


Fig. 4 Pathway 2: Drug action through changes in kinetics of contractile proteins.

Since drugs in groups I and II directly affect MUSICO and MP surrogate parameters, we were able to predict with our tools the outcome on force generation in sarcomeres during twitch contractions.

III Changes in macroscopic parameters

HCM – Entresto®, which acts on remodelling of heart ventricle walls and modulates the elasticity of blood vessels, typically reducing resistance to blood flow and improving cardiac output in HCM. The workflow for testing these types of drugs is shown in Fig. 5. The experimental observations in many clinical trials are used as an input for FE models yielding the precise model of Entresto® action.

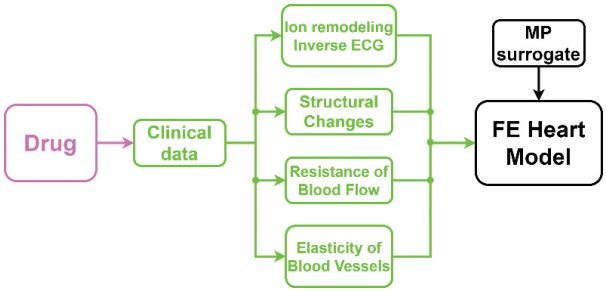


Fig. 5 Pathway 3: Drug action through macroscopic structural and boundary condition changes.

3. Results

The simulations using virtual loading predict left ventricular pressure and volume changes between healthy and HCM and DCM hearts. The predicted traces of the pressures and volumes during heartbeats can be plotted as left ventricular Pressure-Volume loops (Fig. 6). These simulations were obtained using the MP model parameters and experimental calcium transients, with modified lower basal calcium levels and by changing a "force-scaling" mechanical

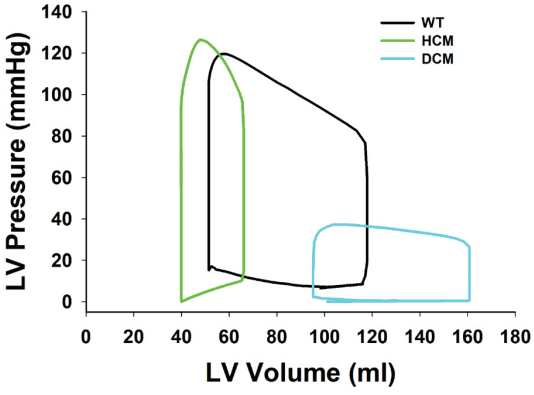


Fig. 6 Left ventricular (LV) Pressure-Volume (P-V) loops during two consecutive heart beats for normal heart (WT, black line), hypertrophic (HCM, green line) and dilated cardiomyopathy (DCM, blue line) obtained with FE coupled with MP surrogate micro model.

parameter, *eta*, in MP model in order to increase twitch peak tensions to observed level.

3.1 Entresto drug influence

ENTRESTO® (Sacubitril/valsartan) has been shown to be superior to enalapril in reducing the risks of death and hospitalization for heart failure (HF). There are also publications which evaluate the effects of sacubitril/valsartan on clinical, biochemical, and echocardiographic parameters in patients with heart failure and reduced ejection fraction (HFrEF).

3.2 Numerical results from the SILICOFCM platform for patients before and after Entresto treatment

Here, we tried to mimic patient cases before and after Entresto drug treatment. Before Entresto treatment PV diagram, pressure diastolic distribution and pressure systolic distribution are presented in Fig. 7. It represents a typical hypertrophic cardiomyopathy patient with decreased ejection fraction and higher systolic pressure (Fig. 7 left, pressure volume diagram).

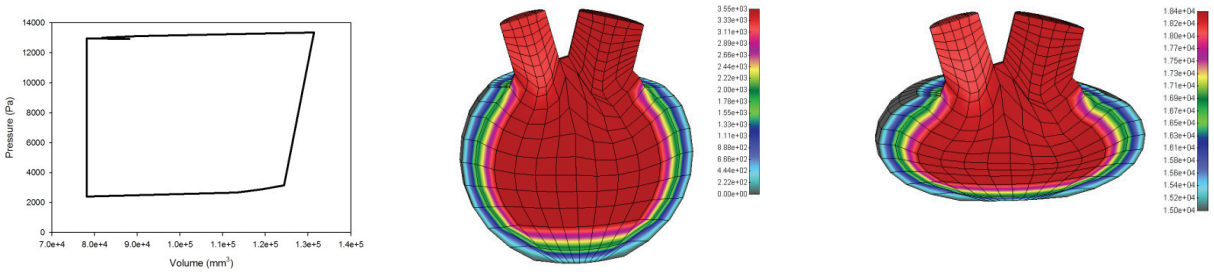


Fig. 7 PV diagram, pressure diastolic distribution, pressure systolic distribution for the case before Entresto treatment.

After Entresto treatment (Fig. 8), we can observe a lower systolic pressure as well as an increasing difference between the end of diastolic and the end of systolic volume. It directly leads to the increase in the ejection fraction. The pressure-volume diagram, velocity distribution in the diastolic and systolic phase for the case after Entresto treatment is presented in Fig. 8.

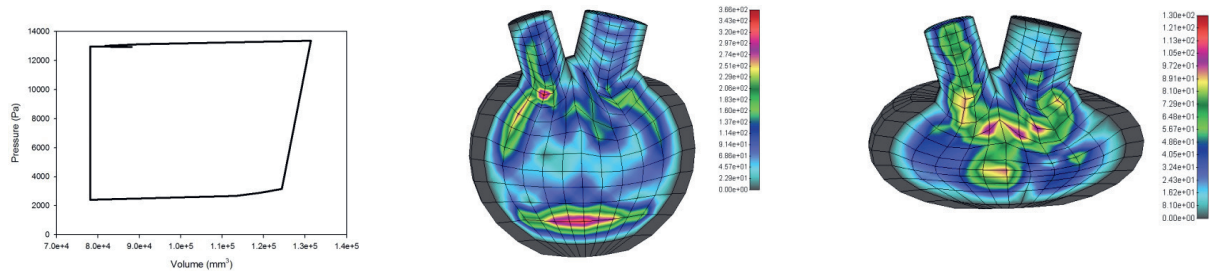


Fig. 8 PV diagram, velocity distribution in the diastolic phase, velocity distribution in the systolic phase for the case after Entresto treatment.

3.3 Realistic geometry of the heart model with left chamber and atrium parts

Using experimental data and DICOM files provided from specific patient, we have reconstructed a realistic heart model as STL format with left atrium (Fig. 9a, marked blue) and chamber part (Fig. 9a, noted yellow) with the accompanying mitral valve cross-section between (Fig. 9a, marked green), and also aortic part (Fig. 9a, marked orange) of the model with aortic cross-section included in fluid part of the model, which is surrounded by solid wall (Fig. 9a, wireframe). Finite element model consists of 1.5M hexahedral 3D elements, divided by 1M nodes. Model geometry is generated using STL files. Solid nodes are constrained around inlet/outlet cross-sections (Fig. 9a; red and magenta rings), and in the zone close to the mitral valve cross-section. Other solid nodes are free. In the Fig. 9c, two cross- section regions are marked to define prescribed inlet and outlet zones. Inside the fluid domain, mitral valve cross-section is presented (part of the model between ventricle and atrium; Fig. 9c, red line). Fiber direction in the solid domain of the realistic heart model is shown in Fig. 9b, and section C on the same Fig. shows distribution of the velocity field in the realistic heart model, at 0.1s. It can be seen that velocity values are the highest at inlet and outlet boundary cross-sections (red and green lines, Fig. 9c), which is logical due to prescribed inlet function and prescribed values at that cross-section at the beginning of simulation. Regarding the material models used, we have selected Holzapfel material model for obtaining passive stresses in the heart wall, and for muscle activation Hunter material model for active stresses is used.

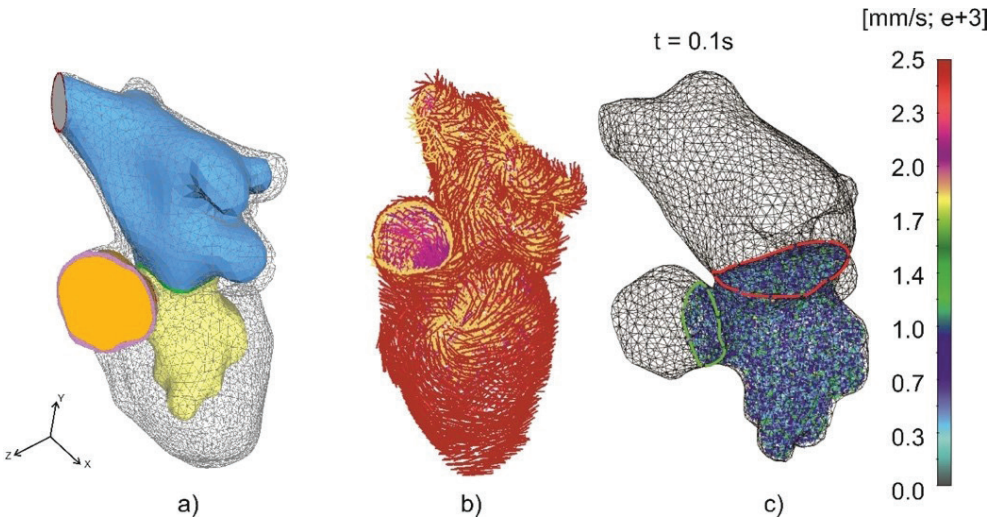


Fig. 9. a) Realistic heart FE model with representative cross-sections and fluid parts; b) Direction of fibres in solid part of realistic model; c) Fluid velocity field at 0.1s (mitral and aortic cross-section noted)

The prescribed inlet velocity function profile is shown in Fig. 10a, and aortic valve cross-section, while outlet velocity function profile is shown in Fig. 10b. Activation of the muscle is achieved using calcium function, displayed in Fig. 10c.

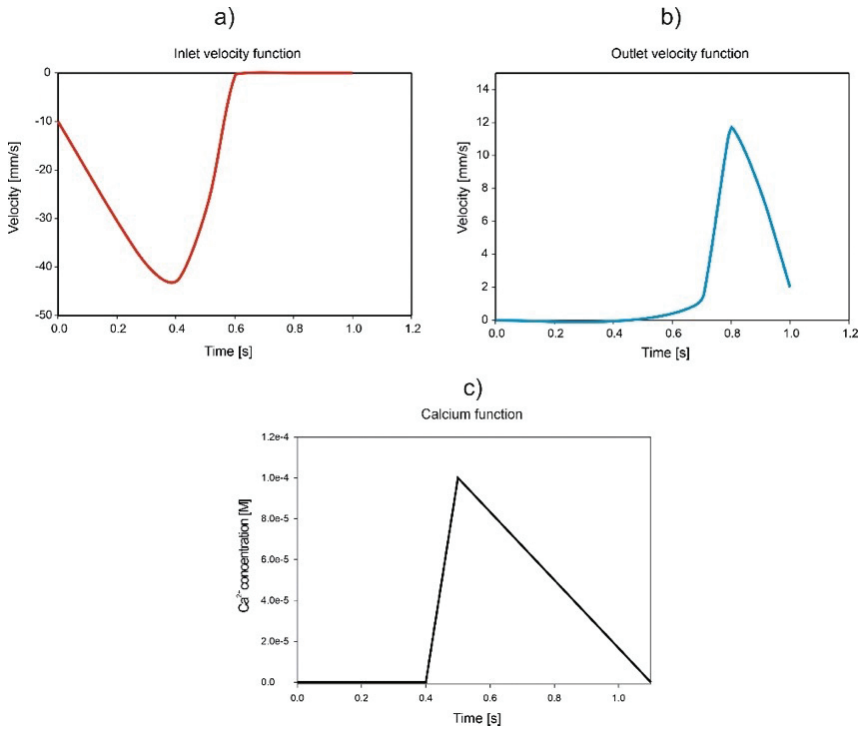


Fig. 10. a) Inlet function of velocity, at mitral valve cross-section; b) outlet velocity function – at aortic valve cross-section; c) Calcium concentration function used for activation of the muscle

Field of displacements in solid wall of realistic model of heart, during four different time steps of one cardiac cycle, is given in Fig. 11. At first step (0.1s), just the passive part of the material model has an impact on solid wall structure and until 0.4s of simulation model volume is increasing until the mitral cross-section is opened and fluid flows into the left chamber part. When the mitral valve is closed and injection of fluid is finished, fluid starts to eject from the chamber through the aortic cross-section, calcium function inside Hunter material model starts to act (0.5s), causing the start of the muscle contraction until the 0.9s of simulation after which model slowly returns to its undeformed state.

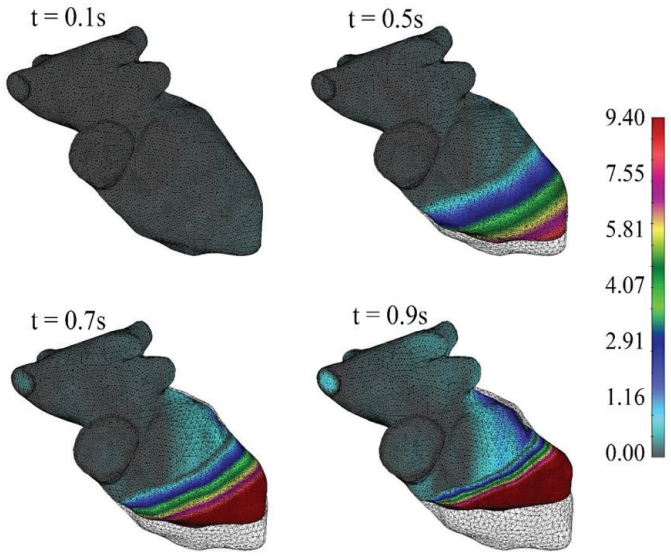


Fig. 11. Field of displacements in the solid wall of the realistic heart model; four different time periods. Undeformed configuration noted as black mesh

The large-scale model of the total heart with deformation simulation and mesh generated with 3M finite elements has been given in the Fig. 12. All of these models represent integration of Big Data technology, HPC and FEM computing. A very specific hardware and software technology has been used to support this integration. Some of the examples are EU projects SILICOFCM for in silico clinical trials.

Fig. 12. Large scale model of total heart. a) Displacement simulation b) Mesh generated with 3M finite elements

4. Discussion and Conclusions

The main result of the SILICOFCM project is a multi-modular, innovative in silico clinical trials solution for the design and functional optimization of whole heart performance and monitoring the effectiveness of pharmacological treatment, with the aim of reducing animal testing and human clinical trials. The SILICOFCM platform is based on the integrated multidisciplinary and multiscale methods for the analysis of patient-specific data and development of patient-specific models for monitoring and assessing patient condition through the progression of disease. The STRATIFYHF project is to develop and clinically validate a truly innovative AI-based decision support system for predicting the risk of heart failure, facilitating its early diagnosis and progression prediction that will radically change how heart failure is managed in both primary and secondary care

Heart modelling for cardiomyopathy and electromechanical coupling of the left ventricle were analyzed in the SILICOFCM (www.silicofcm.eu) and STRATIFYHF (www.stratifyhf.eu) project. Automatic left ventricle segmentation and the geometric parameter model which was extracted from echocardiographic apical and M-mode view was done using the U-net architecture. We have developed a model which includes three kinetic processes of sarcomeric proteins interactions: (i) kinetic transition between three cross-bridge states (a detached state and attached pre- and post-power stroke states; (ii) Ca^{2+} regulation of thin-filament switches between blocked and open states (i.e. by azimuthal movement of regulatory units (RU) containing troponin-tropomyosin complexes); and (iii) process of myosin binding to actin when RUs are in an open state. The drug actions are different for treating a variety of symptoms associated with cardiomyopathies. In particular, drugs simulated with MUSICO (Mijailovich et al., 2021) are divided into three major groups defined by the principal action of each drug, such as modulating calcium transients or changing kinetics of contractile proteins.

We have presented PV diagrams related to different patient cases, pressure diastolic distribution and pressure systolic distribution before and after Entresto and Digoxin drug treatment. Different drug pathways which directly affect the functional heart working have been analyzed. These drugs have a direct influence on the cardiac PV diagrams and ejection fraction. Triangular, parabolic, steep, shifted parabolic, parabolic wider and corresponding PV diagrams for different Ca^{2+} concentration functions have been presented.

Some limitations of the study are the lack of details regarding physical and biological properties of the heart and the need for subject-specific estimation of parameters from limited, noisy data, typically obtained using non-invasive measurements. Also, limitations are large-scale finite element calculations which can take up to several hours.

Computational platforms such as SILICOFCM and STRATIFYHF are novel tools for risk prediction of familial cardiomyopathy and heart failure in a specific patient that will certainly open a new avenue for in silico clinical trials in the future.

Acknowledgments:

This study is supported by the European Union's Horizon 2020 and Europe research and innovation program under grant agreement SILICOFCM,

STRATIFYHF and the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, contract number [451-03-65/2024-03/200107 (Faculty of Engineering, University of Kragujevac)]. This article reflects only the author's view. The European Commission is not responsible for any use that may be made of the information the article contains.

References

Andreu-Perez, J., Poon, C., Merrifield, R., Wong, S., & Yang, G. (2015). Big data for health. *IEEE journal of biomedical and health informatics, 19*(4), 1193-1208.

Armbrust, M., Fox, A., & Griffith, R. (2010). A view of cloud computing. *Commun ACM, 54*(4), 50-58.

Belle, A., Thiagarajan, R., Soroushmehr, S., Navidi, F., Beard, D., & Najarian, K. (2015). Big data analytics in healthcare. *BioMed research international*, 370194.

Bosch J. G, et al. (2002). Automatic segmentation of echocardiographic sequences by active appearance motion models. IEEE Transactions on Medical Imaging, 21 (11), 1374–1383, doi: 10.1109/tmi.2002.806427.

Dean, J., & Ghemawat, S. (2008). MapReduce: simplified data processing on large clusters. *Communications of the ACM, 51*(1), 107-113.

Demchenko, Y., Grosso, P., De Laat, C., & Membrey, P. (2013). Addressing big data issues in scientific data infrastructure. In: *2013 International conference on collaboration technologies and systems (CTS)* (pp. 48-55). IEEE.

Dilsizian, S. E., & Siegel, E. L. (2014). Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Current cardiology reports, 16*, 1-8.

Elliott, P. M. et al. (2014). European Heart Journal, 35(39):2733-79.

Filipovic, N., Sustersic, T., Milosevic, M., Milicevic, B., Simic, V., Prodanovic, M., & Kojic, M. (2022). SILICOFCM platform, multiscale modelling of left ventricle from echocardiographic images and drug influence for cardiomyopathy disease. *Computer Methods and Programs in Biomedicine, 227*, 107194.

Gibbons Kroeker, C. A., Adeeb, S., Tyberg, J. V. and Shrive, N. G. (2006). A 2D FE model of the heart demonstrates the role of the pericardium in ventricular deformation, American Journal of Physiology, vol. 291, no. 5, pp. H2229–H2236.

Hershberger et al. (2010). Official journal of the American College of Medical Genetics, 12(11):655-667.

Kojic, M., Milosevic, M., Simic, V., Milicevic, B., Geroski, V., Nizzero, S., Ziemys, A., Filipovic, N., Ferrari, M. (2019). Smeared Multiscale Finite Element Models for Mass Transport and Electrophysiology Coupled to Muscle Mechanics, Frontiers in Bioengineering and Biotechnology, ISSN 2296-4185, (7) 381, 1-16, 2296-4185.

Kojic, M., Milosevic, M., Simic, V., Milicevic, B., Geroski, V., & Nizzero, S. (2019). Smeared multiscale finite element models for mass transport and electrophysiology coupled to muscle mechanics. *Frontiers in Bioengineering and Biotechnology, 7*(381).

Kouanou, A. T., Tchiotsop, D., Kengne, R., Zephirin, D. T., Armele, N. M., & Tchinda, R. (2018). An optimal big data workflow for biomedical image analysis. *Informatics in Medicine Unlocked, 11*, 68-74.

Lavignon, J. F., Lecomber, D., Phillips, I., Subirada, F., Bodin, F., Gonnord, J., & Muggeridge, M. (2013). ETP4HPC strategic research agenda achieving HPC leadership in Europe.

Luo, J., Wu, M., Gopukumar, D., & Zhao, Y. (2016). Big data application in biomedical research and health care: a literature review. *Biomedical informatics insights, 8*, BII-S31559.

McNally, E. M. et al. (2013). The Journal of Clinical Investigation, 123(1):19-26.

Mijailovich, S M., Prodanovic, M., Poggesi, C., Geeves, M. A., Regnier, M. (2021). Multiscale modelling of twitch contractions in cardiac trabeculae, J Gen Physiol, 153 (3).

Mijailovich, S. M., Stojanovic, B., Nedic, D., Svicevic, M., Geeves, M. A., Irving, T. C., Granzier, H. (2019). Nebulin and Titin Modulate Cross-bridge Cycling and Length Dependent Calcium Sensitivity J Gen Physiol 151(5), 680-704.

Moradi. S. et al. (2019). MFP-Unet: A novel deep learning-based approach for left ventricle segmentation in echocardiography, Physica Medica, 67, 58–69.

Noble, J. A. and Boukerroui, D. (2006). Ultrasound image segmentation: a survey, IEEE Transactions on Medical Imaging, 25, 8, 987–1010, doi: 10.1109/tmi.2006.877092.

Parashar, M. (2014). Big data challenges in simulation-based science. *DICT@ HPDC*, 1-2.

Pullan, A. J., Buist, M. L. and Cheng, L. K. (2005). Mathematically Modelling the Electrical Activity of the Heart – from Cell to Body Surface and Back Again, World Scientific.

Razzak, M. I., Naz, S., & Zaib, A. (2018). Deep learning for medical image processing: Overview, challenges and the future. *Classification in BioApps: Automation of Decision Making*, 323-350.

SILICOFCM: In Silico trials for drug tracing the effects of sarcomeric protein mutations leading to familial cardiomyopathy, 777204, 2018-2022, www.silicofcm.eu

STRATIFYHF: Artificial intelligence-based decision support system for risk stratification and early detection of heart failure in primary and secondary care, No 101080905, 2023-2028, www.stratifyhf.eu

Tchito Tchapga, C., Mih, T. A., Tchagna Kouanou, A., Fozin Fonzin, T., Kuetche Fogang, P., Mezatio, B. A. & Tchiotsop, D. (2021). Biomedical image classification in a big data architecture using machine learning algorithms. *Journal of Healthcare Engineering*, 1-11.

Trudel, M. C., Dubé, B., Potse, M., Gulrajani, R. M. and Leon, L. J. (2004). Simulation of QRST integral maps with a membrane-based computer heart model employing parallel processing, IEEE Transactions on Biomedical Engineering, vol. 51, no. 8, pp. 1319–1329.

White, T. (2015). *Hadoop: The Definitive Guide.* Sebastopol, CA: O'Reilly Media, Inc.

Dragorad A. Milovanović[1]
University Business Academy in Novi Sad,
Faculty of Organizational Studies "Eduka" in Belgrade
Novi Sad (Serbia)

UDC 61:004.8
004.8:174
Review scientific article
Received: 29.04.2024.
Accepted: 02.07.2024.
doi: 10.5937/napredak5-50751

Responsible and ethical application of artificial intelligence systems in healthcare through a case study in diagnostics and personalized medicine

Abstract: The paper discusses new challenges of responsible and ethical application of artificial intelligence (AI) systems in healthcare. Academic research and commercial development focused on medicine are showing exponential growth; however, regulatory requirements for clinical use and commercial introduction are progressing more slowly. After classifying AI applications and identifying challenges and risks, a case study in diagnostics and personalized medicine is presented. The main result is a clear overview of the upcoming trends in medicine that can help legislators prepare to create new regulations and ethical principles of practice for new AI technologies. It is essential that technical experts, policy makers, legislators and other decision-makers accept responsibility.

Keywords: artificial intelligence, regulation, case study, medicine

Introduction

The use of artificial intelligence (AI) systems and software in medicine and healthcare is expanding. Academic research and commercial development are growing exponentially, but regulations for clinical application and global commercial market are growing much more slowly. The application of AI brings specific advantages, but also risks, so that there is a need for responsible and transparent systems which fulfil regulatory and ethical standards (Milovanović & Terzić, 2023).

Ethical guidelines of the development, application and use of reliable and responsible artificial intelligence are aimed at introducing a preventive mechanism which enables responsible development and verification procedures. There is a need for AI systems in compliance with the highest ethical and security standards regarding personal data protection, protection against discrimination in the

[1] dragoam@gmail.com ; https://orcid.org/0000-0001-6430-7421

application of machine learning and the establishment of responsible development of artificial intelligence in compliance with international ethical principles. We stress the importance of policy options the aims of which are the minimization of medical artificial intelligence risks, improvement of management and strengthening of responsible development. It is essential that technical experts, policy makers, legislators and other decision-makers accept responsibility (Milovanović, Terzić, Vučetić, 2023).

The EU strategy is based on high standards of protection against social risks represented by AI, unlike the US strategy which is developed mainly through private sector initiatives and self-regulation, and China's strategy which is essentially run by the government and characterized by a strong coordination of private and public investments in AI technologies (Papakonstantinou & DeHert, 2024). At the beginning of 2024, the EU Parliament adopted the Artificial Intelligence Act which imposes significant penalties for the failure to observe the regulations about the use of the forbidden AI systems as of the beginning of next year (Müller & Kettemann, 2024). This Law applies to AI providers in the EU market, but also to users in the EU and the providers and users in a third country if output results are used in the EU. The Law contains substantial assessment elements for high-risk systems which are completely forbidden, then high-risk systems which are not forbidden in themselves, but are significantly regulated and, finally, transparency and innovation support measures are introduced for all systems. The EU guidelines promote a reliable AI system which, in line with all the applicable laws and regulations, ensures respect for ethical principles and values, both from technical and social perspectives robust in avoiding unintentional causing of damage. It is essential that software and hardware artificial intelligence systems are directed towards man, that they are developed, applied and used in compliance with the key ethical requirements of human action: supervision, robustness and safety, data privacy and management, transparency, diversity, non-discrimination and fairness, social and ecological prosperity and responsibility (Coeckelbergh, 2020).

> Serbia has adopted a strategy of artificial intelligence development by 2025, an action plan and ethical guidelines for the development, application and use of reliable and responsible artificial intelligence. A draft strategy is currently being prepared for the period by 2030.

Challenges of responsible and ethical application

AI technology is being progressively developed and introduced in almost all areas of medicine, from primary protection to rare diseases and biomedical research. It is expected that many aspects of management regarding healthcare administration (increased efficiency, quality control, more conscientious business operations) and policy benefit from the new tools mediated by artificial intelligence (Gerke, Minssen, Cohen, 2020).

AI systems in healthcare can be classified on the basis of user groups of stakeholders: patients and citizens, clinicians and carers, medical workers, public health experts and policy makers. The basic domains of AI technology application in healthcare are clinical practice, clinical research, new medications, personalized medicine, public healthcare and global health, and healthcare administration. The classification of medical AI tools can also be based on the environment in which the tools are used: clinical environments (hospitals, primary healthcare centres, emergency centres), clinical processing and setting management (laboratory, pharmacy, radiology) and administration. We will adopt a more comprehensive classification of AI applications clinical practice, research, public health and administrative applications.

Clinical practice. The potential of AI application in the clinical environment is huge and ranges from the automation of diagnostic processes to therapy decision-making and clinical research. The data necessary for diagnosing and treatment come from numerous sources, including clinical notes, laboratory tests, pharmacy data, medical imaging and genome information. Radiology is a branch of medicine primarily dealing with discovering diseases and injuries – namely diagnostics, and that is why it has undergone substantial development with AI application in the past few years. AI technologies for screens processing help radiologists in interpreting medical images. For example, image segmentation with limited human supervision has been achieved by using deep neural network (DNN) models for automatic localization and marking of anatomic structure borders. AI tools can suggest a priority and monitor findings which require early attention, and enable radiologists to concentrate on the most probable pathological images. Radiomix is yet another image processing technique I which AI has proved to be useful. Although the term is not strictly defined, the general goals are a quantitative analysis of diagnostic images and treatment planning (Larson et al., 2021). The features include the characteristics of tissues and lesions such as their heterogeneity and shape, and can be used for resolving clinical problems independently or in combination with demographic, histological, genomic or proteomic data, which refer to patient's proteins, including their expression, modification and interaction. The importance of Radiomix is even greater when a large amount of information is processed with the aid of AI techniques.

Biomedical research. Medical research has numerous benefits from the solutions derived from artificial intelligence in comparison to clinical applications, while recent progress also show promising AI applications in discovering clinical knowledge. For example, the main sources of medical knowledge still use MI, algorithms for search result ranging, including algorithms which permanently learn from the users' behaviour in the search. For example, PubMed is a widely-used browser for biomedical literature. BestMatch search algorithm uses users' intelligence and most modern MI, technology as an alternative to the traditional data sorting sequence. BestMatch algorithm is "trained" by using previous user searches with dozens of signals (factors) of relevance ranging (the most important are the previous use of an article, its publication date, relevance grading and

article category). The algorithm significantly improves finding relevant information in relation to the implied time sequence in PubMed and with time it has increased the use of relevance search. Thanks to techniques such as information extraction, automatic summing and deep learning (DL) artificial intelligence has the potential to transform static narrative articles into patient-specific clinical evidence. Designers of medications intensively apply MMI techniques in searching chemical information in big data bases of compounds in order to discover new medications. The application focus is on the development of AI approach in implementing innovative modelling based on the diverse nature of data sets for medications. AI models contribute to a better understanding of the wide range of medications and clinical outcomes they may offer.

Clinical research. Randomized Controlled Research (RCT) is the most powerful method for risk assessment and benefiting from any medical intervention. However, the implementation of RCT is not always feasible. Common difficulties of unsuccessful RCT include the wrong selection of patients, inadequate randomization, insufficient sample size and wrong selection of end points. AI models are trained for a better selection of study participants by advanced statistical methods and for assessing the study end points by a data-based method. AI application generates a more efficient performance and a larger statistical strength than traditional RCT. Apart from the efficient selection process, a sufficiently large sample is critical for enabling the discovery of statistically significant differences between the groups.

Personalized medicine. It is important to understand scientifically how unique characteristics of an individual patient, such as molecular and genetic profiles, make the patient vulnerable to a disease and responsibe to a therapeutic treatment. The original concept of personalized medicine has been expanded to cover other properties and individual clinical characteristics, and a new concept has finally been formed and called expanded personalized medicine on the basis of additional sources of information such as clinical sources, demographic data, social data, lifestyle parameters (quality of sleep, physical activity, eating habits) and environmental conditions. AI tools improve progress in personalized medicine by assessing clinical benefits of various research methods and several types of data.

Global health. Public health covers disease prevention, life extension and health improvement through organized efforts and informed choices of society, public and private organizations, communities and individuals. Experiments with relevant AI solutions are currently underway within many areas of public health. AI can help to identify specific demographic or geographical locations with a distributed disease or high-risk behaviour. The scope of AI solutions which can improve supervision of a disease is also significant. Digital epidemiological supervision refers to integration of supervision based on cases and events (news ad online media, sensors, digital traces, mobile devices, social media, microbiological laboratories and clinical reporting) for the purpose of analysing the approach to threat verification. The aim is to build early warning systems for undesired events in relation to med-

ications and low air quality. AI has the potential to intensify contact with patients, as well as to personalize services. The essential component of the initiatives involves contacting a large number of patients via different automated, simply scalable methods, such as text messages and portals for patients. AI application in public health also includes a broader perspective of healthcare policy and management. Studies cover the research of artificial intelligence with the aim of improving the performances of healthcare institutions and allocating resources from the system perspective.

AI healthcare administration. Healthcare systems are characterized by a complex administrative work flow with a wide range of actors and institutions, including patients (insurance collection management), healthcare workers, healthcare institutions and organizations (patient flow), laboratories (supply chain of consumables), pharmacies, taxpayers and regulators. Within primary healthcare, potential problems have been identified as follows: time necessary for financial compensation, data input into various non-integrated information systems based on practice, processing information from hospitals and other external service providers and helping patients in the fragmented healthcare system. AI can perform routine tasks in a more efficient, accurate and impartial manner. One of the arguments in favour of its use in administrative practice is that errors in these activities are less serious than those in the clinical environment. Nevertheless, the questions of security and lack of privacy still remain. AI applications may be critical in the organization of the patient flow. For example, the lack of availability of hospital beds is an important cause of cancelling surgical interventions; however, it may constitute an administrative error which may be avoided in the patient flow. The problem frequently occurs and is also related to delayed release from the clinical department.

Below we will identify the main risks and potential consequences of AI application in medicine and healthcare:

- Patients' injury due to AI errors (main causes are noise and artefacts in clinical inputs and measurements, unexpected variations in clinical contexts and environments; medical consequences are inadequate treatment and wrong appointments or determining the priority of interventions);
- Abuse of medical AI tools (potential causes are the lack of training and sufficient explanation and information; improper use of AI tools may lead to wrong medical assessment and decision-making and, subsequently, to potential harm to patients);
- AI partiality and continuance of existing inequalities (most frequent causes are the lack of transparency, imbalanced data sets based on structural partiality and discrimination, as well as the lack of diversity and interdisciplinarity in technological, scientific, clinical and policy-making teams);
- Issues of privacy and security (the cause of which is the lack of transparency regarding the design, development, evaluation and application of AI tools; specific risks include the lack of understanding and trust, difficulties in the autonomous work

of reproducing and assessing algorithms, difficulties in identifying the source of errors and defining responsibility);
- Gaps in responsibility (main risks include sharing personal data without a fully informed consent, data requalification without the patient's knowledge, data violation both at the individual level and at the hospital or healthcare system level);
- Algorithm responsibility (the key aspect of reliable and applicable AI, legal gaps still existing in the current national and international regulations, not being simple to define roles and responsibilities due to multiple actors involved in the process, from design to application, particularly if the used AI model is not fully transparent);
- Obstacles in the implementation (limited data quality, structuring and interoperability of AI tools with the existing clinical work flows and electronic healthcare systems).

We have clarified basic clinical, social and ethical risks of AI application in healthcare: potential errors and harm suffered by patients, risk of impartiality and increased healthcare inequalities, the lack of transparency and trust, as well as vulnerability to hacking and violation of data privacy. Challenges must be addressed and prevented, so that it is necessary to monitor the processes of international standardization and patent registration.

Analyzing international registered patents with AI application in medicine it is possible to predict the trends of new technologies by specialities which require new regulations. A relatively recent study from 2022 identifies more than 10,000 healthcare patients in the past ten years, which has enabled regulators to perceive clear trends of the upcoming AI technologies and necessary strategic activities. First, it is possible to predict the areas on which to direct the regulatory focus, how patents are followed by products/devices ready for the market with a predictable delay. Medical specialties emphasized by a significant number of registered patents are radiology, oncology and ophthalmology, and therefore these areas will have the largest number of new technologies in clinical practice in the following years.

Case study

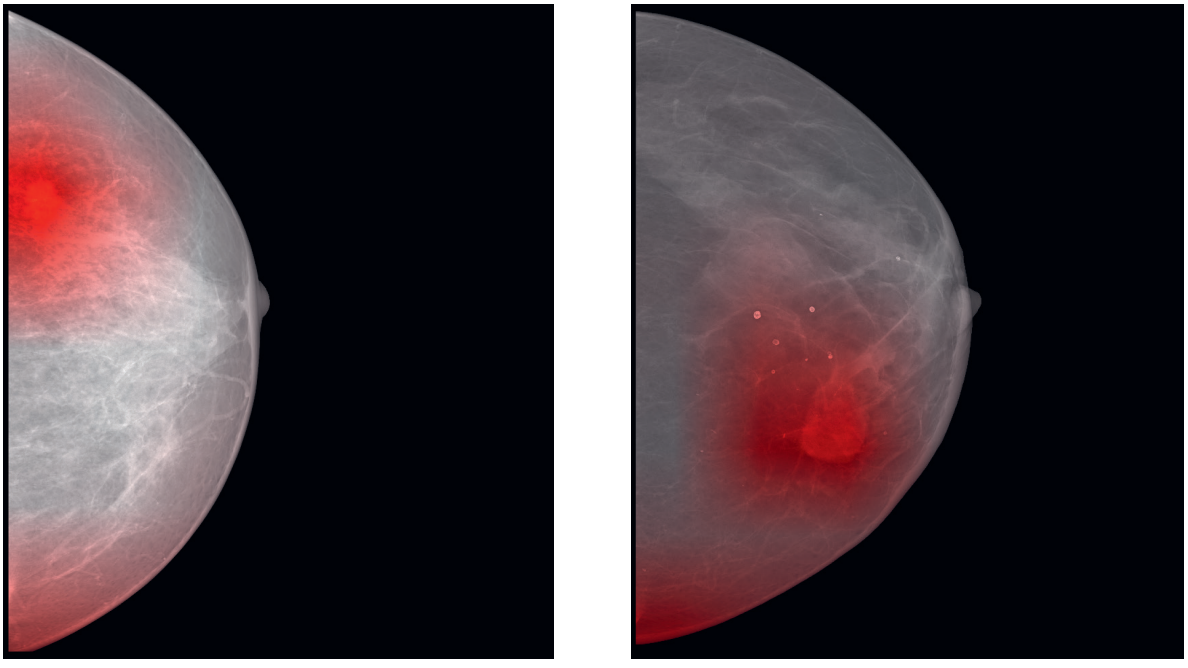
Artificial intelligence possesses significant potential for improving various aspects of healthcare and it covers diagnostics, treatment, data management and increasing efficiency of healthcare systems (Recht et al., 2020). We emphasize the trends of AI effects on healthcare:

- **Diagnostics.** The development of AI systems for analyzing and interpreting medical screens (US, RO, CT, NMR, PET) may improve diagnosis speed and accuracy and be helpful in the work of clinical doctors.
- **Personalized treatment.** The use of AI for the analysis of genetic data and other factors in order to adjust the treatment to patients' individual needs.
- **Predictive analytics.** The implementation of AI systems in the analysis of a large amount of data in order to predict the risk of certain diseases or complications, enabling preventive measures.

- **Telemedicine and mobile applications.** The AI integration in the applications for health monitoring and telemedical platforms ensures better monitoring of patients and providing remote personalized care.
- **Automation of administrative processes.** The use of AI in the automation of administrative tasks, such as making appointments for examinations, invoicing and data management.

Together with the partners from the healthcare sector, they work on the development of AI tools for diagnostics, personalized medicine and improvement of efficiency of healthcare services, as well as the design of new medications. The aim is to ensure faster and simpler carcinoma detection, to reduce the complexity of rare disease diagnostics and to shorten the time until giving an

> The researchers from the Institute for Artificial Intelligence Research and Development of Serbia (IVI, 2024) implement a number of projects in the area of healthcare: diagnostics of cancer and rare diseases, as well as the development of new medications.

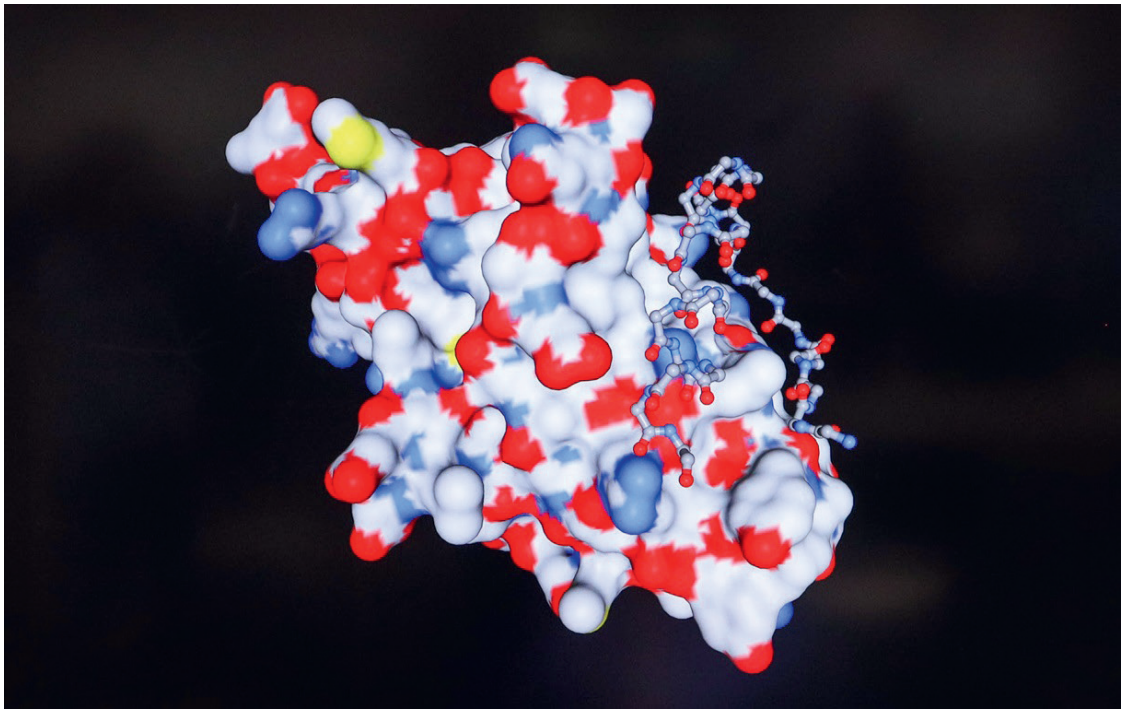


The Institute for Artificial Intelligence Research and Development of Serbia, with the aid of artificial intelligence, analyzes and ranks patients' mammographic images with the aim of faster diagnosis of carcinomas.
Photo: Institute for Artificial Intelligence Research and Development

accurate diagnosis, to reduce the pressure on the healthcare system and, consequently, to redirect treatment costs towards an adequate innovative therapy, as well as to accelerate the process of finding and developing mediations and to reduce pharmaceutical research costs.

The first project includes AI analysis and ranging of patients' mammographic screens. The procedure is that early screening results indicate that no further monitoring of the patient is required or that additional analyses are necessary. The solution is to range screens by their importance: the first place is for the screens which need further attention, interpretation, analysis or biopsy, while the last place is for the patients who do not need an urgent examination and who are probably healthy. Of course, a radiologist makes the final decision and that is why AI tools only ensures better organization and optimization of doctors' working hours.

The project of diagnosing rare diseases is important due to the observed long time interval between the patient's registrations of the first symptoms to the actual diagnosis. The use of



The Institute for Artificial Intelligence Research and Development of Serbia, in cooperation with a startup from the United States of America, develops a platform based on generative artificial intelligence for the design of binding peptides to a desired protein.

Photo: Institute for Artificial Intelligence Research and Development

modern AI techniques is examined in terms of reducing the complexity of diagnosing Fabry disease, shortening the time for giving an accurate diagnosis, reducing the pressure on the healthcare system and, consequently, on redirecting treatment costs towards an adequate innovative therapy. Exceptionally rare diseases are on the increase, but there is still a problem of doctors not examining frequently the patients suffering from rare diseases and, in addition, the symptoms are quite different. Thanks to NLP models, the researchers from the Institute have examined huge amounts of medical data about 50,000 patients and a certain knowledge corpus has been generated about the patients potentially diagnosed with Fabry disease. The ranking list has been made on the basis of the criteria of the speciality doctors about the most important symptoms, as well as the minor ones. The ranking of the patients has been generated for testing at the doctor's invitation. Early detection of rare diseases significantly extends and improves the quality of patients' lives, so that the development of AI tools also contributes to substantial savings in the healthcare system.

The Institute works on AI application projects in cooperation with pharmaceutical industry on discovering and producing new medications. With the advanced techniques of machine learning of PL and AI, huge data sets from biomedical research are analyzed. Trained AI systems then identify patterns and relations of different biological entities and molecules, which further enables the evaluation of the efficiency of molecules as potential medications. There is also an emphasis on the development of generative AI models in creating molecules which are not present in standard libraries of known molecules used for therapeutic purposes. Models expand the chemical space of searching for new medications on a large scale and support therapies for the so far untreated diseases. The project is rather complex and also includes the analysis of a large amount of biomedical data, as well as the development of generative AI models which re3quire laboratory corroboration of potential molecules as candidates for the evaluation of the model quality.

Conclusion

The paper considers current regulatory and socio-ethical aspects of AI in medicine and healthcare. Artificial intelligence may make a revolutionary transformation in the way we provide and receive healthcare services, through personalized medicine, more efficient diagnoses and treatment, better outcomes for patients and the optimization of the healthcare system. Predicting AI trends in healthcare is challenging and exciting.

With responsible development and application, AI can support a more precise, efficient and available medical service. It is difficult to overcome technical and ethical challenges of the AI system application in order to use maximally the potential and avoid potential risks. The future of healthcare is inevitably related to AI and it is our task to shape it in the way that brings benefits to society.

References

Coeckelbergh, M. (2020). *AI ethics*. Cambridge. MA: MIT Press.

Gerke, S., Minssen, T., Cohen, G. (2020). Ethical and legal challenges of artificial intelligence-driven healthcare. In: A. Bohr, K. Memarzadeh (eds.) *Artificial Intelligence in Healthcare* (295–336). Cambridge, MA: Academic Press.

IVI, *Institute for Artificial Intelligence Research and Development of Serbia* (2024). Available at https://ivi.ac.rs/en/news/.

Larson, D. B. et al. (2021). Regulatory frameworks for development and evaluation of artificial intelligence–based diagnostic imaging algorithms: summary and recommendations. *Journal of the American College of Radiology*, XVIII (3), 413–424. DOI: https://doi.org/10.1016/j.jacr.2020.09.060.

Milovanović, D., Terzić, R. (2023). Responsible application of Artificial Intelligence in healthcare: ethical principles of practice and new regulations. In: M. Ivković, D. Drašković (eds.) *Proceedings of the 29th ICT Conference YU INFO 2023* (46–49). Beograd: Informaciono društvo Srbije. [In Serbian]

Milovanović, D., Terzić, R., Vučetić, Č. (2024). Trends of application Artificial Intelligence in healthcare: challenges and regulation issues. In: *Proceedings of the 30th ICT Conference YU INFO 2024.* In preparation [In Serbian]

Müller, M., Kettemann, M.C. (2024). European approaches to the regulation of digital technologies. In: H. Werthner et al. (eds.) *Introduction to Digital Humanism* (623–637). Cham: Springer.

Papakonstantinou, V., DeHert, P. (2024). *The Regulation of Digital Technologies in the EU*, Oxfordshire, England, UK: Routledge.

Recht, M. P. et al. (2020). Integrating artificial intelligence into the clinical practice of radiology: challenges and recommendations. *European Radiology*, XXX (6), 3576–3584. DOI: https://doi.org/10.1007/s00330-020-06672-5

Danimir P. Mandić[1]
University of Belgrade,
Faculty of Education
Belgrade (Serbia)

UDC 37.091.39:004.8(497.11)
Review scientific article
Received: 02.07.2024.
Accepted: 08.08.2024.
doi: 10.5937/napredak5-51939

A New Paradigm of Education and Potentials of Artificial Intelligence

Abstract: This paper addresses the possibility of using artificial intelligence (AI) in education and establishing rules, restrictions and procedures for the safe use of AI in different fields. In the past two decades, we have been trying to improve educational technology by using multimedia, hypermedia, virtual and extended reality, as well as educational software. It has led to the raised quality of learning, increased dynamics and obviousness of the teaching process, as well as encouraged motivation of students. In the past few years, the possibilities of using AI in education have been researched, particularly in the sphere of personalized learning, more dynamic teaching, as well as complex evaluation of students' work. The paper analyzes the benefits, challenges and risks in relation to the use of AI in education and emphasizes the importance of maintaining balance between AI and human interaction. It also analyzes the use of robots in education with the aim of encouraging creativity and critical thinking, as well as developing problem-solving skills through the development of algorithm thinking. In the past year, the application of ChatGPT has been intensified in education and that is why it is necessary to define rules and procedures in which this technology would contribute to teaching and learning.

Keywords: artificial intelligence, educational technology, digital competences

Introduction

Modern technologies bring intense changes in the field of engineering, traffic, economy and medicine, while changes occur somewhat more slowly in education. Today's development of educational technology is harmonized with the development of information technology, in line with the needs of students in the 21st century. Modern educational technology based on AI leads to the introduction of significant changes in the work methods, organization of teaching, as well as the evaluation of students' work. Integration of artificial intelligence technologies in education ensures

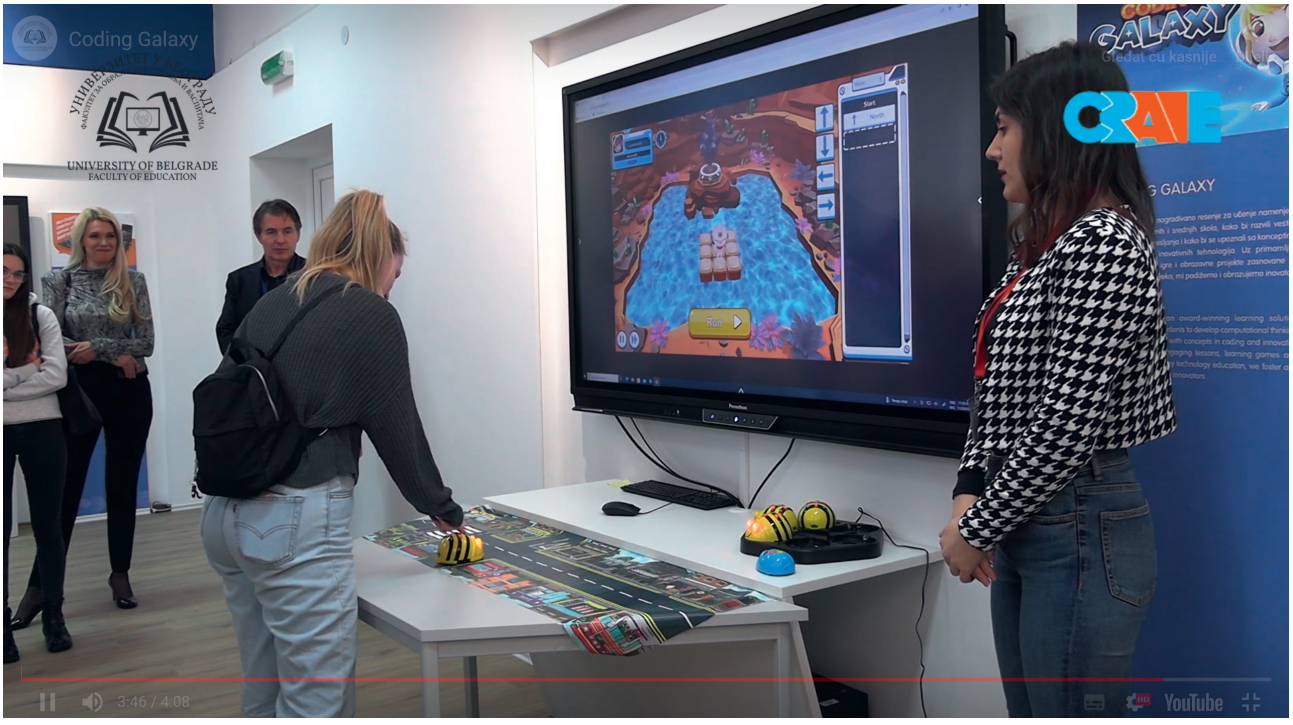
[1] Danimir.Mandic@uf.bg.ac.rs ; https://orcid.org/0000-0002-3986-6077

individualization and differentiation of teaching contents in line with the possibilities, prior knowledge and learning styles of every student. Robots and 3D simulations in education offer possibilities of practical learning and experimenting which are difficult to provide in traditional teaching, while they give special importance to encouraging research and activities of students in the process of acquiring new knowledge. Serbia ranks among the most avant-garde countries which have introduced subjects and content necessary for acquiring digital competences from pre-school age to the university level of education.[2] In pre-school age, through a project approach, the use of the simplest digital devices is ensured, while in the first four grades of primary school the content is studied in relation to the possibility of using digital devices, children's safety and the development of algorithm thinking. In higher grades of primary school and in secondary school, digital competences are further improved through the subject of computing, where students study content related to computer hardware, computer networks, data bases, as well as elementary programming. The curriculum stipulates content related to the manner of functioning and the application of AI in different fields, thus informing students about the way of learning about artificial intelligence, as well as about the possibilities of its application in different fields. Serbia is one of the leading countries in Europe in defining the Strategy of the Application of Artificial Intelligence until 2030, in which the application of AI in education will have a special place, including the effect of AI on teaching and learning technology. At the Faculty of Education in Belgrade, future educators and teachers acquire digital competences in line with the recommendations of UNESCO and OECD. The Republic of Serbia has defined the New Framework of digital competences for teachers, in which 25 digital competences grouped in six fields. Following the legal regulations and recommendations of the Ministry of Education, the Faculty of Education regularly innovates curricula so that future educators and teachers are adequately trained for most modern teaching and learning methods. In the Centre for Robotics and Artificial Intelligence in Education (CRAIE), which was founded in cooperation with the Beijing Normal University (BNU) and the software devel-

> The use of AI leaves more for teacher's interaction with students, the development of value attitudes and personal character traits, thus giving special significance to education, which has been neglected in our schools.

[2] Serbia is the first country in Southeast Europe which has adopted the Strategy for the Development of Artificial Intelligence in the Republic of Serbia for the period 2020–2025 ("Official Gazette of the RS", No. 96/19), and the preparation of a new strategy for the upcoming period is underway. Moreover, Serbia has joined the Global Partnership on Artificial Intelligence (GPAI) and will chair this organization in the next three years.

opment company NetDragon Websoft from China, the most modern equipment and applications have been installed in relation to 3D simulations, virtual reality, holograms, robotics and artificial intelligence. CRAIE is a research-development centre within the Faculty of Education, intended for the incapacitation of students and permanent specialization of teachers in the field of digital technologies. In education, from the pre-school age onwards, programmable robots are used, such as Bee-Bots, which encourage the development of algorithm thinking. Apart from educational robots, advanced humanoid robots are also used, which can help teachers and students in knowledge personalization, programming and performing different tasks and simulating real-life situations with permanent interaction and feedback, which enables monitoring, measuring and evaluation of each student's progress. They can help in teaching coding, mathematics, physics as well as other subjects, and make learning more interactive and pleasant. In their interaction with robots, students develop computing thinking, logical thinking and the ability of solving complex problems. The selection of adequate methods of work and modern didactic systems promotes cooperation among students and teamwork in the robot programming process, task solving and engaging in practical projects. These cooperation



Working with educational robots in the Centre for Robotics and Artificial Intelligence at the Faculty of Education

experiences cultivate communication skills, empathy, critical thinking and the ability to work efficiently in a team. They can be adjusted to students' individual needs and their prior knowledge, as well as abilities, offering adjusted content, activities and feedback.

By analyzing students' progress, robots can identify deficiencies in their knowledge and provide targeted interventions, ensuring that every student receives support necessary for success. In this exciting era of robotics in education, we will write about the possibilities of changing teaching and learning, accept innovation and witness the transformative effect of these intelligent machines. Students who learn with the aid of new technologies perceive the teaching process as a journey on which education encounters technology, thus empowering them to become life-long learners, think critically and solve problems in the constantly changing world.

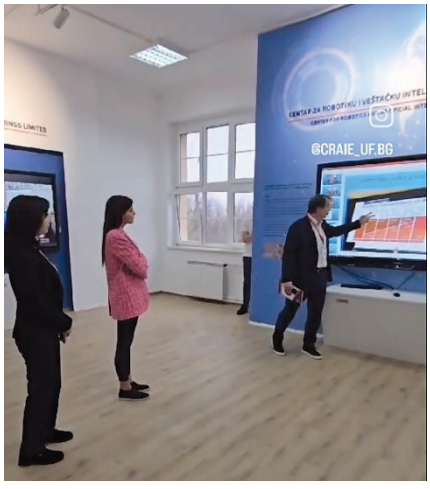
AI and teachers' competences

Since artificial intelligence is becoming more and more important in education, it is of essential relevance for teachers to develop new competences and skills in order to efficiently use and integrate AI technologies in their teaching theory and practice. Teachers should be informed about the possibilities offered by AI and trained for the skilful use of artificial intelligence tools and technologies. This includes understanding the way in which AI algorithms function, being familiar with the platforms and applications created by artificial intelligence and knowing how to use educational resources based on artificial intelligence. Moreover, teachers should be trained how to interpret and analyze AI-generated data in order to monitor, measure and evaluate students' progress and identify problems in learning. Continued feedback ensures students' greater motivation and creates good conditions for self-evaluation, thus making the learning process more efficient and effective.

Virtual reality in in the Centre for Robotics and Artificial Intelligence at the Faculty of Education

Teachers should understand how to integrate AI technologies efficiently in their content and how to adjust teaching methods and the system of evaluating students' work to modern educational technology. This includes harmonizing artificial intelligence tools with learning goals, modelling learning activities based on artificial intelligence and using artificial intelligence as support to different teaching strategies in the use of modern educational technology. With the development of new technologies, teachers should be flexible and ready to learn permanently and update their competences. This involves being informed about the latest technology in education, participating in the possibilities of professional development and cooperation with other teachers for the purpose of

exchanging best experiences. In addition, teachers must be aware of the ethical implications of using artificial intelligence in education and some problems with the applications which are not in line with ethical principles. They should understand the matters related to data privacy, security and algorithm partiality. Teachers need to ensure that AI tools are used with responsibility, ethically and in a manner which respects students' privacy and promotes fairness (Milutinović, Mandić, 2022).



Artificial Intelligence in in the Centre for Robotics and Artificial Intelligence at the Faculty of Education

It is very important that students should develop critical thinking and comparative analysis of information received from AI with the information found in reviewed textbooks. Teachers should be able to facilitate cooperation among students by using platforms and tools based on artificial intelligence. They should also be trained for cooperation with artificial intelligence systems, such as virtual assistants and chat-bots, in order to support individual learning and provide personalized feedback. While artificial intelligence may automatize certain tasks, teachers' role is to develop students' critical thinking skills and encourage their creativity. Teachers need to devise modern learning methods which encourage students to think critically about the AI-generated content, to ask questions about assumptions and to develop creative problem-solving skills. Although there is concern that AI might replace teachers, it is difficult to imagine because human interaction, empathy, socialization and educational work with students are irreplaceable. Teachers should focus on building permanent interaction with students, giving emotional support, cooperation and creating a stimulating and inclusive environment in the classroom. By developing these competences, teachers can efficiently use AI technologies for improving their teaching practices, personalizing learning experience and supporting their students' success. It is important for teachers to accept AI as a tool which supplements their professionalism and improves their ability to fulfil students' different requirements in the modern era (Mandić, 2023).

Teachers and AI

As we have already mentioned, a number of experts in education claim that AI-supported teaching will suppress teachers by giving students an opportunity for self-education and self-evaluation. It might make teachers unnecessary in a certain way, which would cause a revolution in teaching by putting students in the position of independent creators of learning. Based on research, it can be concluded that teaching, supported by smart devices based on AI with new methods, changes the teachers' role,

but cannot replace them. In line with these changes, AI-supported teaching makes it possible for the teacher to change his/her role of a lecturer, evaluator and primary source of knowledge into a new role which implies competences of organizers, planners, researchers, motivators and verifiers of students' work. Although both learning and teaching are key components of the educational process, the ultimate goals are to ensure students' better learning outcomes and their motivation for learning. AI can analyze individual information about students and provide personalized recommendations, content and activities adjusted to their specific needs and learning styles. This personalized approach helps students to be involved more deeply in the content and to optimize their learning outcomes. Evaluations based on artificial intelligence may be adjusted to students' responses, offering questions and tasks which are adjusted to their interests, abilities and prior knowledge. The software based on artificial intelligence can provide momentary feedback to students about their performance, thus enabling them to identify and correct mistakes in real time (Baker & Inventado, 2014).

Extended reality in in the Centre for Robotics and Artificial Intelligence at the Faculty of Education

This timely feedback helps students to understand their advantages and weaknesses, enabling them to make changes and improve their learning. AI can help students to access a wide range of educational resources and information. Platforms based on artificial intelligence can recommend relevant content, suggest further reading and provide access to online libraries, thus improving students' ability to research and deepen their understanding of topics. New technologies, such as virtual reality (VR), extended reality (ER) and play-based learning, may create impressive and interactive learning experiences. Teachers are becoming convinced that teaching and learning based on artificial intelligence, used in combination with other types of teaching, are a purposeful innovation. On the other hand, the curriculum preparation, verification and implementation require teachers' huge efforts, as well as material investments and pedagogically trained staff, which makes it more difficult than traditional education. For this reason, many curricula which are currently criticized have not been prepared professionally enough and have not been tested in practice by authorized experts. That might lead to a situation in which teachers would show resistance to the introduction of innovation if they believe it has not been sufficiently formed and devised in pedagogical terms. Teaching and learning based on artificial intelligence are still a relatively new and insufficiently elaborated concept and that is why many critical objections and some results should be taken conditionally. All the weaknesses of the teaching process we have spoken about will not be manifested in practice if it is professionally programmed and organized (Ristić, Mandić, 2018).

Teaching and learning results based on artificial intelligence fulfil the learning logic, open new learning possibilities, encourage students' activity, ensure progress at students' own pace, make the teaching process more attractive, provide economy and efficiency of the teaching process and thus belong to the category of teaching which is most attractive and suits young people's needs and interests. Most importantly, this new organization of teaching and learning should increase students' internal and external motivation and help them to acquire functional knowledge and to understand facts and procedures. Furthermore, it helps students to have more permanent knowledge. These technologies engage students, promote active learning and make the learning process more pleasant and attractive. AI can be adapted to various learning styles and preferences by being adapted to different students. For example, some students may prefer visual learning while others may prefer auditive or kinesthetic approaches. AI can ensure content and activities which fulfil individual learning preferences. Modern technologies can improve accessibility for students with special needs. Speech recognition based on artificial intelligence and text-to-speech conversion tools can help students with impaired vision or with learning disabilities by providing them with more efficient access to educational materials. AI can analyze large data quantities about students' performance and learning patterns by offering valuable insights to teachers. These insights can help teachers to understand the efficiency of their teaching strategies, to identify improvement areas and adjust their approaches to teaching in order to fulfil students' individual needs in a better way.

Students and AI

As education develops with the integration of artificial intelligence, students' roles also change. The essential change is that students will not be passive actors in formal education, but they should become active participants and researchers. They have access to personalized learning experiences, interactive and adaptable evaluations. Students are encouraged to take initiative for learning, to establish goals and make decisions in line with their interests and prior knowledge, as well as abilities. AI technologies provide students with tools and resources to be included in independent learning. Students can research topics of interest, access educational materials at their own pace and assume responsibility for progress in learning. Platforms based on artificial intelligence can offer recommendations and guidelines, empowering students to manage their own learning experiences. They ensure consistency in the learning pace and details of the presented content, maximum adaptability to students' individual abilities, their psychological traits, speed and learning style. Materials are logically selected in series in which acquiring further knowledge is logically related to prior knowledge, including constant connection of students with their teachers (D'Mello & Graesser, 2012).

AI encourages students to develop critical thinking and problem-solving skills. Instead of passively consuming information, students face the challenge of analysing, evaluating and synthetizing knowledge. AI technologies can offer possibilities to students to be involved in complex


ROBOT EMA
Zdravo! Ja sam Ema, posebno dizajniran humanoidni robot na bazi veštačke inteligencije koji je razvila kompanija NetDragon Websoft Holdings Limited, sa sedištem u Kini. Pored toga što mogu da klimam glavom i trepćem, imam napredne vizuelne i govorne sposobnosti za komunikaciju sa ljudima i interakciju sa njima uz pomoć mobilnog telefona, kao i veštine za izvršavanje određenih instrukcija zadatih od strane čoveka.
Hello! I am Emma, a specially designed humanoid AI robot developed by NetDragon Websoft Holdings Limited based in China. Besides doing simple gestures like nod and blink, I have advanced visual and speech abilities to communicate with people and interact with them through mobile phone, and skills for completing human instructions.

problem-solving tasks and to develop their analytical and creative thinking abilities. AI facilitates the cooperative learning experience, enabling students to work together on projects, share ideas and participate in debates. Students learn to cooperate efficiently, present their ideas and respect different perspectives. Tools based on artificial intelligence can support virtual cooperation, enabling students to be connected and cooperate beyond the boundaries of the physical classroom. With AI integration, students should develop digital literacy skills. They should be trained to use AI tools, to navigate digital platforms and to critically assess online resources. Students learn how to find their way in the digital environment with responsibility, how to understand ethical implications of artificial intelligence and become informed modern citizens. New technologies promote the culture of life-long learning. Students understand that learning is not limited to the classroom and that they should constantly improve their skills and adjust to the rapidly changing world. AI technologies can provide personalized recommendations for further learning, thus helping students to develop a way of thinking about progress and a desire for permanent improvement. Educational technologies may help students in expressing their creativity. Students can use tools based on artificial intelligence for digital art, composing music, narration and other creative enterprises. AI can provide students with new paths in researching and expressing their unique talents and perspectives. Thanks to artificial intelligence, students have the opportunity to express their opinions and to advocate goals in which they believe. They can use AI technologies for researching, data collection and efficient advocating of their ideas. Students become empowered advocates of social and ecological matters, using artificial intelligence to initiate positive changes in their communities. Since education includes artificial intelligence, students' roles shift from passive knowledge recipients to active participants in their own learning. They develop basic skills, learn permanently and contribute to cooperation and a more innovative

The photo on the previous page: Humanoid root Ema was developed by Chinese company NetDragon Websoft Holdings Limited which is used in the education of future teachers in the Centre for Robotics and Artificial Intelligence in Education, within the Faculty of Education, the University of Belgrade.
Photo: Faculty of Education

educational environment. By using artificial intelligence for the improvement of the learning process, teachers can create more interesting, personalized and efficient learning experiences for students. While teaching remains an essential component, the focus is shifted towards ensuring and supporting better learning outcomes for students and developing empathy and social competences (Mandić, Miščević, Bujišić, 2024).

Challenges and risks of AI application in education

In the use of AI to date, it is possible to observe three different approaches formed in certain countries. They differ by the teachers' attitude towards the use of AI, the benefits students have, as well as the risk of plagiarism. In the first group, teachers and experts on education are completely against the use of AI in education and do everything to forbid students' access to AI. They believe that the results gained by students from ChatGPT, Gemini, and Copilot applications help them to reduce their own efforts and engagement and not to achieve desired competences intended by educational goals. Some essays have happened to be graded with the highest grades while teachers were unaware that it was the result of the work of AI, the consequence of which was severe punishment, including expulsion from higher-education institutions. Teachers think that students had no permission to use modern sources of information, and students believed that they would have all technology at their disposal after completing university studies, so they could not understand why higher-education institutions did not prepare them for using available technologies to optimally solve defined tasks. In our opinion, teachers in this particular case do not want to give up traditional teaching methods, in particular ways of evaluating students' knowledge. Teachers in the second group try to ignore the presence of AI, thinking that changes and results given by AI software will not have more significant effects in the near future. In their opinion, AI is under control, students will not use it excessively, so that at the moment there is no need for defining rules regarding its use. Education experts in the third group think proactively and are aware of rapid changes. They know that graduates will complete their projects and solve problems by using artificial intelligence whenever possible and no one will be against it in the companies employing them. Of course, it is necessary to have a critical relationship towards results given by AI in this stage of its development. It is also important to ensure that students' privacy is protected, and that data are stored safely and used in line with relevant regulations. Schools and educators must have clear policies for solving data privacy problems (Mandić, 2023). Artificial intelligence systems and robots must be programmed and ethically used in education. It is important to take into account the partialities and limitations of AI algorithms, as well as potential impact on students' welfare and mental health. Teachers should be aware of ethical implications of artificial intelligence and ensure that its use is in line with educational values and goals.

While artificial intelligence and robots can improve learning experience, possible inequalities may also arise in the approach to these technologies. Schools and districts must consider the questions of fairness and ensure that all students, regardless of their socio-economic background, have equal possibilities in benefiting from artificial intelligence and robotics in education. Integration of artificial intelligence and robots in the classroom requires training and supporting teachers adequately. They need to develop necessary skills and knowledge in order to efficiently use these technologies and to integrate the in their teaching practice. It is necessary to provide possibilities for permanent professional development so that teachers may use the full potential of artificial intelligence and robots in education. Teachers should ensure that technology is used as a means of improvement, and not a replacement of the teacher/student relationship. The application of artificial intelligence and robotics in education can lead to substantial costs, including the purchase of hardware, software and current maintenance. Schools and districts must carefully consider financial implications and sustainability of including these technologies in their school curricula. Schools need to consider the adaptability of their technological infrastructure and the ability to be in line with updates and progress in the field of artificial intelligence and robotics. This ensures that technology remains relevant and efficient in its support to learning outcomes, teachers' educational work, the development of students' critical thinking and creative potential. Integration of AI and robots in education can cause concern about the potential displacement of certain jobs in the educational sector. It is important to consider the social impact and implications of automation for workforce, as well as the need for teachers' retraining and specialization for the purpose of their adjustment to changeable roles and responsibilities. By addressing these risks and considerations, teachers can maximally use the advantages of artificial intelligence and robots in education, while simultaneously ensuring responsible and ethical use of these technologies. It is necessary to carefully plan and implement strategies which prioritize students' welfare, fairness and effective pedagogy in the context of integration of AI and robotics (Mandić, Miščević, Bujišić, 2024).

At the Faculty of Education, research entitled *Possibilities and limitations of modern educational technology from the point of view of teachers in Serbia* was conducted in cooperation with UNESCO. Teachers' self-evaluation is a process in which they form opinions about the adequacy and effectiveness of their own knowledge, performance, beliefs and effects with the aim of self-improvement. This was particularly important in the domain of being informed about the possibilities offered by AI-supported educational technology. It is important to ensure that students receive high-quality education which will prepare them for success in the contemporary world. We used the tool for online surveying called Survey Monkey. In the questionnaire, our respondents/teachers were asked questions about their evaluation of the achieved level of ICT competences, particularly of AI. Our respondents were teachers

at different levels of education, from pre-school, via primary and secondary school, to university. The Figure below gives us answers to one of the key questions.

The research results show that teachers are not sufficiently informed about the possibilities offered by AI nor are they trained for its use. To improve teachers' ICT competence, it is necessary to have a clear action plan and progress indications for the use of technologies in education. The Faculty of Education of the University of Belgrade is ready, in cooperation with the Ministry of Education and the *Institute for the Improvement of Education* to implement teacher training and to monitor their progress on a permanent basis.

Conclusion

In the past few years we have faced numerous changes in different social fields, including education. After the pandemic period during which teachers and students were forced to use new technologies and platforms for online learning, changes were initiated in education which intensified the development of digital technologies. From the relevant world research, it can be concluded that educational technology based on artificial intelligence frees teachers to a certain extent from routine jobs of teaching and grading while giving them more time for interaction with students, the development of critical thinking, value attitudes and creative potential. Teachers

How would you rate your satisfaction with the use of modern technologies based on artificial intelligence, e.g., ChatGPT, Bing or Gemini in the teaching process? Evaluate it on the scale from 1 to 4 (this and all future scales will always be oriented from negative towards positive as follows: 1 - I am not satisfied at all, 2 - I am partly not satisfied, 3- I am neither satisfied nor dissatisfied, 4 - I am partly satisfied, 5 - I am completely satisfied.

	1	2	3	4	5	TOTAL	WEIGHTED AVERAGE
☆	24.74%	7.48%	27.86%	25.36%	14.55%		
	119	36	134	122	70	481	2.98

The research conducted at the Faculty of Education

could have more possibilities to progress in professional terms, engage in creative work, research in the teaching process, solve educational problems, implement educational/upbringing programs, and be more socially involved in their communities. In this way, teachers have more time for planning their work, introducing necessary innovation and enrich work with their creativity. Integration of artificial intelligence and robots in education opens the world of possibilities and advantages both for students and teachers. These technologies have the potential for making revolutionary changes in the learning experience, making it more personalized, attractive and efficient. Platforms and robots based on artificial intelligence offer personalized learning paths, adjusting to students' individual needs and promoting independent learning. They offer practical experiences which encourage critical thinking, problem solving and cooperation skills. Artificial intelligence and robots promote inclusivity in education by ensuring accessible learning experiences to students with different needs. These technologies may be adjusted to support different learning styles, individualized teaching, as well as content adapted to students with special needs. Naturally, it is important to take into account risks and dangers related to the use of artificial intelligence and robots in education. Privacy, ethical use of artificial intelligence, fairness, teacher training and consideration of costs must be carefully taken into account in order to ensure its responsible and efficient application. By accepting artificial intelligence and robots in education we can create dynamic and innovative learning environments which prepare students for future challenges. These technologies have the power to encourage curiosity, motivate for learning and equip students with skills and knowledge needed for their progress in the rapidly-changing world. Modern teaching methods imply reduced time for teachers' lessons who, unlike traditional teachers, do not pass ready-made knowledge to students, but encourage them to research independently and find new information, and to progress in line with their own prior knowledge and abilities. Artificial intelligence might help us to make individualization and differentiation personalized for every student. In the developing teaching process, students learn on their own, using branched person-oriented textbooks which should replace outdated linear textbooks used in the same way for years. We believe that AI will contribute to better-quality teaching, and that teachers will have more time for educational work with students, for teaching them how to learn, how to develop logical thinking, for knowledge functionality, and in particular for addressing the development of children's character traits with the aim of creating best possible persons, which is the primary goal of society on the whole.

References

Baker, R.S., Inventado, P.S. (2014). Educational Data Mining and Learning Analytics. In: Larusson, J., White, B. (eds) *Learning Analytics* (61-75). New York: Springer.

D'Mello, S. & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, XXII (2), 145–157. https://doi.org/10.1016/j.learninstruc.2011.10.001

Mandić, D. (2023). Report on Smart Education in the Republic of Serbia. In: Zhuang, R., *et al.* (eds) *Smart Education in China and Central & Eastern European Countries. Lecture Notes in Educational Technology* (271-291). Singapore: Springer.

Mandić, D., Miščević, G., Bujišić, Lj. (2024). Evaluating the quality of responses generated by ChatGPT. *Metodička teorija i praksa*, XXVII (1), 5–19. DOI:10.5937/metpra27-51446.

Milutinović, V., Mandić, D. (2022). Predicting Acceptance to Use Computers at Traditional and Innovative Levels in Teaching Mathematics in Serbia. *Inovacije u nastavi*, XXXV (2), 71–88. https://doi.org/10.5937/inovacije2202071M [In Serbian]

Ristić, M., Mandić, D. (2018). Readiness of Education System for mobile learning. *Sociološki pregled*, LII (3), 1044-1071. DOI: 10.5937/socpreg52-18707

Željko B. Zdravković[1]
Academy of Applied Technical Studies,
Department Belgrade Polytechnics
Belgrade (Serbia)

UDC 7:004.8(497.11)
338.46:7.05(497.11)
Professional article
Received: 30.04.2024.
Accepted: 02.07.2024.
doi: 10.5937/napredak5-50771

Creative Serbia and AI: Potential of Shaping Creative Generative Artificial Intelligence in Arts of the Future through the Example of Midjourney/Discord Application

Abstract: The paper aims to investigate how artificial intelligence (AI) can shape the future of artistic creativity in Serbia. This analysis will focus on the potential of new algorithmic tools of creative generative AI, more specifically on applications such as Midjourney/Discord, as well as the specific examples and possibilities of artistic expression that they certainly open up. The integration of AI and art opens up new possibilities for creators and audiences, but at the same time raises questions about authenticity, originality and the role of the artist. Moreover, this paper will explore how AI technology affects the process of creating works of art, audience perception and cultural implications. Through an interdisciplinary approach that combines art, technology and social sciences, this paper will provide insight into the potential and challenges brought by the integration of artificial intelligence into art, especially in the context of creative Serbia.

Keywords: creative Serbia, generative artificial intelligence, art, future, Midjourney/Discord, algorithms, processes

Introduction

The need for creativity and innovation in Serbia is the key aspect which may affect the development and progress of society in different spheres, including economy, culture, education and the technological sector. The analysis of this matter requires considering Serbia's actual states of affairs, challenges and potential in the field of creativity and innovation. What should be considered

[1] zzdravkovic@atssb.edu.rs

first is the contextual framework. Serbia is facing different challenges, including economic inequalities, bureaucracy, still insufficiently developed educational systems and brain drain, which poses a serious threat to the future development of the country. In such an environment, creativity and innovation become key factors which may contribute to the resolution of these problems. Creativity can be seen as an ability of generating new ideas, while innovation refers to the application of these ideas in practice for the sake of creating new values. It is necessary to encourage the development of creative skills through permanent creative education and cultural programs.[2] In addition, it is important to create a thoroughly stimulating environment for innovation, which includes support to research and development, facilitation of the business environment, as well as constant strengthening of ties between the academic community, the private sector and the state.

Some of the key sectors in which Serbia can achieve significant progress through creativity and innovation are the technological sector and Industry 4.0[3]. The development of high-tech industries, such as information technologies, biotechnology, artificial intelligence and digital media, may create new workplaces, attract foreign investments and improve the country's competitiveness in the international market. However, to achieve this potential, it is necessary to invest in education, infrastructure and research, as well as in the already-mentioned creation of the stimulating environment for innovative startups. Moreover, creativity and innovation may also contribute to the accelerated development of cultural and creative industry in Serbia, particularly in the field of applied arts, cinematography and music art. Encouraging artistic creativity, supporting cultural events and festivals, as well as promoting cultural tourism, can contribute to the country's economic development and the strengthening of its international reputation. By encouraging creative industries, technological innovation and cultural creation, Serbia may achieve significant progress in economic, cultural and social terms, creating sustainable future for all its citizens.

Serbia has an extremely rich artistic tradition and cultural heritage, but the need also arises for new forms of expression and innovation in art. Introducing creative generative AI[4] in this field can contribute to the creation of new, innovative and even yet unseen works of art which enable artists and designers to explore new forms and fields in art and to express themselves in completely new ways. Digital expression with generated artistic visuals strives towards a new chapter and a new era of creative formation.

[2] See: https://serbiacreates.rs/ (Accessed on 15. 4. 2024.)

[3] See: https://www.weforum.org/about/the-fourth-industrial-revolution-by-klaus-schwab/ (Accessed on 14. 4. 2024.)

[4] AI – artificial intelligence

Artificial intelligence in art

The emergence of artificial intelligence, as it is already known, opens up completely new horizons in art. Generative AI technologies enable computers to create original images, music, texts and other forms of creative expression (Figoli, Mattioli, Rampino, 2022, p. 45). Artists, designers and architects can use AI as a tool for experimenting, installation or even as a creative digital partner.[5] This technology opens up possibilities for creating works of art which go beyond human imagination and bring new aesthetic and conceptual expressions in Serbian art. Comparative analysis of artificial intelligence in art can cover various aspects, including the application of AI in different branches of art, the impact on the process of creating works of art, audience perception and cultural implications. In painting, AI algorithms can generate new pictures based on the style features of eminent artists or even create completely new aesthetics of expression. In music, AI can compose, generate sounds or even emulate the voices of famous performers, while in film it is used for generating special effects, animation, post-production or even creating scenarios based on the analysis of the existing films.

In the process of creating new works of art, AI helps immensely in the experiment. Designers can use AI as a tool for exploring new ideas, conceptual artists can use it for experimenting with different techniques, and it can particularly accelerate certain stages in the process of the creation itself of works of art, particularly in the sector of applied art (Figoli, Mattioli, Rampino, 2022, p. 35).

Artificial intelligence as a new medium definitely brings about a revolution in art, thus opening up new horizons for creators and audience, at the same time posing challenges to the relation and role of the artist in cultural implications of technological change. It is crucial to continue research, dialogues and public debates on this topic so that we, as Serbian society, can understand both the potential and the limitations of this interaction related to artificial intelligence, and ensure that this technology serves general cultural progress.

Above all, it is always interesting to consider the potential of art generated by artificial intelligence through the works of skilful and creative individuals worldwide, for example Professor Joshua Vermillion[6] from UNLV in Nevada. Professor Vermillion is only one in a multitude of those who have begun experimenting with generative AI. Since these technologies continue to develop and become increasingly sophisticated, we can expect even more intensive integration of AI into the world of art. It is inevitable in the future of the majority of creative industries.

[5] See: The Intersection of Art and AI | Ai-Da Robot | TEDxOxford: https://www.youtube.com/watch?v=XaZJG7jiRak (Accessed on 16. 4. 2024.)

[6] See: https://vermillion.faculty.unlv.edu/ (Accessed on 17. 4. 2024.)

Combining art and artificial intelligence: artistic spatial installations by Professor Joshua Vermillion, generated on the Midjourney/discord platform

Source: artistcloseup.com

Creative generative AI and its potential: Midjourney/Discord application as a platform for artistic interaction

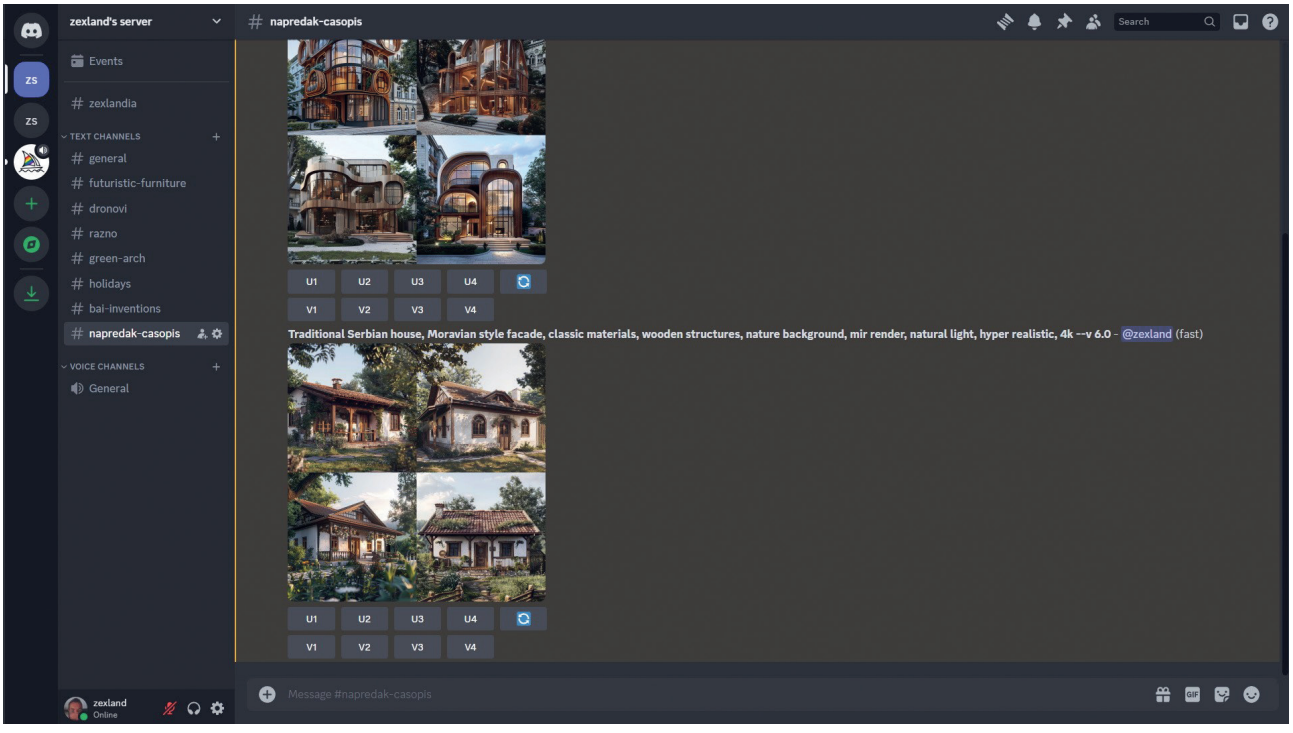
We have defined creative generative AI as a field dealing with the creation of new types of visualization, i.e., works of art by using algorithmic processes and computer learning (Muntañola, 2022, p. 16). However, this technology, apart from being able to analyze the existing works of art and generate new ideas, can also make possible the cooperation between artists and designers in the creation of completely new interdisciplinary works of art. The application of creative generative AI in Serbia can provide artists and designers with completely new possibilities for expressing and improving creative processes.

One of the currently most popular online applications is the Midjourney/Discord[7] platform used

[7] See: https://www.quora.com/How-does-Midjourney-Discord-work (Accessed on 15. 4. 2024.)

for generating different virtual objects. To generate any visualization, we must first open a Discord account and subscribe to the Midjourney server. When we join this server, we can begin generating images, for which the term "prompting" is used, by entering certain textual commands, but with the compulsory first prefix: /imagine <prompt>, where <prompt> is the first textual description of the image we want to generate. For example, to generate an image of a cat sitting on the beach, we would type the following command: */imagine <prompt> /imagine a cat sitting on the beach.* The so-called Midjourney bot will then generate this image based on our query and send it back to the Discord account.

Visualizations generated on the Midjourney/Discord platform are publicly visible in the gallery of the members, but if we have "pro or mega" subscription, we can generate private images visible only to us. We need to be as precise as possible in our queries. The more specific we are in the queries, the better the results will be. We can also use several queries in one command. This may help us to generate more creative and interesting results.



Entering a textual command (prompting)

Source: Midjourney/Discord

The result: an example of the traditional Serbian house built in Morava style, as seen by AI
Source: Midjourney/Discord

Midjourney and Discord as a platform may with time grow into a place for networking Serbian artists, designers and architects, as well as become a creative online hub for discussing art topics and collaborating on projects, either within the academic community or economy. The absolute benefit is that artificial intelligence in this case can definitely help in finding and connecting similar talents from different artistic, as well as technical-technological fields based on similar interests and professional goals in the future.

Arts of the future

Gilles Lipovetsky, French philosopher, sociologist and theoretician of culture, is well-known for his papers dealing with the phenomena of modern society, individualism, post-modernism and changes in cultural patterns. On the other hand, the topic of the potential of shaping creative generative artificial intelligence (AI) in arts of the future opens up a wide range of discussions about how this technology can transform the artistic

process and expression. Comparing these topics gives us insight into the ways of interconnecting technological progress and cultural changes. In his extraordinary book *L'ère du vide*, Lipovetsky points in a visionary way to the future tendencies of "digital arts" and states that "the use of the dispositive for automatic processing of information in the beginning attracted only experts such as Michael Noll or Manfred Mohr,[8] capable of mastering not only this technique, but also the combinatorics of mathematical and logical type". Furthermore, it is emphasized that "the works created by them, inspired by abstract geometry rather than interest in figuration, play with graphic elements, series and algorithmizing of colour" (Lipovetsky, 2011, p. 317). In that context, art plays a key role as a form of expressing and articulating individual experiences, identities and broadening the boundaries of traditional artistic practices in the forthcoming times.

Lipovetsky also analyses the impact of technological changes on culture and society, including digitization and globalization. However, the question arises as to how these two aspects can be in mutual cohesion. One of the potential trends is that AI will enable artists to explore new dimensions of individuality and identity through interaction with the technology. For example, AI can analyze patterns of human behaviour and make personalized works of art which reflect unique aspects of individual experiences. Are they synthetized simulacra of the already known works of art? In addition, AI can serve as a toll for experimenting with completely new aesthetic trends and forms of expression, which contributes to the diversity and innovation, as well as to simulation in artistic creation and, thus, new and still unknown arts of the future. On the other hand, there is also a potential challenge regarding the use of AI in art, and that is the question of authenticity and originality. Since AI technology is becoming more and more advanced in the creation of works of art, the question arises as to whether those works can really be considered authentic or original, having in mind that they have been generated by algorithms, i.e., simulated from billions of pixels invisible to the naked eye from the *world wide web*. This question can provoke discussions about the topic of simulation of art and the role of the artist in the creation of a new algorithmic work of art.

Simulation and simulacra in the context of producing the surreal is perfectly described by Jean Baudrillard) when speaking about a three-dimensional simulacrum: "But why would a three-dimensional simulacrum be closer to the real one than the two-dimensional one? It would like to be like that, but its paradoxical effect is, on the contrary, to make us more sensitive to the fourth dimension as a hidden truth, a secret dimension of each thing, which suddenly acquires the strength of obviousness" (Baudrillard, 1979, p. 109).

[8] See: The Art of Generative Thinking | Manfred Mohr in Conversation with Margit Rosen. http://www.emohr.com/ (Accessed on 18. 4. 2024.)

An example of exploring the form of an amorphous modern façade in the colours of the Serbian flag, as seen by AI
Source: Midjourney/Discord

It is evident that the hidden creativity which is invisibly intertwined among trillions of pixels, reproduced in an unbelievably short time, also has its other hidden truth which we still cannot see, at least not legally, within regulating rights and obligations of the future "AI prompt designers"[9] in the course of this form of the generative AI creative process. Just as in any new field, we believe that this matter will be regulated in the near future, both in the world and in our country, and that it will not remain lost in the fourth dimension.

Above all, the fourth dimension concept in art is often related to the idea of time, space and motion. Throughout history artists have tried to express these ideas through different media, such as paint-

[9] See: https://www.internimagazine.com/features/che-cose-un-prompt-designer-e-come-diventarlo/ (Accessed on 20. 4. 2024.)

ing, sculpture, performance or architecture. The examples of the works of art exploring the fourth dimension include the works such as *Nude Descending a Staircase (No. 2)*[10] by Marcel Duchamp (1887–1968) from 1912, or *Portrait of Daniel-Henry Kahnweiler*[11] by Pablo Picasso (1881–1973) from 1910, which shows movement and motion in several dimensions of the Cubist expression.

At the very end, we made a small generated AI experiment on the Midjourney/Discord platform, where we put Picasso's *Portrait of Daniel-Henry Kahnweiler* in the application as an inspiration for the modern sofa design.

By analyzing different trends, technological innovation and styles, we have observed key characteristics and potential brought by new arts in the future. They constitute a fascinating, dynamic and infinitely unexplored space of creativity that is constantly shaped under the effect of technological, cultural and social changes in which we live. Being interdisciplinary, under the effect of AI as a new medium, they will give us new findings, create new artistic skills and new artists for the 22nd century. Finally, the arts of the future promise an exciting and inspirational future which invites us to explore, dream and create together.

The result of the experiment of generating a work by Picasso
Source: Midjourney/Discord

[10] See: https://philamuseum.org/collection/object/51449 (Accessed on 21. 4. 2024.)

[11] See: https://www.artic.edu/artworks/111060/daniel-henry-kahnweiler (Accessed on 22. 4. 2024.)

Conclusion

Through the analysis of the potential of shaping creative generative artificial intelligence in arts of the future using the example of Midjourney/Discord application, we gain insight into various possibilities provided by this technology, but also into challenges we encounter in the process. As we have already stated, art in Serbia has deep roots and rich heritage which goes back to different periods and cultural influences. The integration of AI into art can only lead to even greater enrichment of the already acquired Serbian heritage by creating new forms of expression, experimenting with new aesthetic approaches and interactive experiences for the audience. The Midjourney/Discord application is an example of how AI can be used for generating extremely creative content which inspires, provokes and engages artists, designers, architects and future users.

We conclude that the potential of shaping creative generative artificial intelligence in the arts of the future through the example of the Midjourney/Discord application provide insight into the dynamic combination of technology, local culture and creativity. Despite challenges, this interaction brings possibilities for innovation, changes and sustainable development of creative industries in Serbia, thus producing a rich and inspirational cultural scene for our future generations.

References

Baudrillard, J. (1979). *Simulacra and Simulation*. Beograd: Svetovi. [In Serbian]

Lipovetsky, G. (2011). *L'ère du vide: Essais sur l'individualisme contemporain*. Novi Sad: Izdavačka knjižarnica Zorana Stojanovića. [In Serbian]

Muntañola, J. (2022). *Arquitectonics Mind, Land & Society. Artificial intelligence and architectural design: an introduction*. Barcelona: Universitat Politècnica de Catalunya

Figoli, F., Mattioli, F., Rampino L. (2022). *Artificial Intelligence in the Design Process. The Impact on Creativity and Team Collaboration.* Milano: FrancoAngeli.

Internet sources

https://serbiacreates.rs/

https://www.weforum.org/about/the-fourth-industrial-revolution-by-klaus-schwab/

https://www.youtube.com/watch?v=XaZJG7jiRak

https://vermillion.faculty.unlv.edu/

https://www.quora.com/How-does-Midjourney-Discord-work

http://www.emohr.com/

https://www.internimagazine.com/features/che-cose-un-prompt-designer-e-come-diventarlo/

https://philamuseum.org/collection/object/51449

https://www.artic.edu/artworks/111060/daniel-henry-kahnweiler

Miloš Z. Agatonović[1]
Academy of Applied Preschool Teaching and Health Studies
Kruševac (Serbia)

UDC 121:004.896
Professional article
Received: 23.04.2024.
Accepted: 02.07.2024.
doi: 10.5937/napredak5-50621

Is True Robot Autonomy Possible? (A Philosophical Consideration)

Abstract: The paper aims to demonstrate that autonomy in action requires self-location so that it is possible to construct autonomous systems only if they are capable of self-localization. The paper briefly presents the history of the idea of humanoid robots, and then in the main part it conceptually discusses the question of the possibility of robot autonomy starting from the problem of indexicals. The paper starts from John Perry's thesis that indexicals are necessary for action, and subsequently, presents Jenann Ismael's concept of an agent representing an information system in which indexical information, for example in self-location, connects an information model with the environment. The paper points to examples of robots, such as the Figure 01 model, as systems described by Ismael, which are capable of autonomously performing actions as well as of self-locating.

Keywords: robots, autonomous action, indexicals, self-location, Figure 01 robot model

Introduction

At the beginning of 2024, some of the world's biggest technological companies announced the projects in which they intended to produce humanoid robots that would be autonomous in the true sense of the word. With its chatbot based on artificial intelligence (AI) and large language models (LLM), Company OpenAI, known for its popular Internet platform ChatGPT, began a cooperation with a smaller robotic company Figure to create an autonomous robot in a human form which is capable of performing various actions and tasks on its own.[2] Thanks to Figure and Open AI joining forces and to merging robotics and AI, the Figure 01 robot model is able to speak, draw practical conclusions, perform tasks in line with the orders formulated in a natural language,

[1] milos.agatonovic@vaspks.edu.rs ; https://orcid.org/0000-0001-6598-1792

[2] See: https://www.figure.ai/

to notice and recognize objects and acts accordingly.[3] Big technological companies, such as Nvidia, Amazon, Microsoft and OpenAI, recognized the importance of the Figure project and invested $675 million, thus substantially increasing its value which, according to the current estimates, is $2.6 billion.[4]

The idea of humanoid robots and AI, which attracts large investments and encourages economic development in the modern world, has been the subject of considerations by thinkers, philosophers and scientists ever since ancient times. In this paper we will briefly present the history of the idea of robots and then consider the question of the potential autonomy of robots, starting from the problems of indexicals and self-location. The goal of this paper is to show that for autonomy of action, either of men or robots, self-location is necessary, so that it is possible to construct systems with autonomy only if those systems are capable of self-location.

Brief history of the idea of robots

In ancient times, the idea of humanoid machines was originally present in mythological and religious narratives. Heron of Alexandria, 1st century AD), an ancient Greek mathematician and inventor, was the first to devise practical automata outside mythology, and that is why he is considered the founder of cybernetics. His famous automata were applied in the theatre and music. In the Middle Ages, Ibn ar-Razāz al-Jazarī (1136–1206) also construed practical automata, described in *The Book of Knowledge of Ingenious Mechanical Devices* (*Kitab fi ma 'rifat al-hiyal al-handasiya*). From the beginning of the modern era, thinkers such as Leonardo da Vinci, René Descartes and Nikola Tesla were obsessed with the idea of humanoid robots (Rosheim, 1994, p. 1). Mark Rosheim introduced the term "anthrobot" for anthropomorphic robots (in old Greek, *antropos* – man, *morphe* – shape) (Rosheim, 1990, p. 2162) to determine the ideas of humanoid robots by the above-mentioned thinkers (Rosheim, 1994, p. 1).

The research aimed at construing the humanoid robotic arm began in the 1960s, with the pioneering work of the members of "Mihajlo Pupin" Institute, professors from the Faculty of Electrical Engineering in Belgrade (Tomović, 1960; Tomović & Boni, 1962; Rakić, 1964). Namely, in 1963 they construed the first bionic hand, or adaptive prosthesis with external power supply, known as the "Belgrade Hand".

[3] See: "Figure Status Update – OpenAI Speech-to-Speech Reasoning", available at https://youtu.be/Sq1QZB5baNw?si=g_nhbZSNRWMlNVpV (Accessed on 11. 4. 2024)

[4] See: https://www.cnbc.com/2024/02/29/robot-startup-figure-valued-at-2point6-billion-by-bezos-amazon-nvidia.html (Accessed on 11. 4. 2024)

Rosheim himself construed a robotic arm which was able, in terms of mobility, to perform every task just as the real human hand (Rosheim, 1990, p. 2162). In terms of kinematics and structure, Rosheim's "omnidirectional arm" is equivalent to the human arm and, therefore, it is "anthrobotic" (Rosheim, 1990, p. 2162).

Today's robotic arm models may perform different movements and strokes in different positions, with great precision and skill. Moreover, apart from the Figure 01 robot model being able to use its arms to sort out and put dishes into a basket, to collect scattered litter and put it into a bin, to catch and give an apple, as it can be seen in the company's promotional video,[5] it can also *recognize* objects, dishes, a basket for dishes, litter, a bin, an apple, *draw conclusions* about all objects, *treat* them *practically* in an adequate manner,

Belgrade hand was constructed in 1963 by the researchers from Institute "Mihajlo Pupin", part of the collection in the Museum of Science and Technology in Belgrade.
Photo: Museum of Science and Technology

[5] See: "Figure Status Update – OpenAI Speech-to-Speech Reasoning", available at https://youtu.be/Sq1QZB5baNw?si=g_nhbZSNRWMlNVpV (Accessed on 11. 4. 2024)

determine the location of an object in relation to itself and its own location in relation to objects, act *upon general orders*, talk and *understand* the interlocutor's intention and *assess* its actions. This robot has the above-listed abilities not only thanks to the construction of robotic hands and sensors, but primarily to the neuron network models which have been trained on a sufficiently large amount of data to resolve problems autonomously. By joining robotics and machine learning technology and neural networks on which artificial intelligence relies, the idea strived for by thinkers ever since ancient times will be realized – that machines can be autonomous in their action like a human, that they can autonomously resolve problems, make decisions and perform practical tasks.

From indexicals to autonomy

According to an influential belief in modern philosophy, an individual's ability to translate his/her understanding of the world into action depends on "indexical thoughts (*indexicals*)", or thoughts in which something or someone is indicated, shown or pointed to. One of the proponents of this belief is John Perry, an American philosopher who claims that indexical expressions, such as the personal pronoun "I" and demonstrative pronouns ("this", "that", "here", "there"), are indispensable and irreplaceable in indexicals. If we replaced indexical expressions in indexicals by descriptive terms, we would get an inadequate explanation of behaviour.

Perry provides graphic examples in which he points out the essential difference between the descriptions of the content of beliefs and beliefs themselves in which indexical expressions cannot be omitted. Based on the thought experiment about the "messy shopper" (Perry, 1979), since the shopper John Perry noticed sugar spilled on the shop floor, he believed that it had been spilled by a messy shopper. However, after realizing that the sugar bag in his basket was perforated, he concluded that he was responsible for the mess. Believing that someone spills sugar and believing that John Perry spills sugar essentially differ from his belief expressed by the sentence: "I spill sugar". After realizing that it was he who spilled sugar in the shop, and changing his belief, Perry also changed his behaviour and stopped spilling sugar. The change of belief into the belief expressed by the statement: "I spill sugar" explains the change of behaviour. Without this indexical belief, which points his/her place to the agent, his/her role and relationship towards the described situation, there would be no change of behaviour. The component expressed by the indexical expression in the indexical belief cannot be omitted since with its aid the agent determines his/her place, his/her role and relationship towards the situation, thus motivating his/her acting.

Starting from Perry's understanding that the agent's ability to translate his understanding of the world into actions depends on indexicals, Jenann Ismael considers problems related to an adequate answer and attitude towards circumstances when managing the system, the circumstances which are determined by different contextually dependent contingent facts, whereas solutions to these problems are related to the corresponding context (Ismael, 2007, p. 18). Ismael tries to integrate the self into the natural order with the aim of explaining the computing model of the self as a component of the embodied system which has been construed to solve primarily practical problems (Ismael, 2007, p. 18). She points out that the

role or the function of the structure which represents the world in the head is not to give a reflection of the nature, but to establish a connection between the elements of the internal structure and the outer space in order to direct the movement of the body in the right direction (Ismael, 2007, p. 18). For the agent to be able to understand his/her environment, it is enough to provide him/her with the source of information whose states implicitly have an *"indexical spatial content"* (Ismael, 2007, p. 18). The content of the states changes dynamically with the agent's location, enabling his/her behaviour to be aligned with the current position. The agent receives only the information directly relevant for his/her actions.

Ismael introduces a distinction between semantic and architectural relations: semantic relations refer to the relations between the elements within the "representational medium", while architectural relations connect the elements within "the representational medium" with their subject or with the elements of our media which are not part of the same semantic network (Ismael, 2007, p. 33). Within this concept, central notions are representational media and coordination, whereas representational media perform the function of information channels, while medium coordination in relation to its subject includes the establishment and maintenance of causal and contextual relations which ensure that its states correspond to information in different contexts they are used in (Ismael, 2007, p. 6). Man's consciousness is such a medium, with a role of creating models of the self and situation, connecting them with self-locating actions. These models may be composed of different elements, while self-location establishes architectural relations between the model and world components (Ismael, 2007, p. 34).

Paul Teller attempts to explain Jenann Ismael's ideas on the specific examples of artificial systems which may act autonomously, whereas such system is integrated in the "natural order" (Teller, 2011, p. 768). Teller describes an example of a "representational medium" in which the function performed is analogous to indexical expressions and indexical beliefs. Namely, Teller describes a group of robots which autonomously perform actions following the instructions relying only on the ideas of a third person (Teller, 2011, p. 767). Teller's robots are equipped with the following devices: *identification register,* which provides robots with relevant information; *deliberation modules*, which assess whether it is necessary to do the prescribed task; and *perception modules*, which enable recognizing and reporting about a certain event. These devices make it possible for Teller's robots to establish semantic relations, which include the information collected by the robot and interpreted with the aid of the above-list3ed devices, and architectural relations, which include the relations of the robot and its devices with the environment. Similarly to Perry's scenario of the "messy shopper", in the situation in which the robot with the perforated bag spills sugar, the event is recorded in the information register in the third person, with the accurate identification of the robot spilling sugar. The recording of this event in the register causes the initiation of the sub-module of the action module, the "sugar bag patching module", with the robot whose sugar is spilled (Teller, 2011, p. 767). According to it, Teller concludes that every robot acts appropriately on the basis of the third-person reports about the robot with a certain identification spills sugar, so that the real robot should perform the task of patching the sugar bag

Humanoid robot Optimus was produced by American multinational company Tesla; its prototype was first presented in 2022. In May 2024, Tesla Company presented a video in which the robot performs a number of different tasks in their factory.
Photo: Shutterstock

while using no indexical information (Teller, 2011, p. 767). However, Teller draws an exaggerated conclusion that robots do not use indexical information, since the entire setting has been devised to enable robots to interpret reports in the third person from the information register so that they self-locate.

Teller's robots are autonomous since they perform actions on their own. However, without the setting that enables their self-location, acting would not be possible. The Figure 01 robot model is autonomous even without the external setting in the form of the register network and modules, primarily thanks to large language models which enable not only the understanding of the language, either speech or language recognition of objects, but also the activation of action based on language components which are a "trigger" for action in an adequate context. The "trigger" may refer to orders, normative statements, questions, but also indexical statements. Figure 01 takes an apple from the table in front of it and gives it to the interlocutor after being asked for some food.[6] In all actions and justifications of those action, Figure 01 relates to objects, actions and the interlocutor, determining its relationship with the aid of indexical expressions, explicitly or implicitly. Moreover, Figure 01 assesses its actions by a statement which starts with a personal pronoun "I", justifying why it has completed the tasks successfully. This definitely does not imply that the robot has consciousness, but it points to the fact that it self-locates while acting. Without self-location, action would be impossible: if Figure 01 does not know that it is next to the apple and the interlocutor who asked for some food, it would not give this apple to the interlocutor.

Conclusion

That self-location is necessary for robots' autonomous action is indicated by the importance of the problem of simultaneous localization and mapping (SLAM) in robotics. This problem raises the question whether it is possible for a mobile robot on an unfamiliar location in an unfamiliar environment to make a map of the environment by researching it and simultaneously determining its own location. This method of the movement of robots and vehicles differs from the movement which is externally navigated, for example with the aid of the Global Positioning System (GPS). SLAM enables robots and vehicles to move autonomously in an unfamiliar environment in which external navigation is not possible (Durrant-Whyte & Bailey, 2006).

Although the autonomy topic still sounds like science fiction even today, autonomous robots have become reality. Furthermore, as intended to show in this paper, autonomy of robots is not an unreasonable idea, but the one that is possible to elaborate in detail and realize by constructing information systems which may act autonomously, as simple examples of Teller's robots or the sophisticated Figure 01 model analyzed here. In the future, it may be expected that autonomous robots will become part of our everyday life and even to assume most human jobs. Economist Guy Standing claims that the forthcoming technological revolution, which brings along the application of robots and AI, will disturb everyday life and make it insecure, and that is why it is necessary to introduce a system of wealth distribution in the form of

[6] See: https://youtu.be/Sq1QZB5baNw?si=g_nhbZSNRWMlNVpV (Accessed on 11. 4. 2024)

a universal basic income which would ensure everyone's right to the part of the economic benefit brought by automation and AI (Standing, 2020, pp. 31-32). Philosopher Nick Bostrom (Bostrom, 2024) speaks of post-work utopias as visions of society have accomplished full automation and thanks to it, the need for work has been eliminated. According to Bostrom, from the perspective of technological maturity, post-work utopia is a realistic vision, while from the same perspective it would be unrealistic to assume that human work is necessary. There is definitely a larger number of ominous visions of the future regarding the application of autonomous robots, which are commonly known and do not need mentioning. In any case, it may be expected that robots and AI will change every society in the world, and that is why it is necessary to consider seriously their application and society's organization about their application.

References

Bostrom, N. (2024). *Deep Utopia: Life and Meaning in a Solved World*. Washington, DC: IdeaPress Publishing.

Durrant-Whyte, H., Bailey, T. (2006). Simultaneous Localization and Mapping: Part I. *IEEE Robotics & Automation Magazine*, XIII (2), 99–110. doi: 10.1109/MRA.2006.1638022

Ibn ar-Razāz al-Jazarī. (1974). *The Book of Knowledge of Ingenious Mechanical Devices: (Kitab fi ma 'rifat al-hiyal al-handasiya)* (D. R. Hill, Trans.). Dordrecht/Boston: D. Reidel.

Ismael, J. T. (2007). *The Situated Self.* Oxford: Oxford University Press.

Perry, J. (1979). The Problem of the Essential Indexical. *Noûs*, XIII (1), 3–21. Available at: https://www.uvm.edu/~lderosse/courses/lang/Perry(1979).pdf

Rakić, M. (1964). An Automatic Hand Prosthesis. *Medical Electronics and Biological Engineering*, II, 47–55. doi: 10.1007/BF02474360

Rosheim, M. E. (1990). Design of an omnidirectional arm. *Proceedings of the IEEE International Conference on Robotics and Automation*, III, 2162–2167. doi: 10.1109/ROBOT.1990.126324.

Rosheim, M. E. (1994). *Robot Evolution: The Development of Anthrobotics*. New York: John Wiley and Sons.

Standing, G. (2020). *Battling Eight Giants: Basic Income Now*. London & New York: I. B. Tauris.

Teller, P. (2011). Robots, Action, and the „Essential Indexical". *Phenomenological Research*, LXXXII (3), 763–771.

Tomović, R. (1960). Human Hand as a Feedback System. *First IFAC Congress Moscow,* I (1), 624–628. Available at: https://www.sciencedirect.com/science/article/pii/S1474667017701425?__cf_chl_tk=t9bi6xCM.BKCUKHeaDRyYVHnHPAMIeajZ-VbNsDDNp9w-1721823180-0.0.1.1-3967

Tomović, R., Boni, G. (1962). An adaptive artificial hand. *IRE Transactions on Automatic Control*, VII (3), 3–10. doi: 10.1109/TAC.1962.1105456

Internet sources

https://www.cnbc.com/2024/02/29/robot-startup-figure-valued-at-2point6-billion-by-bezos-amazon-nvidia.html

https://www.figure.ai/

https://youtu.be/Sq1QZB5baNw?si=g_nhbZSNRWMlNVpV

Slobodan Đ. Prodić[1]
Faculty of European Legal and Political Studies
Belgrade (Serbia)

UDC 159.942.5:2
004.896
Review scientific article
Received: 12.06.2024.
Accepted: 09.07.2024.
doi: 10.5937/napredak5-51599

Fear of artificial intelligence initiated by religious feelings

Abstract: Artificial intelligence is a specific phenomenon which largely affects modern society. Being a relatively recent phenomenon, inside communities undergoing the desecularization process, artificial intelligence inevitably leads to the question as to how and in what way it has been accepted by believing people. Believers' opinions about artificial intelligence are divided. In a number of them it causes fear and extremely negative feelings. The most frequent reason for it is the insufficient familiarity about what artificial intelligence is and what its place in modern society is. As for religious communities, huge responsibility about the attitude towards artificial intelligence is actually assumed by clergy and their willingness to speak about it to people they have been entrusted with for spiritual guidance.

Keywords: artificial intelligence, religion, fear

1. Introductory notes

Artificial intelligence (AI) is undoubtedly a field about which is rightfully spoken about frequently nowadays. The reason for it is the fact that AI is becoming ever more present on a daily basis in different areas of life and that, as such, it is something that attracts attention of almost entire society. AI is reality and therefore it will remain part of society. What AI has largely found its position in are mainly the areas of different branches of industry, healthcare etc.

On the other hand, AI is evidently finding its place in other segments which are, so as to say, *closer* to ordinary people. This fact has substantially contributed to the fact that it is exactly in the broadest layer of society, or among people without a narrow specialization, and thus without being sufficiently familiar with what AI is and what areas it is applied in, a feeling of fear arises that is characteristic of

[1] porodicaprodic@gmail.com ; https://orcid.org/0009-0007-2297-7460

encountering the unknown (Luknar, 2023, pp. 151-152). In this brief paper, our intention is to focus on only one of the forms of fear observed among people who encounter AI for the first time or who have only listened or read about it – the fear initiated by religious feelings. In a society like ours, i.e., in a society that is, statistically speaking, very religious and in which, according to the data from several years ago, 72% people declare themselves as believers, it is necessary to pay attention to a detail such as fear of AI initiated by religious feelings and to take certain steps in order to control that fear and prevent it from making an individual feel uncomfortable.[2]

2. Phenomenon of fear in Christianity

A substantial number of historians of religion advocate the attitude according to which religion, in different forms and systems, is inseparably connected with man ever since ancient times. For a believer, it exists from the very moment of his creation by God, while for the non-believing ones religion is the man's answer to the phenomena he has not been able to control and which he believes to come from a higher force (Prodić, 2024, p. 7). Without going deeper into the resolution of the question which of these two attitudes towards religion and its existence in human society is older and, so as to say, scientifically more founded, what is distinguished as an element present in both attitudes is exactly the existence of fear of what man considers supernatural. This fear is perhaps best described in Christianity, i.e., in the terminology of the Orthodox Church, where we encounter a very interesting term – *fear of God.*

A believing man accepts the dogma about the entire world, and thus a man, actually being a God's creation. As such, it was given a clear and unambiguous task from the Creator: "Be fruitful and multiply! Fill the earth and subdue it! Rule over the fish of the sea and the birds of the air and every creature that moves on the ground" (Book of Genesis, 1:28). To put it more simply, God as the creator of everything, according to Christian religion, has placed a man in the position of a responsible master of the entire creation. Moreover, what is often "forgotten" by a modern man is that the Creator has also given the task to a man to be his own master, i.e., the master of his thoughts, words and acts. Everyone is more or less familiar with the Biblical story about the Fall and its consequences. This act of our progenitors constituted an essentially important relationship of a man towards God and God's relationship towards a man. A man, or Adam and Eve, disobeyed the Creator's commandment and that is why they lost a direct relationship with Him they used to have before the Fall. On the other hand, God as the man's Creator, does not leave the man without the possibility to re-establish the relationship which used to existed before the Fall because, as it will be announced by the promised arrival of the Saviour of the world, Jesus Christ, His resurrection and the establishment of the eucharist and the church, each of us,

[2] http://www.021.rs/story/Info/Srbija/105636/Koliko-ljudi-u-Srbiji-je-religiozno (accessed on 12.05.2024.)

as the descendants of the progenitors, is given the possibility to realize the union with God in the Kingdom of Heaven.

In shortest terms, these are the axioms of Christian soteriology and eschatological dimensions of the Orthodox faith. What is omnipresent in a very specific manner as a segment of the salvation process is, among other things, the necessary existence of fear in a man. This fear manifests itself in several ways and in relation to the topic we speak about it is the fear focused on the fact that a man as an individual striving to get salvation and be once again in union with the Creator, should not be recognized by the Creator as His at an essentially important moment. The moment we refer to is the time when the man's soul, after its separation from the body, comes to God hoping that He, as the Judge, will recognize that person or not as "someone who is his", someone who with the faith in God made efforts to spend life on this earth in doing everything that is expected from a believing man in the salvation process.[3] It is in this salvation process, which actually lasts throughout man's entire life, that the fear from not being with God is the essence of what we determine by the term *fear of God.* It is exactly from this fear, among other things, that the man's need for correct conduct (orthopraxy) is founded, which implies the fulfilment of different God's commandments both from the times of the Old Testament, and particularly from the times of the New Testament of God with a man.

3. Christian morality as a stabilizer in the face of challenges and temptations

All of us have encountered the term *morality* on countless occasions to date. This term basically, from the perception of Christian faith, implies a set of regulations which quite specifically exist as a certain type of stabilizer which has the task to keep the man within, to put it graphically, the boundaries of the territories in which he can be accomplished as a God-created being, as a crown of God's creation. Morality is a necessity in the man's relationship with God, with the nearest ones (other people) and with himself. With these important determinants, of course, if we look at morality from the perception of a believing Christian, it becomes clear why there is a specific connection between ethical/moral norms and the above-mentioned fear of God. Today's testimony of it is still dominant Christian foundation of European civilization, based exactly on Biblical values. Despite an ever more pronounced attack on these values, European society still has the feeling of self-accountability, as swell as of thoughtfulness in the challenges brought along by certain situations or, to put in in terms of theology, temptations. For the sake of illustration, we will emphasize the following: society in the above-mentioned region more or less still reacts to the injustice incurred against the other/others, primarily in everyday situations. Observed empathy largely has its foundations in Christianity, i.e., in the individual's consciousness about the

[3] A good example would be the New Testament parable about foolish and wise virgins (Matthew 25: 1–13).

existence of unambiguous commandment about Christian virtues, whose realization in everyday life constitutes a specific bastion in the face of the temptation of selfishness or any other form of *sklerokardia*. In fact, "hard-heartedness", which is the literal translation of this Greek word, is really an illustrative example of the condition of a large number of population today, which, unfortunately, seems to become dominant in many parts of our continent.

On the other hand, a question arises as to the relation between Christian morality and a novelty such as AI. First of all, we should wonder whether AI is one of the challenges or, more precisely, one of the countless temptations encountered by an individual, as well as the society in which he lives. Seen from the perspective of theology, AI is really both a challenge and a temptation. If we see it as a challenge, it becomes something that can be classified as one of the number of phenomena which *attract* the man (men and women equally) towards itself. AI is definitely "attractive for the eye" and "desirable for making one wise" (Book of Genesis 3:6). As such, it may easily turn into what is implied under the notion of *temptation*. In fact, there are countless things and phenomena which were used by a man for reasons known only to himself and which are still used in a manner that is not good, the manner that is equally directed to causing harm to others, but also to himself. To put it more simply, there are countless things which have been *abused* by a man.

The word *abuse* itself etymologically refers to significant notions of *deception* or *misuse*.

Use implies utilization and realization of a thought or an idea which occurs to one or several individuals and which then turns into action, becomes tangible and which, as such, results in a certain manifestation.

On the other hand, the term *deception* or misuse is of essential importance not only when speaking about AI, but about other accomplishments made by the man, thanks to the common sense given by the Creator, during his history, most frequently for the sake of benefit or doing good, for facilitating his existence on Earth. However, because of the constant presence of temptations, numerous accomplishments originating from the man's mental actions and skills, have turned into something destructive, devastating and, above all, soul-destroying. Would it have happened if an individual (or several individuals) had been restricted by Christian morality? This very question is of essential importance for the topic we discuss here.

5. Justification and unjustification of the feeling of fear of artificial intelligence

There is little probability that someone without any fear of God will think at all about AS and fear of it. As a matter of fact, if we pay attention to the stereotype with which most of us perceive experts from the domain of information and electronic technologies – that they are cold and almost heartless people who have solely numbers or formulas in their minds, with very little or often no empathy at all – it is not difficult to understand why we have a certain sort of fear of them. What we have said about people engaged in science is certainly a bias which has substantially originated from, for

example, film industry. Namely, there is a countless number of films with the scenes of laboratory research being conducted by people in white coats, focused solely on projects and realization of sometimes strange and destructive ideas. On the other hand, if we leave aside the fiction so characteristic of film art, it is necessary to take into account the factor of the past, i.e., the man's memory of the events which occurred and proved in practice that the line between human and inhuman use of a scientific discovery is fragile and easily broken. It is these experiences, undergone by humanity, that leave not only "a bitter taste in the mouth" in a large number of believing people, but they also cause fear that is essentially the fear of AI abuse.

The communities of believing people, and therefore the Church, in the sense it has in Christian theological thought, are often seen, to put it mildly, as conservative and "backward", i.e.., as the communities which prevent progress. Just as in many other things, there is a grain of truth in this assumption. Namely, particularly when it comes to the societies undergoing the desecularization process (including Serbian society as well), we recognize specific signs of elitist hermeticization and the existence of a specific culture of the closed society. Nevertheless, we should not forget that the individuals, primarily those highly positioned in the in the hierarchical structure of religious communities, those who are, so to say, ahead of their time, are actually the ones who not only evidently support technical-technological progress of civilization, but also strengthen it themselves with their work (an illustrative example Russian Orthodox priest is Pavel Florensky). Such people are naturally a minority in terms of percentage, but they need to be mentioned not only to show their noticeability in the domain of science, but also to indicate that religion and religiosity are not an obstacle for engaging in some other fields, such as art,[4] science etc.

What should be emphasized in relation to the topic we speak about is the circumstance that many religious communities, for their own reasons, have actually fenced themselves off in a sort of their own area in which they feel well. The result of it, among other things, is that such communities accept with difficulty, and often with open resistance do not accept what has been recognized by the leading figures of the community as something "bad". For example, inside some protestant communities, the followers used to be forbidden to watch TV (Đurđević, 2004, p. 125). On the other hand, with the emergence and expansion of computing industry and the recognition of the possibilities for profits, some protestant communities in the former Socialist Federal Republic of Yugoslavia turned to the hardware segment of computing industry because, as they thought, it had a different meaning of "looking at the screen". As a result of these changes, specifically in Adventist communities, the attitude was changed towards television and radio programs.[5]At the same time, this attitude, but totally reversed when it comes to the attitude towards television, electronic media, and even AI,

[4] We should use the opportunity to mention Bishop Irinej (Ćirić), PhD, who was canonized by the Serbian Orthodox Church several years ago and declared for a saint. He was also successful at painting.

[5] https://www.adventisti.net/o_nama/verovanja/zivot/hriscansko-ponasanje/ (accessed on 28.05.2024)

can be noticed today among a number of Orthodox Christians. In fact, especially from the late 199s, if we focus on the territory of the former SFRY and the countries established after its breakup, the movement of the so-called zealots emerged and strengthened within the corpus of Orthodox Christians. Zealots are people who advocate, so as to say, the "harder" trend in practising religion. If we leave aside their attitude primarily towards the higher hierarchy and certain part of the Serbian Orthodox Church clergy, within these *parasynagogues* (Jevtić, 2020) in the full meaning of the word, an evident change occurs in the attitude towards, for example, modern technologies.[6] Apart from famous "TV watching", i.e., the negative attitude towards the content of the TV program (namely, individuals do not accept having TV sets at home), these people have a clearly shown negative and even hostile attitude towards certain medical elements of modern civilization (primarily towards children's vaccination). They have this attitude towards AI as well because they believe these are elements appearing and existing with a reason – to change/endanger humanity. When speaking about zealots, it is interesting that almost every person who is a member of these communities also "exists" in social media and actively uses a mobile telephone and has access to the Internet. This is in itself quite a strong contradiction and definitely a separate topic for research.

If we return to the question of religious people's fear of AI, we will notice that fear of AI is simultaneously justified and unjustified. Justification or unjustification of the believers' fear of AI is basically founded on the problem of abuse or, more specifically, abuse of AI in those spheres where it may be or is directed against a man. The realistic justification of this fear derives from the experience undergone by humanity in the past and in the present. For example, the development of the nuclear energy research seems to have been redirected too easily towards its destructive, devastating application, i.e., the destruction of man/humanity. For a believer, the very existence of AI and its increasingly larger application in everyday life causes discomfort, first when it comes to who and in what manner will introduce it into society or, to put it simply, who is the one who will "control the controllers". AI, as it is presented to the broadest audience, strives to become a man's creation which, from the believer's perspective, will control or exist together with the man as God's creation. A substantial percentage of believers, when asked about AI and its broader application in everyday life, asks himself and others whether AI controllers are people who are guided by the moral norms of religion in their lives and professional activities. For example, the question is whether the sixth commandment "Thou shalt not murder" is something that those people, who shape AI and are members of Judaism or Christianity, observe in their own lives (Janković, 2010, pp. 81–82). The question is whether they, if they are followers of Islam, find this commandment something they will truly observe (Halilović, 2016, p. 7). Believers are also concerned, and therefore have the feeling of fear, whether and to what extent AI will affect

[6] http://www.eparhija-prizren.com/sr/vesti/episkop-atanasije-artemijeva-sekta-parasinagogaparacrkva/ (accessed on 13.05.2024)

many other ethical matters which have an impact both on social trends and on each of us individually. AI indisputably has a huge potential in essentially important elements of life, for example in medicine. Moreover, the possibility of overlooking a problem in radiology is noticeably smaller in AI than in a specialist doctor in this field who, for instance, examines many patients during one day and, because of exhaustion or some other human weakness, can make a mistake in reading an MRI etc. It is in these terms that we may speak of unjustification of fear of AI primarily having in mind that a doctor uses AI as a special tool, but that a radiologist is the one who makes the final decision about whether something has been diagnosed in the examination or whether the result of that examination is negative. Practically speaking, in the above-listed example AI can be used for the purpose of performing a larger number of examinations, but so that the doctor is the one who will describe the change in case he/she observes an anomaly in the screening.

In relation to AI application in everyday life, believers are also concerned about the episodes such as the one that took place in St. Paul's Church in

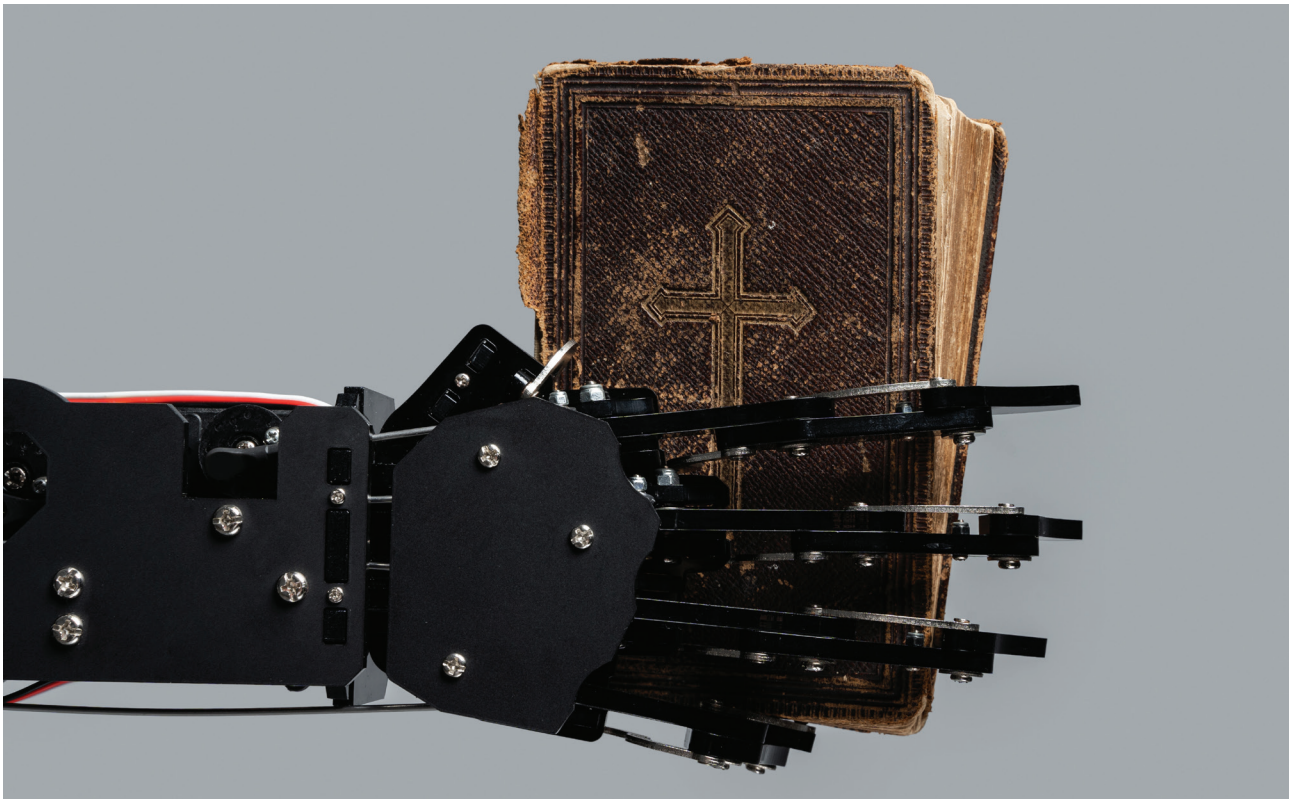


Photo: Shutterstock

Fürth, Germany, in 2023. Namely, the sermon within the Protestant service, in this church, full of believers, was performed by one of four avatars (two boys and two girls). The sermon words coming from the avatar generated by artificial intelligence referred to the questions of leaving the past behind, focussing on present challenges, overcoming fear of death and not losing trust in Jesus Christ.[7] The question which arises definitely refers to the purpose of such an experiment, i.e., what was supposed to be achieved by it. Leaving aside the short-lived element of the phenomenon and fascination, the use of AI in religion is, among other things, reflected in an attempt not only to "replace" the man as an essentially important factor in the worship service, but also to "replace" God in a certain manner, which is impossible and unacceptable from the perspective of the orthodox and the believing ones. "I am the way and the truth and the life. No one comes to the Father except through me" (John, 14:6). On the other hand, what such experiments bring up to the surface is exactly what causes fear among believers, and that is abuse of AI by those who set its "task", i.e., those who directed it towards certain segments of man's life. This challenging attitude of AI controllers towards man's religious feelings is counterproductive for all not only because they cause unpleasantness and discomfort among potential "consumers" of what AI offers, but also because among those who direct and control AI it actually creates a soul-destroying feeling that they awaken "like divine beings" (Book of Genesis, 3:5), the feeling which has been abused in the original sin.

6. Conclusion

"Fear can be an important motivational factor in social and political behaviour. Fear culture is a tool used by elites to control and direct the attention of the lower strata of society in the manner which suits elites and facilitate their maintenance of power" (Luknar, 2023, p. 153). If we add the circumstance that today's public life functions on the basis of information consumed by an individual online, i.e., by taking in information whose reality and accuracy are often disputable, we can very easily find ourselves in a situation when due to such semi-information we become inundated by all sorts of fears. Fear of AI has its justification primarily when it comes to its abuse. There may be most diverse kinds of abuse and addressing them entails a broader analysis supported by a multidisciplinary approach. If in broadest terms, from the perspective of a believer, we look at the above-mentioned kinds of AI abuse, we can clearly see that it is actually a tool in the hands of those who, for their own reasons, stand up against the God-established order. It is for this reason that responsibility in religious communities is borne by the people who are placed for spiritual guides of believers because they are expected primarily to perceive the problems faced by the modern man and then, through communication with believers, to direct them towards how and in what manner to avoid challenges and abuses of what is offered by technological progress in the broadest meaning of

[7] https://www.danas.rs/svet/vestacka-inteligencija-odrzala-propoved-u-crkvi-u-nemackoj/ (accessed on 15.05.2024.)

the word. Religious communities more and more frequently address the question of AI and challenges it brings along. In order to avoid prejudice and a priori refusal of the AI use, it is desirable that clergy, i.e., believers in broader terms, become familiar more thoroughly with the given phenomenon. From today's perspective, we find it incomprehensible, but we should recall that something like that is customary, just as plain potatoes and their introduction into people's nutrition used to be a huge problem in our region. An almost identical situation is also with the acceptance of AI by believers. It is definitely good to have the feeling of thoughtfulness, caution, and even fear of AI, first of all because of its potential abuse. Namely, we should recall the words of Apostle Paul: "I have the right to do anything, but not everything is beneficial; I have the right to do anything, but I will not be mastered by anything" (1 Corinthians, 6:12), and "All things are lawful for me, but all things are not expedient. All things are lawful for me, but all things do not edify" (1 Corinthians, 10:23).

References

Đurđević, S. (2004). God's Church of the Seventh-day Christians – Anthropological Approach. *Religija i tolerancija*, No. 2, 119–133. Available at: https://www.ceir.co.rs/index.php?strana=images/stories/rit.php&ostalo=02 [In Serbian]

Halilović, N. (2016). God's commandments in terms of developing moral education among students. *Novi Muallim*, XVII (68), 3–9. Available at: https://ilmijja.ba/ojs/index.php/casopis1/issue/view/75 [In Bosnian]

Janković, B., Janković, N. (2010). Homicidum, *Engrami*, XXXII (3), 79–88. Available at: https://www.engrami.rs/downloads/engrami-3-2010.pdf [In Serbian]

Lipij, A., Bojić, Lj., Cvetković, V. (2023). Artificial intelligence and Holy Fathers: Holy Fathers' guide through the challenges of the artificial intelligence and super-intelligence. *Crkvene studije*, No. 20, 193–209. DOI: https://doi.org/10.18485/ccs_cs.2023.20.20.11 [In Serbian]

Luknar, I. (2023). Sociocultural conceptualisation of fear. *Srpska politička misao*, LXXIX (1), 143–161. DOI: 10.5937/spm79-42727 [In Serbian]

Prodić, S. (2024). *Essays on the relationship between religion and politics*. Beograd: Fakultet za evropske pravno-političke studije. [In Serbian]

Internet sources

Artificial intelligence held a prophet in the church in Germany. Available at: https://www.danas.rs/svet/vestacka-inteligencija-odrzala-propoved-u-crkvi-u-nemackoj/ (Accessed on 15 May 2024) [In Serbian]

Christian Behaviour (2024). Available at: https://www.adventisti.net/o_nama/verovanja/zivot/hriscansko-ponasanje/ (Accessed on 28 May 2024) [In Serbian]

How many people in Serbia are religious (13 April 2024). http://www.021.rs/story/Info/Srbija/105636/Koliko-ljudi-u-Srbiji-je-religiozno (Accessed on 12 May 2024) [In Serbian]

Jevtić, A. (2020). *Artemije's sect – parasynagogue and parachurch*. Available at: http://www.eparhija-prizren.com/sr/vesti/episkop-atanasije-artemijeva-sekta-parasinagogaparacrkva/ (Accessed on 13 May 2024) [In Serbian]

Uroš V. Šuvaković[1]
University of Belgrade,
Faculty of Education,
Department of Philosophy and Social Sciences
Belgrade (Serbia)

UDC 004.8:316
327:929 Кисинџер Х.
Review scientific article
Received: 19.08.2023.
Accepted: 23.08.2023.
doi: 10.5937/napredak5-52859

Creator of the present in an attempt to understand the future or what can be counted as, but not reduced to, Kissinger's legacy[2]

Summary: The text was written with an idea of reviewing the book *The Age of A.I. and Our Human Future* (2021/2022) by Henry Kissinger and his eminent associates Eric Schmidt and Daniel Huttenlocher, which was published in the USA when Kissinger was 99 years old. Since in the meantime this US statesman passed away at the age of 101, it was impossible not to take into consideration and, at least briefly, not point to Kissinger's main accomplishments in international relations and diplomacy, his attitudes relevant for the Yugoslav crisis, and then look at this diplomat's attempt in the book about artificial intelligence to perceive its potentials and, by understanding its advantages and shortcomings, to model the direction of the development of humanity. That is why this paper consists of two parts. The first part sketches Kissinger's influence of US foreign affairs and making far-reaching foreign policy decisions of the USA, which created the world we used to know in the second half of the 20th century, with a particular emphasis on the establishment of relations between the People's Republic of China and the USA, and the détente policy of the USA and the USSR. In the second part, attention is dedicated to Kissinger's analyses and warnings about the emergence and development of artificial intelligence. Although he delved into this field at a rather advanced age, his insights are extremely important, particularly given his striving for international cooperation in the regulation of this matter.

Keywords: Henry Kissinger, diplomacy, artificial intelligence

[1] uros.suvakovic@uf.bg.ac.rs ; https://orcid.org/0000-0002-2141-8698

[2] This paper is the result of the research conducted with the financial support of the Ministry of Labour, Technological Development and Innovation of the Republic of Serbia, according to the Agreement on the transfer of funds for scientific research work of teaching staff at accredited higher education institutions, which was signed with the Faculty of Education of the University of Belgrade for the year of 2024 (No. 451-03-65/2024-03/200138).

1. Creating the present

Famous American diplomat Henry A Kissinger (1923–2023) was called, primarily by his critics, the hawk of US foreign affairs. His role in political violence in different parts of the world – as the national security advisor and the US Secretary of State in the administrations of President Nixon [3]and President Ford – is rather contentious, to say the least. From his influence on the ending of the Vietnam War, direct involvement in overthrowing from power Salvador Allende, the legally elected socialist president of Chile[4] and the coup[5] in which Augusto Pinochet, the profascist dictator, came to power in the CIA scenario, via his support to the Argentinian neo-Nazi regime ("the worst of all Latin American monsters those years" [Chomsky, in: Chomsky, Waterstone, 2022, p. 42]), participation in the decisions about the interventions in Cambodia and Laos, which caused the death of tens of thousands of people, to his support to Indonesia in East Timor, to Pakistan against Bangladesh, the coup against Archbishop Makarios in Cyprus etc.

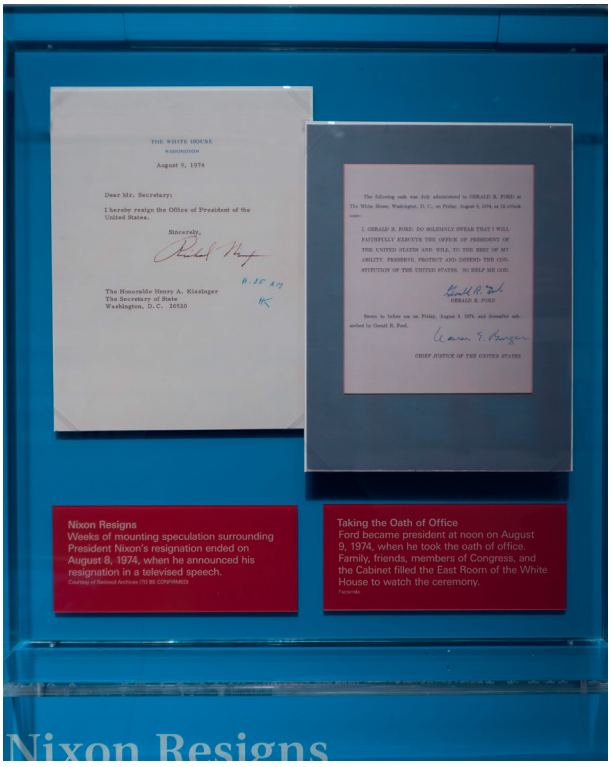
On the other hand, those in favour of somewhat different allegories, described him as a pigeon (of peace), having in mind his creator role in the end of the Vietnam War, when the USA needed "an honourable exit" from it. Kissinger ensured the end of the conflict through negotiations with Lê Đức Thọ, the leader of North Vietnam, in Paris in 1973, which earned them both the Nobel Peace Prize that year[6].

[3] In President Nixon's administration, during a period of time, he had two most important state functions at the same time: the national security advisor and the US Secretary of State, while President Ford left him in charge of the latter function. Chomsky paraphrased Kissinger's definition of an expert as someone who is able to "articulate the consensus of the powerful", concluding that it consequently enables him also "to manage jobs on his behalf" (Chomsky, in: Chomsky, Waterstone, 2022, p. 51). Although this is certainly an exaggeration, which is also indicated by Professor Visković in the same source, we cannot help pointing out the extent to which Kissinger's influence was valued – there is a famous anecdote from the 1970s: "What would happen if Henry Kissinger suddenly died? Richard Nixon would become President of the USA!" (RTSa, 2023).

[4] "An example of a successfully elected Marxist government in Chile would definitely affect other parts of the word, and even the value of the precedent, particularly for Italy; in addition, it would substantially affect the world balance and our position in the world", wrote Kissinger in his confidential memorandum to Nixon, after which Nixon decided to support the coup against Allende (according to Chomsky, in: Chomsky, Waterstone, 2022, 86–87).

[5] A long time ago, Dragan Simeunović established a characteristic distinction between a coup, as a higher gender concept, and a *military coup*, "as a particularly militant form of a coup d'etat" (Simeunović, 1992, p. 139).

[6] There are three historical curiosities in relation to the Nobel Peace Prize that year. The first one, relevant for us, is that Kissinger's main counter-candidate was Yugoslav President Tito, and that the voting result was 3 : 2 in Kissinger's favour, whereas two members of the Nobel Committee publicly resigned, and the announcement of that decision provoked protests in the USA and worldwide. We do not agree with B. Dimitrijević, who believes that the decision of the Yugoslav top leadership to nominate Tito for this prize actually resulted from their sycophancy, and not from Tito's role in the establishment and promotion of the Non-Aligned Movement (RTS, 2023). On the contrary, Tito's merit in that respect was inambiguous and the decision adopted by overvoting would certainly not have caused the protests. Of course, this does not exclude the attempt of sycophancy a, but cannot be reduced solely to it. In fact, Serbia is still benefiting today from Tito's Non-Aligned politics since this large group of countries is considered the successor of SFRY and, although without a particularly solid foundations, is associated with Tito. The second curiosity was that the formal award ceremony was not attended by either of the Nobel



Gerald R. Ford Presidential Museum. Richard Nixon resignation letter to Henry Kissinger and transcript of Ford's oath of office to become president, Michigan, USA
Photo: Shutterstock

Kissinger certainly contributed to the end of the Arab-Israel conflict which broke out because of Israel's occupation of the Golan Heights. That is when the so-called shuttle diplomacy was promoted, which produced results due to Kissinger's skilful performance. Finally, there is and the détente policy with the Soviet Union, and reaching agreement SALT I signed by Nixon and Brezhnev. Some time before that, Kissinger also reached agreement about the normalization of relations between PR China and the USA in the negotiations with the then Chinese Prime Minister Zhou Enlai. Namely, in 1972, Nixon and Mao Zedong signed the agreement, reaching the position of the recognition of "the policy of one China" – the People's Republic of China, which was preceded by its membership in the OUN, when China replaced Taiwan in this organization, with the ritual opposition of the USA. During his last visit to China in 2023, several months before his death, where he was welcomed as an "old friend" (Xi Jinping), Kissinger said that the USA and China "cannot afford being hostile to each other" (BBC NEWS, 2023). This message sounds substantially different from the conclusions of the analysis performed by Brzeziński (Brzeziński, 2013).

It should be observed that in his diplomacy Kissinger was guided by pragmatics. He did not hesitate to change his initial position in the course of the process to the point of unrecognizability. It was important for him to achieve the set goal of US foreign policy, while, in line with the Machiavellian principle, any means justified the ends. If something could be decided in negotiations. Fine; if a war was necessary, there was no need to hesitate; if the war has begun – America must, at least in some way, end it successfully. Therefore, the question of moral principles *in foreign policy* (italics by the author) was

Peace Prize winners. Namely, Lê Đức Thọ refused to accept the prize because of the continued US support to Saigon after the signing of the Paris Peace Accords. Finaly, in 1975, after the troops of North Vietnam entered South Vietnam, Kissinger himself wanted to return the Nobel Prize. As a matter of fact, he immediately donated the prize money to charity – to the children of the US soldiers killed in the war.

not the guiding one; it was necessary to achieve the goal. We call it Machiavellianism, some others speak about moral relativism, even "amorality" (Visković), while Kissinger's followers call this approach Realpolitik, and the same is done by the largest number of theoreticians of international relations.

It should also be said that Kissinger was an old-time statesman. It means: he was very well educated, which is not the characteristic of the members of the political class in modern Europe[7] and the USA. He earned his PhD at Harvard University, where he also worked as a professor. Among other things, he wrote *Diplomacy*, one of the best and definitely most translated textbooks of the history of diplomacy and international relations, drawing both on the presentation of different theoretical positions and on his own experience acquired in his practice of being involved in diplomatic affairs.

This education was exactly what helped Kissinger to understand the Yugoslav crisis, or at least the essence of its stages so far, since the author of this text does not believe that this crisis has ended yet. However, it should be noted that Kissinger's youth in Germany, and even his subsequent conservative attitude I ideology, also caused his, so as to say, "historical misconceptions" which he kept until his death, for example, that the assassination of Ferdinand in Sarajevo was "an act of terrorism" (Kissinger, 1999a, pp. 174–175) and that the assassin was a "Serbian nationalist" (Kissinger, Schmidt, Huttenlocher, 2021/2022, p. 83). This is definitely a German paradigm of war instigators who transferred the blame on those refusing to be enslaved, the view of the historians of the country that does not accept the defeat in the First World War[8]; however, the historical truth is diametrically opposite: the assassination against the occupier was an act of liberation, and Gavrilo Princip was a Yugoslav nationalist by all parameters of that time. In the above-mentioned textbook, this is how Kissinger describes the formation of Yugoslavia after the First World War, while at the same time listing causes for it:

"The new Yugoslavia fulfilled the aspirations of South Slavic intellectuals. But to create that state, it was necessary to cross the fault line of European history, which divided the Western and the Eastern Roman empires, the Catholic and the Orthodox religions, the Latin and the Cyrillic scripts – a fault line running roughly between Croatia and Serbia, which had never in their complex histories belonged to the same political unit. The bill for this came due after 1941, in a murderous civil war which started all over again in 1991" (Kissinger, 1999a, p. 202).

It is due to the application of the principle of historicity in interpreting events, i.e., contextualization[9], that he understood the civil war in Bosnia and Herzegovina much better than other politicians and theoreticians in the West.

[7] Apart from few honourable exceptions, which include President Vučić, President Putin and Prime Minister Orban, and I apologize to all those current European leaders who did not occur to me while writing this paper.

[8] And in the Second World War as well, as it seems to us?

[9] Here we can recognize his sentence in the book about artificial intelligence: "To be useful or at least sensible, information must be understood as an objective of culture and history" (Kissinger, Schmidt, Huttenlocher, 2021/2022, p. 33), which confirms our claim about the methodological approach to understanding problems applied by Kissinger.

Uroš V. Šuvaković
Creator of the present in an attempt to understand
the future or what can be counted as, but not reduced to,
Kissinger's legacy

US Secretary of State Henry Kissinger at the reception by Yugoslav President Josip Broz Tito, in Belgrade, on 4 November 1974.
Photo: Museum of Yugoslavia, Photo-archives of Josip Broz Tito

"We act as if we're trying to force the Serbs back into a mythical Bosnia that has never existed in history. There is no Bosnian language. There is no Bosnian culture. Bosnia is an administrative entity which contains Croats, Muslims, and Serbs, artificially created as a subdivision of Yugoslavia and foolishly recognized as a state by the Western powers in 1991. If you had looked at Serbian history, for 600 years they have fought not to be dominated by Muslims. And why the United States should violate its own principle of self-determination to bomb them back, why our media should call them the separative Serbs, what are they separating from that has ever existed? So I believe that what we should do is create a Muslim state or recognize a Muslim state, permit the other nationalities either to make themselves independent or join Croatia and Serbia, as the case may be, and not get ourselves

involved in a Balkan war that cannot end, even if we win it. We cannot stay there in permanent occupation. The Germans required 17 divisions to police Serbia in the last war, and we do not have a stomach for anything like that, nor should we have" (Kissinger, 1994).[10]

From this slightly longer quote we may see that Kissinger was extremely familiar with the Yugoslav situation and the relations in Bosnia and Herzegovina, and that he advocated for its solution in the most democratic way possible – by applying the principle of self-determination of a nation, which is an "American principle" in foreign policy ever since Woodrow Wilson.[11] The statements from this interview with Kissinger are in full compliance with the policy pursued by official Serbia of the time (as well as the Federal Republic of Yugoslavia) and that, if they had been accepted in the past, they would have meant stability in the Balkans and in contemporaneity.

He will even more sharply oppose the NATO aggression against the Federal Republic of Yugoslavia in 1999. We believe that Kissinger's assessment of the document (or so-called agreement), offered by Madelaine Albright to the Yugoslav delegation in the castle of Rambouillet, presents its essence most concisely and precisely, while at the same time he expresses indignation at the fact that it was the product of the US diplomacy of the time. It is clear that Kissinger's assessment, just as the document itself, constituted integral part of the history of diplomacy of the 20th century:

"The Rambouillet text, which called on Serbia to admit NATO troops throughout Yugoslavia, was a provocation, an excuse to start bombing. Rambouillet is not a document that an angelic Serb could have accepted. It was a terrible diplomatic document that should never have been presented in that form" (Kissinger, 1999b)[12].

Moreover, he thoroughly analyzed Clinton's address on the occasion of beginning the aggression, placing it, of course, into the context of the document offered in Rambouillet, as well as its potential implications. Establishing, first of all, that, contrary to the US mainstream propaganda, "Slobodan Milošević is not Hitler", and that, contrary to Clinton's claims, "neither Milošević nor any other Balkan leader is in a position to threaten global balance", Kissinger emphasizes that, "unlike Bosnia, Kosovo is a war for the territory considered by the Serbs a national sanctity. That is why in Belgrade there were few, if any signs of opposition to Mi-

[10] Explaining "Yugoslav freedom of action" in relation to the USSR after the Second World War, Kissinger admits that it derived from the fact that "Yugoslavia liberated itself from German occupation thanks to its own guerilla forces" (Kissinger, 1999a, p. 342).

[11] In fact, before him, if we take into account the American "independence war", Wilson converted this principle in a typical professor's way into the principle of foreign policy, which was at that moment in line with Lenin's principle of "self-determination of a nation". Finally, the OUN Charter included two equally important principles: respect for the territorial integrity of the UN member states and a nation's right to self-determination.

[12] It would be good if those from the ranks of the current opposition in Serbia and, unfortunately, from parts of the current academic community, who claim that we could have accepted this document as a state, from time to time, rad this statement by Kissinger, just "for the sake of being smarter". All of them should thoroughly think about this statement.

lošević's policy regarding Kosovo. The Serbs have rejected the Rambouillet agreement because they see in it a prelude to independence for Kosovo. They also see the presence of NATO troops as the sort of foreign occupation Serbia has historically resisted against the Ottoman and Austrian empires, Hitler and Stalin. Even if they are bombed into capitulation, they can hardly be expected to be willing supporters of the outcome. As for the KLA, its goal is independence, not autonomy; it acceded to Rambouillet as a tactical device to unleash NATO air power against the hated Serbs. The KLA is even less likely to agree to autonomy under Serbian rule now that Serbia has been so weakened by the NATO air campaign. The KLA will not turn in its weapons to NATO forces. And NATO forces will have no domestic support if they fight the KLA to impose disarmament. Nor will the KLA acquiesce to Serbian forces policing its frontiers" (Kissinger, 1999c). In the same text, Kissinger predicts that "as Kosovo moves toward independence, the pressures on Macedonia, a third of whose population is Albanian, will increase. Namely, they will want to be granted the same self-determination as their brethren inside Serbia".

Despite such understanding of the problem, Kissinger advocated for the continuation of the air campaign – so as to prevent the compromising of the NATO – while at the same time passionately rejecting the beginning of the land invasion, but also leaving the possibility of its consideration for the purpose of maintaining "the NATO credibility". Finally, he predicted the outcome: long-term presence of the NATO in Kosovo and Metohija or in Macedonia[13] "to prevent the Balkan conflict from widening", as well as long and difficult negotiations which he believed would result in "some form of Kosovo independence".

Kissinger similarly addressed the problem of Ukraine. Ten days before the referendum in Crimea in 2014., where a decision was made to return this strategically important peninsula to the territory of Russia, in his text published in the *Washington Post* Kissinger pointed to the premises as a starting point in preventing the conflict, since it was likely to occur in the near future. He also analyzed the history of Ukraine, showing that it was a country with only 23-year-long history of independence. In addition, he recalls Kiev Russia and the fact that the Russians received Christianity in that territory, stressing that "the Wests must understand that for Russia, Ukraine can never be just a foreign country". He asserts that Ukraine was part of Russia for centuries and warns that this country is divided into western and eastern parts in terms of identity, and that the differences are huge: while the western part is Catholic and speaks Ukrainian, the eastern part is Orthodox and speaks Russian. He believes that the problem is in the fact that both parts are trying to impose itself to the other one in terms of rule, but that "any attempt by one wing of Ukraine to dominate the other – as has been the pattern – would lead eventually to civil war or breakup. To treat Ukraine as part of an East-West confrontation would scuttle for decades any prospect to bring Russia and the West – especially Russia and

[13] As it has turned out, in both regions.

Former Secretary of State Henry Kissinger, chairman of President Reagan's Bipartisan Commission on Central America, presides over a meeting at the State Department, Washington DC, January 6, 1983

Europe – into a cooperative international system" (Kissinger, 2014). Unfortunately, Kissinger's warnings, issued in a timely manner, were not taken into account. What he had predicted actually occurred and today there is an ongoing war between the NATO (the USA + the EU) and Russia in the territory of Ukraine (see Šuvaković, 2023), which has lasted two and a half years and its end cannot be predicted. In the same text, Kissinger opts for a solution that will not be "absolute satisfaction but balanced dissatisfaction", considering it a test for the agreement sustainability.[14] Accordingly,

[14] If we are able to follow the guiding idea of President Vučić's policy regarding Kosovo and Metohija, then It coincides with Kissinger's idea. When searching for a compromise, "it is impossible for someone to gain everything, and for someone to lose everything" – that is exactly what the West keeps "offering" to Serbia.

he advocated for Ukraine's right to access different international organizations, including the EU, but not the NATO since he opposed any change in Ukraine's borders as it would undermine the world order and, in that respect, he was in favour of keeping Crimea as its integral part, but with the guarantee of great autonomy for the people living there and for resolving the status of Sevastopol as the base of the Black Sea fleet in the long run. Finally, he believed that Ukrainian government should be a reflection of the national will and that it would be wise for the government to take a position similar to the then position of Finland: being turned towards the West, but not challenging Russia.[15] Two years later, he revised his position and advised the then US President Donald Trump to accept that Crimea belonged to Russia, but without any official recognition, thinking that Crimea must not be a problem on the relation Washington–Moscow (Kissinger, 2016).

After the outbreak of the Russian-Ukrainian conflict, several statements by Kissinger were differently interpreted. However, evolution existed and was evident throughout the length of the conflict: he became insistent that Russia must not be allowed to have benefits from the war; he advocated for negotiations and allowed for the possibility of Ukraine becoming a member of the NATO after the end of the negotiations. Therefore, the negotiations were his primary choice but as the conflict escalated, his attitude also shifted closer to the attitude of the collective West.

2. Looking at the challenge of the future that has already started

It is difficult to predict social consequences
of artificial intelligence.
(Kissinger et al., 2022, p. 56)

Being accustomed to seeing Kissinger's practical and academic work through the prism of dealing with diplomacy and international relations, our academic public did not show particular interest in the book by Henry Kissinger, Eric Schmidt[16] and Daniel Huttenlocher,[17] *The Age of A.I. and Our Human Future*, published in the USA in 2021. In our country, it was translated and published in 2022 by Klub plus, which serves a credit to this small publisher. In Serbian social periodicals, as far as we know, only three reviews of this quite important monograph have been published – one from the perspective of political science (Orlović, 2023), the second from the aspect of the safety relationship between man and artificial intelligence (Marković, 2023) and the third from

[15] Unfortunately, Finland has recently abandoned this position and become a member of the NATO. We firmly believe that this decision will prove to be rather bad for the Finns in the future.

[16] CEO and President of Google Company (2001–2011) and subsequently its executive manager and technical advisor. He deserves credit for convincing Kissinger to attend a lecture about artificial intelligence at the Bilderberg Meeting in 2016 (Bisenić, 2021).

[17] The first dean of MIT Schwarzman College of Computing. He worked on the foundation of Cornell Tech, post-graduate studies in the field of digital technologies within Cornell University, New York, where he was the first dean and the vice-chancellor.

the perspective of historical science (Dimitrijević, 2024). This book certainly deserves more attention for minimum two reasons: first, it was written by the people who have plenty to say about this issue, so that much can be also learned from them and, not less importantly, the issue of artificial intelligence goes deep not only in every science and scientific discipline, but the question of its emergence and development also goes into the foundations of science as a human creation, ultimately questioning it and radically changing human society. Therefore, artificial intelligence is a technological challenge, although being a revolutionary technology,[18] but it is much more than a social challenge poses, because it revolutionizes our life and our view of humanity, even our existence. That is exactly what is suggested by the authors of this book and it is absolutely a pity that the book has not been sufficiently recognized by our scientists.

It is too bad that the authors did not clearly which of them wrote specific parts of the book, thus leaving us to make conclusions about it on the basis of knowing their work in the past. However, Kissinger's signature is recognizable in certain parts (historical development, worldview from human perspective, international agreement on regulating nuclear weapons, how to control the development of artificial intelligence etc.), just as that the final editing of the text seems to have been done by him (definitely not without the co-authors' consent).[19] That is why the introductory pages contain the note that there is no full agreement among them: "to some degree, we three differ in the extent to which we are optimists about artificial intelligence. But we agree that technology is changing human thought, knowledge, perception, and reality – and, in so doing, changing the course of human history" (Kissinger at al., 2022, p. 3). However, as it can be subsequently seen from the book, the authors do not ascribe these revolutionary changes to every technology. On the contrary, they indicate that the emergence of various new technologies has led to changes, but that "only rarely has technology fundamentally transformed the social and political structure of our societies" (Ibid., 12). The authors mention Gutenberg's printing press as an example of a revolutionary technology (Ibid., 125).[20] The potential for such dramatic change, according to the authors, is possessed by artificial intelligence,[21] whose "outcome will be the alteration of human identity and the human experience of reality, at levels not experienced since the dawn of the modern age" (Ibid., 2).

[18] This insistence on *artificial intelligence technology* is of great relevance because every technology is value neutral, so it can be used both for noble and destructive purposes, unlike science which is value "utral", i.e., based on at least some most common human values, such as humanism (Marković, 1994a). This is exactly what the authors of this book insist on, from its introductory pages onwards.

[19] That is why in the paper hereinafter, as well as for the sake of applying the principle of economy, when citing this book, we will mark it as (Kissinger et al., 2022), without the slightest intention of minimizing the contribution made by Eric Schmidt and Daniel Huttenlocher to its creation. As far as we know, their work on this monograph lasted for as many as five years and we must honour it by saying that their work was quite fruitful.

[20] About the impact of the media on the change in the social structure, see Debray (2000) and Vučković (2020).

[21] Hereinafter: AI.

The authors of the monograph first present AI successes using the examples of several programs based on it: AlfaZero in chess, against the previously most powerful form of artificial intelligence created for this game; the discovery of a novel and most comprehensive antibiotic – halicin, at the Massachusetts Institute of Technology (MIT), in which artificial intelligence participated; finally, GPT_3 creation by OpenAI. What was new here? Unlike previous chess playing programs, based on fast processing of *human understanding* of this game, acquired during a millennium-long period, AlfaZero only knew the rules of the game. The rules were given to it and its instruction was to win or get the best position possible. Hence this program in some games also sacrificed those pieces human players considered vital, including the queen. "It had a logic of its own, informed by its ability to recognize *patterns* of moves across vast sets of possibilities human minds cannot fully digest or employ" (Ibid., p. 6). As for the discovery of halicin, according to Kissinger et al., as shown on the example of 2,000 different molecules, AI, "*learned* the attributes of molecules predicted to be antibacterial". What is pointed out by the authors of the monograph is that AI simultaneously, without being asked, established new attributes of those molecules, neither identified nor encoded by the researchers. When faced with 61,000 different molecules, FDA-approved drugs, and natural products for molecules:

"1) that AI predicted would be effective as antibiotics,

2) that did not look like any existing antibiotics,

3) that AI predicted would be nontoxic",

AI chose only one molecule fitting these criteria, and it was named halicin. The same scenario was repeated once again: "artificial intelligence identified relationships that had escaped human detection – or possibly even defied human description" (Ibid., pp. 7–8). Even after the antibiotic was discovered, humans could not articulate precisely *why* it worked[22]. Finally, with the introduction of GPT_3, AI progressed to the level of a *generative model*. It is now able to offer the end of a sentence, to write a passage or a shorter tractate, to answer the questions posed about everything that has corresponding information contained on the Internet. That is why it may be classified as a *virtual autodidact* – "it learns" from the information left by humans – consciously or unconsciously – about themselves or others in the web-space. However, there is also an important limitation here: what does not exist on the web can hardly be a subject about which GPT_3 will be able to discuss.[23]

Kissinger et al. found their book on three above-mentioned examples of AI application in

[22] It is not clear from the book whether AI may have been assigned the task of finding a medicine against COVID-19, since it derives from the context that the discovery of halicin coincided with the peak of the pandemic or immediately before it. If so – what was its result? If that task had not been assigned at all – why not? Kissinger's co-authors who are still alive would have to answer this question.

[23] That is, for example, what escapes the attention of colleague Dimitrijević (2024) in his review of the book, when he writes about the dangers borne by AI to historiography. Dangers definitely exist (not only when it comes to historical science), but they are related only to what is available in any way through the web. Of course, in the large historical archives worldwide, all most important collections have been digitalized, and thus everything has become available; it is by no means the case

chess, pharmacology and language, pointing out that *in those cases* AI proved to possess the possibilities which are *similar to human ones or exceed them*, while humans *do not understand the methods, or procedures through with AI achieves that.* Furthermore, it rightfully causes their concern for the future of everything created by man – from the existing relations between countries to the meaning of existence of science, philosophy and even man himself. Namely, even if the development of AI does not reach the degree of AGI,[24] in the light of Descartes' maxim "I think, therefore I am", the authors quite reasonably wonder: "Who are we here?" Is human identity brought to question since the intellect has always been (and still is) what distinguished man from other living beings on the planet. Now, all of a sudden, there is something that "understands>", but it not the human intellect? Is man encountering competition on planet earth? Wil man still create his own future? Will he rule the technology or vice versa? How will AI affect our culture, our concept of humanity and, after all, our history?" (Kissinger et al., 2022, p. 11). All these questions are raised by this monograph while offering possible answers to some of them.

The authors point to AI being based on the previously created computers and the Internet, as well as on networking.[25] According to them, artificial intelligence "learns", and the basis of that are big datasets, which are digitalized, networked and available for machine learning. It is stated that there are three forms of machine learning currently used: a) supervised, b) unsupervised, and в) enhanced learning (Ibid., pp. 40–42). These models have been applied in the creation of different types of AI which were very well selected by the authors for the purpose of illustration. They point out that "AI fragility" lies in the "shallowness of what is learned", since the connection between *input* and *output* in supervised and enhanced learning is essentially a

either in the largest Serbian archives or in the local ones. Therefore, what has not been the subject of digitalization will still be explored by humans. It does not mean, of course, that in the forthcoming period the archive material will not be digitalized, but I only point to the fact that the human researching space has not been exhausted and, moreover, that our findings about ourselves would be expanded if our historians were slightly more involved in local history, or even social history, which have somehow escaped their attention in favour of political history, either general or national. Naturally, this also refers to the research in the fields of other sciences, at least from the domain of social sciences and humanities, where scientific methods in the research of certain subjects will still have to be applied by humans. It is certain, for example, that the meaning of many archaisms or Turkish words in the Serbian language is not familiar to GPT3, so that the functions It can have in the modern language remain reserved solely for it, at least for some time in the future.

[24] AGI (Artificial General Intelligence), which is rather unlikely having in mind the speed of the development of AI to date, which has occurred exponentially. The predictions speak in favour of the fact that this degree will be reached in the next several years, and that it will be followed by Artificial Super Intelligence (ASI) that, by all accounts, surpasses human potentials. About classification see (Mandić, Miščević, Bujišić, 2024, p. 6). Kissinger et al. do not consider the degree of ASI in this book, but aspire to find, so to say, *modus vivendi* between AI and man, and to offer options as to how to ensure security of humanity from the huge possibilities of (ab)use possessed by the still unachieved AGI.

[25] We might safely claim that Castells by no means considered this implication of networking when, inspired by the "success" of the Arab Spring, he wrote his monograph *Networks of Outrage and Hope* (Castells, 2012/2018), particularly not the countless possibilities of the personal identity abuse offered by AI nowadays.

different property from "true human understanding with all its numerous levels of conceptualization and experience". In addition, it also derives from the fact of its "unreasonableness" – "AI does not know what it does not know" (Ibid., p. 50). Different kinds of machine learning and the pace of AI development assert that in the future decisions will be made in three main manners: „via people (known to human minds), then via machines (which is becoming known) and via cooperation between people and machines (which is not only unknown, but also unprecedented)" (Ibid., p. 14). Due to the unprecedented nature of the process, machines transform from human tools into – *human partners* (Ibidem). The main challenge is, to put in in the language of management, how people will manage to preserve the status of the "older partner". The development of procedures which will reliably test whether AI will function in line with our goals and expectations is imposed as imperative (Ibid., pp. 50–52).

The authors particularly address the question of network platforms and the "network platform geopolitics". They indicate that the most significant global network platforms were made in the USA or in China, which are, together with Russia and India as a "new power in this arena", also the "main stakeholders" in the creation of artificial intelligence, while the EU[26] has not even entered the race yet, if we exempt the adoption of certain European regulations in this field (Ibid., pp. 74–77). The result is that network platforms are created so as to cover the regions which are, primarily in commercial terms, a priority for the USA and China.[27] According to the authors of the book, this adds another important segment to making foreign policy decisions – the commercial interest of different platforms, which is further compounded given that the interest is often based on "the priorities of buyers and research and technological centres, both of which may be far from capital cities" (Ibid., p. 60). The key question for the authors is platform administration, giving priority to some content while disturbing or removing other content, criteria by which selection proceeds, which is increasingly less human and being left to AI. Platforms are the creations which cannot be compared to anything

[26] As for the role of the EU in this field, it is important to observe the statements by the authors of the book, as well as the terminology used by them (italics by the author of the paper): "So far *historical global powers* such as France and Germany have appreciated freedom and independence in their technological politics. However, *peripheral European countries* with the recent and direct experience of foreign threats – such as post-Soviet Baltic and Central European countries – have proved more willing to identify themselves with the 'techno-sphere' led by the USA" (Kissinger et al., 2022, p. 77). This is Rumsfeld's division into "old" and "new" Europe, only communicated somewhat more impertinently. Regardless of that, this definitely makes it more difficult for the EU to assume the role of the unique actor in this domain of competence.

[27] The largest Chinese platform is WeChat, with the purpose similar to WhatsApp. Last year, during my visit to China as a representative of the Foundation for the Serbian People and State, at the invitation of the Ministry of Foreign Affairs of the Communist Party of Cina, I asked my kind hosts why they had their own application instead of using the globally existing one. "Why would Americans collect data about our market when we can do it on our own?!", they answered, thus actually explaining everything in a counter-question. Of course, this is the interest, as admitted by the authors of the book as well, only great powers which make unique regions-continents, since a small number of users of a platform would by the authors of the book would render meaningless the interest in its existence.

in the "predigital age".[28] It results in the "standards of their community becoming equally influential as state laws" (Ibidem). However, "what seems intuitive to a software engineer may be confusing to a political leader or inexplicable to a philosopher. What a consumer greets as an amenity, a national security official may see as an unacceptable threat, or a political leader may reject it as unfavourable for national goals. What one society may accept as a welcome guarantee, the other society may interpret it as a loss of choice or freedom" (Ibid., p. 61). Here it is definitely necessary to take into account the role of network platforms in the placement of disinformation, the possibility of AI doing it quite convincingly and beyond state borders, globally, but that it can also be used in the defence against such "attacks". However, here a question also arises: what is truth and what is a lie in the age of semi-truth? The authors wonder whether there is a legitimate interest of the public in reading "lies" generated by AI. Isn't their prevention actually censorship in a new guise? Or in an old guise, through taking protectionist measures for certain platforms, e.g., TikTok-y in the USA and India (Ibid., p. 73).

In our opinion, the key chapter of this extremely important book is the one addressing the question of security and the world order, which is questioned by the advent of AI. Namely, the existing system is based on different agreements; it is the result of numerous and laborious diplomatic negotiations, secret and public connections which have created current international relations. It is particularly important when it comes to nuclear weapons, their control and relations between nuclear superpowers. "No great country cannot afford ignoring security dimensions of artificial intelligence. The race for the strategic advantage of artificial intelligence is already underway, particularly between the USA and China and, to a certain extent, Russia" (Ibid., p. 85). The authors point to the existence of nuclear, cyber technologies and artificial intelligence technology, and that each of the will no doubt play its role in the security strategy. That is why they apodictically believe that "the USA should further try to shape them" (Ibidem). Just as in the Cold War period, it is necessary to achieve the "balance of powers". This is something typical of Kissinger. However, during the Cold War, the danger was measurable, at least roughly. It was possible to count the number of missiles in someone's possession, their range, their ability to carry nuclear warheads and how many at the same time (Ibid., p. 93). With the warning that the principle of the "non-use of nuclear power is not an inherently permanent accomplishment" and that "it requires real and recognized balance" (Ibid., p. 92), the authors also point out that with the advent of cyber weapons and, so as to name it as the general concept of *artificially intelligent weapons*, it is substantially more difficult. First of all, cyber weapons are non-transparent. They may also be used from

[28] This dichotomy in pre-digital and digital age, analogously to pre-industrial and industrial age, points to the importance assigned by Kissinger et al. to the change occurring by the creation of the digital world. In an earlier text dealing with the matters of surveying public opinion and the use of the Internet for these purposes, we established the distinction between *virtual* and *real* public opinion, which becomes evident in the countries where the Internet is insufficiently developed, so that it is not possible to develop a representative sample in that manner (Šuvaković, 2008). Serbia used to belong to this group of countries, while the situation is completely different nowadays.

an office or a forest. It can very easily fall into the hands of terrorists (just as access to AI) or smaller countries which cannot develop nuclear potentials, but can have relatively significant possibilities for e.g., cyber warfare. Cyber weapons activation does not require large computers, although computer systems with a large memory power are definitely necessary to support their functioning. Kissinger et al. illustrate cyber warfare with the example of DDoS attacks, which look like legitimate requests for access to information, but are emitted simultaneously in such numbers that the system falls.[29] The authors point to the similarity between cyber weapons and chemical and biological weapons, in terms of effects of the attack expanding horizontally, and on the victims which were not the target of the attack, thus causing harm "in unintentional and unknown ways" (Ibidem).

Kissinger et al. emphasize that AI introduces "new horizons in the information space, including the field of disinformation. Generative artificial intelligence may create huge amounts of fake, but convincing information. Disinformation and psychological warfare with the aid of artificial intelligence, including the use of artificially created persons, images, video-clips and speech, are ready to produce upsetting new weaknesses, *particularly to free societies* (italics by the author).[30] Widely shared demonstrations produced seemingly realistic images and video-clips of public figures who say things they have never said… If a synthetic image of the national leader is manipulated by an opponent in order to entice dissent or issue misleading instructions, the question is whether the public (or even other governments and officials) will observe the deceit in a timely manner"[31] (Ibid., p. 97).

The authors rightfully express concern about who controls artificial intelligence. Regardless of whether artificial intelligence is used in conventional or (God forbid) nuclear warfare, "it is imperative to ensure an adequate role of human judgment in the supervision and use of force". But it will be insufficient and one-sided, and that is why it is necessary that the "governments of technologically advanced countries explore the challenges of mutual constraint with the aim of applicable examination" (Ibid., p. 101).

In its strategy, the USA has distinguished between *artificial intelligence guided weapons*, "which make war led by men more precise, deadly and

[29] Elon Musk's interview with Donald Trump, who is once again running the candidacy for the US President, was scheduled within Trump's campaign on 13 August 2024 on Musk's network X. The interview was delayed by as many as 40 minutes due to, as Musk explained, a mass cyber-attack, DDoS attack.

[30] Another term remaining from the Cold War. The author of this paper thinks that the countries with most freedom nowadays are exactly those countries definitely not seen as such by the authors of the book (Russia, China, Serbia, Hungary, Slovakia, Turkey, to name but a few), while the least free countries are those which have been considered "the free world" ever since the end of the Second World War (the collective West countries in general, without singling out any particular one).

[31] "Colourful revolutions", including the one in Maidan, by its consequences the most destructive European revolution, primarily to Ukraine, used manipulations and disinformation placed through "networks of outrage and hope". Today we can only imagine the potential for provoking unrest caused by cyber warfare, with the use of generative AI. It seems to us that those who would use it would be exactly the same as those in Maidan in the past.

efficient", and *artificial intelligence weapons*, i.e., those "which make deadly decisions autonomously and independently of human operators. The USA have announced its goal of limiting the use to the first category" (Ibid., pp. 104–105). Kissinger et al. assess this distinction as "wise", while expressing a dose of fear that AI's ability of self-learning "might render insufficient the limitation to certain abilities" (Ibidem). The authors believe that defence "will have to be automated, without conceding the basic elements of human control" (Ibid., p. 106). In that respect, they propose that countries should define "six primary tasks in the control of their arsenals":

1. "Leaders of rival and enemy sides *must* (italics by the author) be ready to regularly speck to one another", as it was done during the Cold War, "about the forms of wars they do not want to wage";
2. Nuclear strategy must be once again be dedicated attention and recognize it as "one of great strategic, technical and moral challenges";
3. "Leading cyber and AI powers need to strive to define their doctrines and borders (even if all their aspects have not been publicly announced) and to identify points of coincidence between their doctrines and doctrines of the rival powers". Of course, the terminology sometimes needs adapting to the "characteristic aspects of cyber intelligence and artificial intelligence";
4. Internal revision of own weapons by the states which possess them, primarily in the domain of "commanding and control, and early warning". This would act preventively against potential cyber-attacks, as well as reduce the danger of "unauthorized, unintentional or accidental use of mass destruction weapons";
5. It is necessary, particularly for technologically most developed countries, to create "robust and accepted methods for maximizing the decision-making time during the period of enhanced tension and in extreme situations... Opponents should particularly try to negotiate a mechanism which will ensure that decisions that may prove to be irrevocable are made at a pace corresponding to human thought and consideration – and survival";
6. The authors suggest that the chief AI powers should consider "how to limit continuous expansion of military artificial intelligence or whether to make systematic efforts in its non-expansion, with the support of diplomacy and the threat by power" (Ibid., pp. 106–107).

The conclusion reached by the authors in this chapter is actually the guiding idea of the entire book. That is what it was created for. Everything else is more or less an analysis, but this is Kissinger's message to the generations of future leaders:

"The will to achieve mutual containment of the most destructive capabilities must not wait for tragedy to strike. While humanity is beginning to compete in creating new, evolving and intelligent weapons, history will not forgive the failure in an attempt of establishing boundaries. In the era of artificial intelligence, permanent search for national advantage must be based on the *ethics of maintaining people* (italics by the author)" (Ibid., p. 108).

In the following chapter Kissinger et al. address the effect of AI on the transformation of human identity. They point to faith and reason as two traditional ways in which people knew the world. Now there is a new way as well: AI, which "understands" what man is unable to understand and which is a non-believer since it has no embedded values. Therefore, it decides by some criteria that are not understandable to people (at least majority of them) or, to put it mildly, unclear. One of the consequences is certainly the transformation of professions. Not only will they not be the same as we know them today, but many today's professions will no longer exist at all, and they will be replaced by some others, the performance of which we cannot even assume today. Does it mean that we will face a new Luddite movement? Perhaps, although for modern man it would much less reasonable than for man more than two centuries ago, when workers in England destroyed their machines, believing that they, and not capitalism, were to blame for their social status. Therefore, not because something by humans would be unreasonable (just as the attempts to prevent development of AI at all costs), but also because it would be inefficient. Man must face the consequences of its emergence. AI will definitely create a large number of unemployed people whose work tasks will be taken over and performed by AI, perhaps more successfully than them. The authors offer society's care for helping people affected by such consequences. Of course, the question is whether such interventionism will please people in the 21st century. On the other hand, we should also ask whether it may be an occasion for Marx's de-alienation. Will such technological change also lead to the change of the social system from global capitalism to the maximum developed by artificial intelligence, in emerging global society towards a different system – new, fairer and better? I will not write down its name because you will remember it. Namely, as Merton used to assert previously, apart from manifest functions (in our case, AI manifest functions), there are also those latent ones, which no one has anticipated. It remains to be seen.

The authors also look at the subcultures such as the Amish. They point out that some societies can simply decide not to apply AI achievements. However, they believe that the omnipresence of AI will be such that with time its use will become inevitable, despite our strong will not to use it.

Observations regarding scientific discoveries are extremely important. "Science has traditionally been the ultimate amalgam of human professionality, intuition and knowledge. In a deep mutual relationship of theory and experiment, human intellect moves all forms of scientific research" (Ibid., p. 115). AI brings something new: "a non-human, different-from-human dimension to scientific research, discoveries and understanding of the world" (Ibidem). What was not written by the authors but only indicted in an earlier passage is that "contextualized information becomes knowledge. When a belief is based on knowledge, it is called wisdom... Only beliefs – in combination with wisdom – enable people to access and research new horizons. *The digital world has little patience for wisdom; its values are determined by the degree of acceptance, and not by thinking. It fundamentally disputes the Enlightenment thesis that reason is the most important element of consciousness...* The digital world does not offer a thesis that connectedness itself is important" (Ibid., p. 33, italics by the author).

This implication unconditionally means that knowledge is unnecessary and, accordingly, that science as "objective, critical, methodologically derived knowledge" is excessive (Marković, 1994b). Why do we need knowledge when it is enough to be networked and to have information? The sea of information in which it is impossible to distinguish the important from the unimportant. Why do we need re-examination and criticality when all of it is done by AI instead of us, as the authors of the book warn?! Even if in the beginning professors and scientists have a role in the development of criticality towards received information, if they make effort to inspire some human values in their students or own children, it will no longer be possible in the following generations: how will those who consider networking instead of knowledge a value be able to consider knowledge and science a certain value? And why would they do it? There is no doubt that science has been gravely brought into question, much more than the authors of this book pointed to or were ready to perceive. And, together with it, our, human paradigm of life. It is also related to the dramatic changes in the sphere of education and upbringing. "AI can serve as a playing mate when a child is bored and as a monitor when the parent is absent. With the introduction of education provided by AI, average human abilities will be put to test or they will increase…. With time, individuals may start preferring their digital assistants to people" (Kissinger et al., 2022, p. 117). However, the authors are right in noticing that "irony is in the fact that even while digitalization makes available an increasing amount of information, it reduces space necessary for deep, concentrated thinking" (Ibid., p. 118).

Answering the question posed about the new human future, the authors believe it is necessary to ensure supremacy of people over AI, but in that respect, they show special care when it comes to what we call *political democracy*. It means ensuring that "key government decisions should be separated from the structures permeated by artificial intelligence and limited to human administration and supervision…; ensuring human supervision and decisive participation in basic elements of power will be of essential significance for maintaining legitimacy…; democracy must preserve human qualities. At the most basic level, it will mean protection of integrity of democratic debates and elections" (Ibid., pp. 120–121). The main fear expressed by the authors is the development of AGI. "*Access to certain powerful artificial intelligence, such as general artificial intelligence, will have to be strictly protected in order to prevent its abuse*" (Ibid., p. 122, italics by the author). This actually means that the USA must keep such monopoly for itself. The USA believes that it will be too expensive and that is why its expansion will be market limited, but it also advocates for reaching international agreement about the limited use of artificial intelligence in some fields, such as the production of biological weapons. In fact, here we can also see Kissinger's Cold War view of the world: reach agreement and consensus between countries, impose their binding quality, naturally with the exception of the USA. However, there is no answer to the question: What will happen if machines begin to communicate on their own?! Is that futurism? Perhaps, but that is exactly what AI used to be about half a century ago.

Towards a conclusion

Henry Kissinger indisputably had a very important role in the shaping of the world in the second half of the 20th century, even the world in which we are still living nowadays, in the third decade of the 21st century. His diplomatic accomplishments brought him plenty of recognition and honour: he pursued both political and academic careers; the education he had and the wisdom he gained gave him the foundation for broad generalizations, sometimes made while even neglecting certain facts, but drawing proper conclusions on the whole. There is no need to speak too much about his moral principles and ethics concerning his politics, since his sole imperative was forever to defend the position of the USA, even when it was opposed to his view of the US national interests (as was the case with the Yugoslav crisis, both when it comes to the USA wanting to preserve Bosnia and Herzegovina as a shole, and when it comes to the 1999 aggression against Serbia and FRY). It seems that he accepted (although he would not have admitted it even if tortured) the Leninist principle of "democratic centralism": "In discussions I defend my own attitude, giving arguments for it consistently and firmly, but when a decision has been made – I will defend it although I was against it previously". The trouble is that in a certain period he was the figure with the greatest influence on decision-makers and that some of those decisions were fatal for many people, as well as for his ethical credibility. On the other hand, some solutions created by him have stood the test of time and he will certainly be remembered because of them.

His co-authored book *The Age of A.I. and Our Human Future* is an extraordinary view of the understanding of *today's degree of AI development*. It is clear that virtually everyone who claims to know something about this technology and wants to stay "in the saddle" of new technological achievements, must read this book very carefully and try to understand it. It is exactly why we have made an unusually detailed overview of the conclusions and attitudes, findings and data contained in this book.

The reader will naturally wonder why we have not written a special review of the book but instead connected the analysis of Kissinger's accomplishments in diplomacy and international relations and Kissinger's insights into today's degree of AI development. There are several reasons for it.

First of all, Kissinger remains remembered as a statesman, even more than as a professor, although he was able to pass much of this theoretical-practical knowledge to his students. That is why this book has been insufficiently perceived in our academic community.

Second, the feature of all Kissinger's efforts was to organize the world order, to create a system which was primary in the interest of the USA. These efforts of his are evident in this book, particularly when he writes about the development of AGI.

Third, his warnings about the potentials of artificial intelligence and warfare with artificially intelligent weapons (particularly because Americans classified and declared their renouncement, *weapons* of artificial intelligence), are quite convincing. No one would survive an "ordinary" nuclear war, let alone the one in which decisions about the use of nuclear missiles are not made

by humans, but are relegated to a generation of artificial intelligence.

Fourth, his requirement for regulating the use of AI in warfare, despite other topics opened up by the book, crucial in it and its guiding idea, is the consequence of his role during the Cold War. Without knowing that role and Kissinger's Cold War logic, it is difficult to understand the recommendations and ideas presented by him in relation to AI.

Fifth and last, a question arises as to whether it is possible to use Cold War means of communication, balance of powers and deterrent effects of nuclear potential to prevent potentially undesirable effects of AI in the 21s century wars. Kissinger et al. offered the solutions based on the experiences of the previous century and on the hypothesis that legitimate democratic decision-making is "reserved" for humans, that it must not be transferred to artificial intelligence. We will see whether they will function in this century as well.

References

BBC NEWS-World Service-In Serbian (2023, July, 20). "America, China and politics: Veteran US diplomat Henry Kissinger in Beijing with Xi Jinping, details of the conversation are not known". BBC NEWS in Serbian. Available at: https://www.bbc.com/serbian/lat/svet-66255601. [In Serbian]

Bisenić, D. (2021, November 14). Kissinger and AI. *Novi Standard.* Available at: https://standard.rs/2021/11/14/kisindzer-i-vestacka-inteligencija/. [In Serbian]

Brzezinski, Z. (2013). *Strategic Vision*, transl. by S. Divjak. Beograd: Albatros plus [In Serbian]

Castells, M. (2012/2018). *Networks of Outrage and Hope*, transl. by V. Savić, M. Nikolić. Beograd: Službeni glasnik. [In Serbian]

Chomsky, N., Waterstone, M. (2022). *Consequences of Capitalism: Manufacturing of Discontent and Resistance*, transl. by P. Milidrag. Novi Sad: Akademska knjiga. [In Serbian]

Debray, R. (2000). *Introduction à la médiologie*, transl. by M. Ivanović. Beograd: CLIO. [In Serbian]

Dimitrijević, B. (2024). Henry A. Kissinger, Eric Schmidt, Daniel Huttenlocher, The Age of A.I. and Our Human Future, Beograd, Klub plus, 2022, 150. *Istorija XX veka* 42 (1), 269–272. Available at: https://istorija20veka.rs/wp-content/uploads/2024/01/2024_1_18_dim_269-272.pdf. [In Serbian]

Kissinger, H. (1994). An Interview. Available at: https://www.youtube.com/watch?v=gnHakmO2_Fs.

Kissinger, H. (1999a). *Diplomacy*, Vol. One. Translated by M. Gligorijević and V. Gligojević. Beograd: Verzalpress. [In Serbian]

Kissinger, H. (1999b, June 28). The Statement. *The Daily Telegraph.*

Kissinger, H. (1999c, April 4). Doing Injury to History. *The Newsweek International.* Available at: https://www.newsweek.com/doing-injury-history-164822 See also an extract in *The Guardian.* Available at: https://www.theguardian.com/world/1999/apr/01/balkans12.

Kissinger, H. (2014, March 5). Opinion: How the Ukraine crisis ends. *The Washington Post*. Available at: https://www.washingtonpost.com/opinions/henry-kissinger-to-settle-the-ukraine-crisis-start-at-the-end/2014/03/05/46dad868-a496-11e3-8466-d34c451760b9_story.html.

Kissinger, H. (2016, December 27). Kissinger advises Trump to accept Crimea as Russia – Bild. Available at: https://web.archive.org/web/20161228103639/http://uatoday.tv/politics/kissinger-advises-trump-to-accept-crimea-as-russia-bild-854458.html.

Kissinger, H.A. Schmidt, E., Huttenlocher, D. (2021/2022). *The Age of A.I. and Our Human Future*, Serbian edition, transl. by J. Novačić and D. Bisenić. Beograd: Klub Plus. [In Serbian]

Mandić, D. P., Miščević, G. M., Bujišić, Lj. G. (2024). Evaluating the quality of responses generated by ChatGPT. *Metodička teorija i praksa* 27 (1), 5–19. DOI: 10.5937/metpra27-51446.

Marković, M. (1994a). *Humanistic sense of social theory.* Beograd: BIGZ, Genes-S štampa, Prosveta, SKZ. [In Serbian]

Marković, M. (1994b). *Philosophical foundations of science.* Beograd: BIGZ, Genes-S štampa, Prosveta, SKZ. [In Serbian]

Orlović, S. (2023). AI (ChatGPT) and our future with it. *Politički život* 24, 95–101. Available at: https://www.fpn.bg.ac.rs/wp-content/uploads/Politicki-zivot-24-KORICE1.pdf. [In Serbian]

RTS (2023). TV "Oko": Henry Kissinger (1923–2023). Available at: https://www.youtube.com/watch?v=nKfJc09UCh4 12`25``–13`57``. [In Serbian]

RTS (2023a). TV "Oko": Henry Kissinger (1923–2023). Available at: https://www.youtube.com/watch?v=nKfJc09UCh4 13`57``–14`17``. [In Serbian]

Simeunović, D. (1992). Coup d'état as a military-political phenomenon. *Vojno delo* 3/92, 119–142. Available at: https://www.vojnodelo.mod.gov.rs/pdf_clanci/vojnodelo285/vd-285-1992-44-2-8-Simeunovic.pdf. [In Serbian]

Šuvaković, U. (2008). Possibilities of Using Internet in Implementation of Scientific Investigation of Public Opinion. *Politička revija* 4/2008, 1523–1546. https://doi.org/10.22182/pr.1842008.24 [In Serbian]

Šuvaković, U. (2023). The War That Broke Out with the Delay of Three Decades. *Balkanske sinteze* 1/2023, 13–31. https://doi.org/10.46630/bs.1.2023.01 [In Serbian]

Vučković, B. (2020). McLuhan and Debray: Theoretical (dis)harmony. *Sociološki pregled* 54 (4), 1311–1328. DOI: 10.5937/socpreg54-29186