MISSION4.0

MISSION4.0

Semi-structured interviews with key stakeholders on socio-ethical aspects of the use of intelligent robotic systems in Serbia

FINAL REPORT

Jelisaveta Petrović

Belgrade, 2022.

Project acronym:	MISSION 4.0
Project full title:	Deep Machine Learning and Swarm Intelligence-based Optimization Algorithms for Control and Scheduling of Cyber-Physical Systems in
Project No:	6523109
Funding scheme:	Science Fund of the Republic of Serbia - Program for Development of Projects in the Field of Artificial Intelligence
Project start date:	01/09/2020
Project end date:	31/08/2022

Abstract	During May and June 2022, a total of 15 semi-structured interviews were conducted with representatives of relevant stakeholders: representatives of companies that use technologies based on artificial intelligence (AI), as well as representatives of scientific and research institutions that participate in the development of AI-based technologies or deal with the social, legal and ethical aspects of the use of these technologies in the business settings in Serbia. In addition to the interviews, a literature review and desk-top research aimed at mapping key actors and AI projects implemented in Serbia were conducted. The research findings are presented in this report alongside the recommendations for introducing AI-based technologies in Serbia ethically and socially justly.

Title of document:	Socio-ethical aspects of the use of intelligent robotic systems in the manufacturing environment
Work package:	3 – The stereo vision-based control algorithm for an intelligent mobile robot used for transportation tasks in a manufacturing environment
Activity:	3.5. Socio-ethical aspects of the use of intelligent robotic systems in the manufacturing environment
Last version date:	30/08/2022
File name:	MISSION40_07_30082022_V01
Number of pages:	27
Dissemination level:	Consortium

VERSIONING AND CONTRIBUTION HISTORY

Version	Date	Revision description	Responsible person(s)
v.1.0	30/08/2022	First draft of the document	Jelisaveta Petrović

1. INTRODUCTION

In his bestselling book *Homo Deus: A Brief History of Tomorrow,* Juval Noah Harari asserts that, due to fast technological development, humanity is on the verge of a new era. This new era is characterised by technological breakthroughs, covering fields such as artificial intelligence, robotics, nanotechnology, quantum computing, etc. These rapid and disruptive changes bring many uncertainties. Among recent technological advancements, artificial intelligence (AI) generates the most public interest but also many controversies. Therefore, we commence this report with some clarifications regarding definitions of AI and respond to some of the most frequent misconceptions.

Artificial Intelligence is the study of the relationship between computation and cognition with the overarching intention of creating or replicating human intelligence (Barr & Feigenbaum, 1982; Kurzweil, 1990). This further implies that AI-based technology should: *act humanly* (e.g. natural language processing, knowledge representation, automated reasoning, and machine learning); *think humanly* (the ability to provide an intelligent solution, as well as the reasoning behind that solution); *think rationally* (the power of logical reasoning), and *act rationally* (the ability to achieve the best outcome, or the closest to it) (see more in Stuart & Norvig, 2003).

European Commission defines artificial intelligence as: "systems that display reasonable, intelligent behaviour by analysing their environment and taking actions — with some degree of autonomy — to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g., voice assistants, image analysis software, search engines, speech and face recognition systems), or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones, or Internet of Things applications)." [1]

[1] High-Level Expert Group on Artificial Intelligence (2018) *A definition of AI: Main capabilities and scientific disciplines*, Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission, 2018. Available at: https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf (accessed 20/08/2022)

Many attempts have been made to assess the intelligence of AI-based technologies. Most famous is the “Turing test”, which asserts that an intelligent machine is one that could trick people into believing they were human. Many present-day AI-based systems have passed the Turing test. However, philosopher John Searle believed that machines could never become truly intelligent. To support this argument, Searle (1980) developed the famous “Chinese room” thought experiment demonstrating that the imitative behaviour of AI was not ‘thinking’ or ‘understanding’ but merely ‘non-consciously’ manipulating symbols. Intelligent decision-making requires actions to be intended with conscious self-awareness (Iphofen & Kritikos, 2019). Despite the fast development of AI since the 1980s, many believe that the Chinese Room Argument showed once and for all that, at best, computers can simulate human cognition.

Today, there is a general distinction between two types of artificial intelligence: narrow artificial intelligence (also called soft AI) and general artificial Intelligence (strong AI). Narrow artificial intelligence is focused on specific tasks in limited fields such as gaming, finances, medicine, etc. (Baum, 2017). On the other hand, general artificial intelligence aims to achieve true human-level intelligence (Goertzel, 2014).

2. RESEARCH DESIGN AND METHODOLOGY

Regarding the development of the concept of Industry 4.0 in the Serbian manufacturing environment, the goal of the MISSION4.0[2] project is to build a systematic methodology for integrating deep machine learning and swarm intelligence-based techniques to achieve adaptable, reconfigurable, and intelligent Cyber-Physical Production Systems. Within Work package 3 (activity 3.5.), sociological research of potential social risks (e.g. loss of jobs, safety concerns, ethical considerations) and benefits (e.g. economic development, employee burden reduction, sustainable development etc.) of introducing AI-based technology in Serbian companies was conducted.

Given that different stakeholders have different perspectives on artificial intelligence and its implementation in business settings, in this research, we started from the assumption that it is essential to bring together alternative frames of thinking - from the various communities of developers, researchers, and business leaders - to properly start acknowledging AI and the readiness of Serbian society for its implementation.

In May and June 2022, semi-structured interviews (N=15) were conducted with stakeholders (developers, company management, and representatives of relevant scientific and research institutions in Serbia). The interviews were conducted with representatives of 2 domestic companies employing AI-based technology – Infostud and Servoteh. Unfortunately, due to complex bureaucratic procedures, foreign-owned companies ZF Serbia and Robert Bosh declined to participate in the research despite lengthy negotiations and initially positive feedback. Interviews were conducted with AI developers from research institutions focused on different AI-based technologies: industrial robotics and AI; application of AI in the legal system; human-machine communication; application of AI in environmental protection; AI in organisational decision-making; AI and cyber security (10 respondents). The interviews were

[2] Full project title: Deep Machine Learning and Swarm Intelligence-based Optimization Algorithms for Control and Scheduling of Cyber-Physical Systems in Industry 4.0

also conducted with social scientists/philosophers exploring the social and ethical implications of using AI (3 respondents).

In addition to the interviews, desk research was conducted alongside a thorough literature review, with the aim of mapping key stakeholders and research projects implemented in the field of artificial intelligence in Serbia.

Based on the research results and complementary material, a set of ethical, socio-economic, and legal recommendations for integrating AI-based technologies in Serbia is presented in the final chapter of this report.

3. RESEARCH FINDINGS

Within the scope of the qualitative research, the following four broad themes were discussed with the interviewees:

1. Development of artificial intelligence in Serbia from a comparative perspective;
2. Institutional and policy framework for the development of AI-based technology in Serbia;
3. Challenges and opportunities for the development of an AI-based economy in Serbia;
4. Ethical considerations in implementing AI-based technologies in Serbia.

In the following sections, we will discuss each of the four distinct topics in more detail. The discussion will be followed by recommendations for the ethically and socially just integration of AI technologies in Serbia.

3.1. Development of artificial intelligence in Serbia from a comparative perspective

The beginnings of modern artificial intelligence can be traced back to the 1940s when the American science fiction writer Isaac Asimov published the short story “Runaround”, where he proposed the famous Three Laws of Robotics. Asimov’s work inspired generations of scientists in robotics, AI, and computer science (Haenlein & Caplan, 2019).

One of the critical milestones in the early days of AI was the work of English polymath Alan Turing, who discussed how to build intelligent machines and test their intelligence in his paper “Computing Machinery and Intelligence” (1950). The next important step was the Logic Theorist program designed by Allen Newell, Cliff Shaw, and Herbert Simon to imitate human problem-solving skills. It’s considered by many to be the first artificial intelligence program. It was presented at the Dartmouth Summer Research Project on Artificial Intelligence (DSRPAI) in 1956. This event was a catalyst for the next twenty years of AI research (Haenlein & Caplan, 2019).

From the Dartmouth conference up until the mid-1970s, AI flourished. Computers became faster, cheaper, and more accessible, and machine learning algorithms improved. However, after the 1973 oil crisis and accompanying economic instabilities, governments of countries that were forerunners in AI development considerably reduced financial support for AI research (Haenlein & Caplan, 2019). Besides the economic crisis, one reason for the cutting of investments and consequent stagnation in the field of AI was the fact that AI technologies were based on logic-based approaches (the 1950s and 1960s) and knowledge-based expert system approaches (during the 1970s and 1980s) that were not able to create general artificial intelligence, as was expected. The recent technological development enabled the use of data-based approaches (from 2000 onwards) and artificial neural networks, considered a paradigmatic shift that allowed the fast growth of AI. In 2012, another breakthrough in machine learning methods occurred, known as *deep learning.*

Today artificial neural networks and deep learning form the basis of most applications of AI (Haenlein & Caplan, 2019).

In Serbia, research in artificial intelligence started to develop in the 1980s, mainly within the research institutions such as the Institute “Mihailo Pupin” (the famous “Belgrade’s school of robotics”), Institute “Jožef Štefan” in Slovenia, and some faculties (e.g. Faculty of Mechanical Engineering (Laboratory for Industrial Robotics and Artificial Intelligence); Faculty of Electrical Engineering in Belgrade). The first university course on industrial robotics (that included lectures on intelligent robot systems) was developed 32 years ago at the Faculty of Mechanical Engineering in Belgrade. As our respondents recall, in those years (before the Internet) in Serbia, it was challenging to access scientific publications (for instance, from MIT). Nevertheless, putting additional effort into their work, Serbian researchers successfully kept pace with the international AI community. In the early 1990s, prof. Srđan Stanković established the first conference, “Neurel”, which focused on the research and development of artificial intelligence.

It is essential to remember that Serbia has a long tradition in robotics. Professors Rajko Tomović and Miodrag Rakić began the critical research at the Laboratory for automatic control at the Institute “Mihailo Pupin” in 1963. One of the most significant achievements of “Belgrade’s School of Robotics” is the "Belgrade Prosthetic Hand", which was used for medical purposes, but also significantly influenced the further development of robotic hands worldwide (Stanić, 2019: 6). Another vital contribution to robotics was made by professor Miomir Vukobratović and his team, developing bipedal movement and exoskeletons. His work on the zero-point movement is recognised as one of the critical contributions to world robotics and is still used today[3]. Since then, in Serbian academic and research institutions, five generations of roboticists have been educated and contributed to the development of robotics and artificial intelligence.

[3] https://www.011info.com/en/get-to-know-belgrade/the-short-and-exciting-history-of-serbian-robotics (accessed 18/08/2022)

Today, the Serbian research community is keeping up with trends in AI and creating very advanced technology. To illustrate this, we will present two projects awarded by the Science Fund of the Republic of Serbia within the Program for Development of Projects in the Field of Artificial Intelligence (2020-2022) (Boxes 1 and 2).

Box 1 Speaker/Style Adaptation for Digital Voice Assistants Based on Image Processing Methods (Acronym: S-ADAPT[4])

Project S-ADAPT uses advanced artificial intelligence methods (machine learning, algorithms based on deep neural networks, especially convolutional neural networks) to improve human-machine speech communication. The goal is to increase the strength of speech recognition in terms of adaptation to the specifics of the speaker and the acoustic environment on the one hand, and to improve speech synthesis, on the other. The research results are verified through their incorporation into the existing digital voice assistant application (Axon Voice Assistant) for Android mobile phones.

In addition to the benefits that this technology has for persons with disabilities, as significant contributions of this project, the following should be mentioned: preservation of the Serbian language in the digital era; independence from technological giants (e.g. Google); and a possibility of using "in-house" solutions for public institutions and private entities that wish to avoid uploading sensitive data on cloud platforms.

[4] The project leader is professor Vlado Delić (Faculty of Technical Sciences, University of Novi Sad). The Faculty for Technical Sciences is the main scientific research organization (SRO) and the partner SRO is Mathematical Institute, Serbian Academy of Sciences and Arts, Belgrade. More information about this project can be found on the official website https://www.ktios.ftn.uns.ac.rs/sadapt/SADAPT_sr.html (accessed 20/08/2022)

Box 2 Deep Machine Learning and Swarm Intelligence-based Optimization Algorithms for Control and Scheduling of Cyber-Physical Systems in Industry 4.0 (Acronym: MISSION 4.0)[5]

The Fourth industrial revolution, which implies complete digitisation and automation of production processes, is in an advanced stage. One segment, however, is not yet sufficiently developed, neither in our country nor in the world, and that is autonomous internal transport based on intelligent mobile robots. The Mission 4.0 project is aimed precisely at the development of such systems.

Two intelligent mobile robots are developed in the Laboratory for Industrial Robotics and Artificial Intelligence of the Faculty of Mechanical Engineering in Belgrade. With the use and analysis of the information they receive from sensors, primarily two cameras, RAICO (Robot with Artificial Intelligence based COgnition) and DOMINO (Deep learning based Omnidirectional Mobile robot with INtelligent control), can determine at what distance is the object that they need to approach and manipulate, and which obstacles need to be avoided during autonomous movement. Two laboratory prototypes can successfully communicate, prevent cyber-attacks, avoid each other while performing robotic tasks, and solve "conflicts". Based on laboratory prototypes, industrial robots will be designed and implemented in domestic industrial production.

As stated in the Strategy for the Development of Artificial Intelligence in the Republic of Serbia for the period 2020-2025: “increasing multidisciplinarity about classical divisions of professions and areas of competence, and in particular the linking of knowledge and skills in the socio-humanistic and artistic fields with the knowledge and skills in the fields of natural sciences and mathematics as well as technical and technological sciences is seen as important” (Strategy, 2019: 24[6]). However, despite policy support, it seems that

[5] The project leader is professor Zoran Miljković (Faculty of Machine Engineering, University of Belgrade). The Faculty of Machine Engineering is the principal scientific research organisation and the partner is the Faculty of Philosophy, University of Belgrade. More information about this project can be found on the official website http://mission4-0.mas.bg.ac.rs (accessed 20/08/2022)

[6] https://www.srbija.gov.rs/tekst/en/149169/strategy-for-the-development-of-artificial-intelligence-in-the-republic-of-serbia-for-the-period-2020-2025.php (accessed 20/08/2022)

interdisciplinarity in Serbia is in its early stages. To illustrate this, we shall look at the twelve AI projects that have received two-year grants from the Science Fund of the Republic of Serbia. Only three projects engage researchers from humanities and social sciences, even though some deal with AI's application in data privacy, language processing and education.

3.2. Institutional and Policy Framework for Development of AI-based technology in Serbia

The policy initiatives to develop information society in Serbia began with its accession to the Electronic South Eastern Europe Initiative in 2000. Since then, policy and institutional frameworks for digitalising different social spheres (telecommunication and media, economy, government, education, environmental protection, etc.) have constantly been improving.

Regulation and planning of the development of AI-based technologies have become interesting for policy-makers in the last few years in Serbia, in parallel with the advances in this sphere on the European level. The policy framework for developing and applying AI in Serbia is relatively developed. According to the Government AI Readiness Index for 2021[7], Serbia scored 55.98 and is in 52nd place among 160 countries. Regionally, Serbia is well positioned as it is best ranked among the countries of Southeast Europe.

In 2019, the Serbian government adopted the *Strategy for the Development of Artificial Intelligence in the Republic of Serbia for 2020-2025*. The Strategy is in line with the European Artificial Intelligence Initiative and other relevant policy and legal documents. As stated in the Strategy: "The general objective of the strategy is the use of artificial intelligence in favour of economic growth, employment and improvement of the quality of life." (Strategy, 2019: 23). The Strategy has also defined five specific objectives: "1. Development of education geared to the needs of modern society and economy conditioned by the advancement of artificial intelligence; 2. Development of science and innovation in artificial intelligence and its implementation; 3. Development of the economy based on artificial intelligence (where this is a key competence and where it is used in different industrial branches); 4. Improvement of assumptions for developing artificial intelligence and public sector services through implementing artificial intelligence; 5. Ethical and safe application of artificial intelligence"(Strategy, 2019: 23).

[7] https://static1.squarespace.com/static/58b2e92c1e5b6c828058484e/t/61e95661c567937d21998d14/1642681965033/Gov_AI_Readiness_2021.pdf (accessed 20/08/2022)

Other essential strategies and laws governing development and implementation of AI are the following: The Strategy for the Development of Education in Serbia by 2020 ("Official Gazette of the RS", No. 107/12); Scientific and Technological Development Strategy of the Republic of Serbia for the period 2016–2020 – Research for Innovation; Strategy for the Development of an Information Society in the Republic of Serbia by 2020; Strategy for the Development of Information Security for the period 2017-2020; Strategy for the Development of the Information Technology Industry for the period 2017-2020; Strategy for Development of New Generation Networks by 2023; The Law on Personal Data Protection ("Official Gazette of the RS", No. 87/18); The Law on the Fundamentals of the Education System ("Official Gazette of the RS", No. 88/17, 27/18, 27/18, 10/19); The Law on Primary Education ("Official Gazette of the RS" No. 55/13, 101/17, 27/18, 10/19); The Law on Secondary Education ("Official Gazette of the RS" No. 55/13, 101/17, 27/18); The Law on Higher Education ("Official Gazette of the RS", No. 88/17, 27/18, 73/18, 67/19); The Law on Science and Research ("Official Gazette of the RS", No. 49/2019-3); The Law on Innovation Activity ("Official Gazette of the RS", No. 110/05, 18/10, 55/13); The Law on the Science Fund of the Republic of Serbia ("Official Gazette of the RS", No. 95/2018-353); The Law on Electronic Government ("Official Gazette of the RS", No. 27/18); The Law on Electronic Communications ("Official Gazette of the RS", No. 44/10, 60/13 (Constitutional Court), 62/14, 95/18); The Law on Information Security ("Official Gazette of the RS", No. 6/16, 94/17, 77/19) (See more in Strategy, 2019).

The Government of Serbia has established the Institute for Artificial Intelligence Research and Development of Serbia[8] based on the initiative from the AI Strategy. The Institute was formally founded in 2021 and is headquartered in Science and Technology Park in Novi Sad. The Institute is envisioned as an incubator for new ideas and educational place providing young experts with necessary resources. The Government of Serbia also established the Council for Artificial Intelligence, which, in addition to the Institute for Artificial Intelligence, is the responsible institution for monitoring and implementing the Strategy (Radun, 2019).

[8] https://www.ivi.ac.rs (accessed 21/08/2022)

In 2020, the Science Fond of the Republic of Serbia granted 12 projects (6 for fundamental and 6 for applied research in the field of artificial intelligence) with a total budget of 2,207,774 euros. The Program for Development of Projects in the Field of Artificial Intelligence was created with the aim to: "enhance excellence and relevance of the scientific research in the domain of artificial intelligence in the Republic of Serbia, as well as to support the implementation of the scientific results in the economic development of the Republic of Serbia, enhancement of human resources development and improvement of international development in the field of artificial intelligence"[9]. The following projects have received funding: 1. AI4TrustBC – Advanced artificial intelligence techniques for analysis and design of system components based on trustworthy blockchain technology; 2. ATLAS – Artificial intelligence theoretical foundations for advanced spatial-temporal modelling of data and processing; 3. AVANTES – Advancing novel textual similarity-based solutions in software development; 4. Com-in-AI – Advanced methods of quantization, compression and learning in artificial intelligence; 5. DECIDE – Decentralized machine learning control for intelligent multi-agent dynamical systems; 6. GRASP – Graphs in space: graph embeddings for machine learning on complex data; 7. AI4WorkplaceSafety – Artificial intelligence for managing workplace safety; 8. ARTEMIS – Artificial intelligence in energy management innovative services; 9. CERES – Eo-based information for "smarter" agriculture and carbon farming; 10. Clean CaDET – Clean code and design educational tool; 11. MISSION4.0 – Deep machine learning and swarm intelligence-based optimization algorithms for control and scheduling of cyber-physical systems in industry 4.0; 12. S-ADAPT – Speaker/style adaptation for digital voice assistants based on image processing methods.

To summarise, policy and institutional infrastructure for developing and implementing AI-based technologies in Serbia are growing in line with the EU regulations and policies. Since the application of these technologies is in an early phase in Serbia, the limitations of the existing infrastructure will be more visible in the future. However, early preparations for exploiting this innovative and disruptive technology are critical. Moreover, scientific projects

[9] http://fondzanauku.gov.rs/poziv/2020/01/program-for-development-of-projects-in-the-field-of-artificial-intelligence/?lang=en (accessed 22/08/2022)

in artificial intelligence gained a certain level of state support, and this good practice must be maintained in the future. Besides the amount of funding, continuity and predictability of funding are of great significance for adequate planning of future research.

3.3. Challenges and opportunities for the development of an AI-based economy in Serbia

Digitisation and automation are expected to fundamentally change the nature of work, business, and society in the coming years (Brynjolfsson & McAfee, 2014; Ford, 2015). Since 1962, when the first industrial robot was employed on a General Motors assembly line in New Jersey, the main idea of robotics was to replace "dull, dirty and dangerous" jobs (Takayama et al., 2008). However, some authors argue that the key difference from previous industrial revolutions is that AI-based technology aims to replace cognitive work (Brynjolfsson & McAfee, 2014).

Popular narratives regarding AI, robotisation and automation usually revolve around potentially adverse effects on jobs and employment. These narratives span from predicting a "jobless society" to expectations of labour market polarisation (Ford, 2015; Frey, Berger, & Chen, 2018).

Various methodologies have been used to assess the effects of automation on occupations and corresponding employment, with varying results and conclusions. For instance, focusing on entire professions, Frey and Osborne (2017) estimated that about 47% of total U.S. employment is at risk of being automated. On the other hand, starting from the assumption that automation usually affects specific tasks but not entire professions, Arntz and colleagues (Arntz et al. 2017) estimated that only 6-12% of people in developed countries perform jobs that are at high risk of automation.

Despite different estimations, labour economists agree that structural changes in the labour market are inevitable. However, mass unemployment is unlikely to happen in the next few decades (Arntz et al., 2017; Autor, 2015). What can be expected is an increase in job polarization[10], on the one side, and the emergence of new professions, on the other (Autor &

[10] This means that the introduction of AI-based technologies is reducing demand for routine and manual tasks while increasing demand for low- and high-skilled tasks and for problem-solving and interpersonal skills.

Dorn, 2013). Moreover, some studies show significant economic benefits of introducing AI-based technologies due to the increased productivity and reduced employee burden (OECD, 2017; European Commission, 2018).

Central and East European countries (CEE) lag behind more developed countries in implementing AI-based technology. Nevertheless, technological changes will affect the job market in these countries shortly (Cséfalvay, 2020).

Due to the high share of manufacturing jobs in total employment in Central and Eastern Europe, these countries are classified among those economies where jobs are at the highest risk of being automated. Similar to developed economies, the estimations vary significantly from the calculated 49-69% of workers threatened by automatisation (Nedelkoska and Quintini, 2018) to only 7-11% (Arntz et al., 2017).

One of the factors influencing a slower pace of work automation in CEE is labour costs. Since high labour costs boost robotisation, and lower wages have the opposite effect, the rate of industrial robot adoption is much weaker in CEE with low salaries and labour-intensive industries compared to Western Europe (Cséfalvay, 2020).

To illustrate the obstacles to implementing new technologies in the Serbian industry, we can use an example from one of our interlocutors. This is the case of a factory from Jagodina, which received a robotic production line of 13 robots through the project funding. However, immediately before the commissioning of this plant, the state launched a campaign for subsidised employment of workers. The factory management decided to sell the product line and hire new workers. In this way, they made money from the sale of technology and did not have additional expenses because the state entirely covered the labour costs. As seen in this example, the employment policy in Serbia is not well aligned with procedures regulating technological innovation.

In CEE countries, robot deployment is mainly concentrated in the automotive industry run by multinational companies (Cséfalvay, 2020). The situation is similar in Serbia. Our research

shows that the connection of the foreign-led automotive industry with domestic research institutions and universities is relatively weak, with significant difficulties in cooperation mainly due to complex bureaucratic procedures in those companies. In the Serbian industry sector, the application of intelligent robots is still not present. This technology is considered too advanced; however, in the next 5 to 10 years, it will be a standard part of industrial processes in our country.

3.4. Ethical considerations in implementing AI-based technologies in Serbia

The views regarding the impact of AI on human life span from technological optimism to apocalyptic predictions (Madridakis, 2017). Nevertheless, in much of the developed world, algorithmic decision-making systems already govern various aspects of people's lives (e.g. healthcare, finances and employment), and it is expected that the importance of AI will continue to grow in the future (Noble, 2018).

When considering the ethical development of AI-based technologies, it is crucial to consider the nature of their underpinning algorithms since automated decision-making can result in harmful discrimination. The lack of transparency around algorithms is an important ethical issue (Iphofen & Kritikos, 2019). Among AI researchers and developers, the general understanding of automated systems is that, if not carefully designed, they tend to reproduce social biases. This is particularly the case for machine learning algorithms, which are trained based on datasets that inevitably reflect human judgments, priorities, and conceptual categories (Crawford, 2021).

Many worldwide initiatives are working toward enabling responsible and ethical development and employment of AI-based technologies (Iphofen & Kritikos, 2019). One of the forerunners in regulating this domain is undoubtedly the EU. At the EU level, the General Data Protection Act (GDPR) was made in 2016 and implemented in 2018. The GDPR is the first legislation explicitly addressing algorithmic discrimination (Iphofen & Kritikos, 2019). In 2018, Serbia adopted the new Law on Personal Data Protection, which complies with the GDPR.

In 2019, the High-Level Expert Group on AI produced the *Ethics Guidelines for Trustworthy AI*.[11] The Ethics Guidelines introduced the concept of Trustworthy AI based on seven essential requirements: 1. human agency and oversight; 2. technical robustness and safety; 3. privacy

[11] https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (accessed 22/08/2022)

and data governance; 4. transparency; 5. diversity, non-discrimination and fairness; 6. environmental and societal well-being; 7. Accountability[12].

In 2020, the Guidelines were supplemented with an *Assessment List for Trustworthy Artificial Intelligence* (ALTAI) that guides developers and deployers of AI in implementing such principles in practice[13].

For now, Serbian research institutions and companies developing and implementing AI-based technologies mainly rely on the Law on Personal Data Protection. Other EU regulations and checklists are usually considered about AI's future development and application. The respondents from research institutions state that rigorous legal and ethical standards guide their work. Most of the respondents stressed that they are concerned about their technology's ethical application. They are developing new technologies having humanistic values and society's best interest as the top priority. Regarding respondents from the business sector, compliance with the Law on Personal Data Protection is emphasised.

[12] https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment (accessed 22/08/2022)
https://eucrim.eu/news/ai-high-level-expert-group-publishes-ethics-checklist/ (accessed 22/08/2022)
[13] https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment (accessed 22/08/2022)
https://eucrim.eu/news/ai-high-level-expert-group-publishes-ethics-checklist/ (accessed 22/08/2022)

4. CONCLUSIONS AND RECOMMENDATIONS

The implementation of AI-based technologies is in the early stages in Serbia. However, given the global trends, these technologies' fast development and cross-sectoral implementation can be expected in Serbia shortly, meaning that all actors should be well prepared. One of the first steps in preparation is developing the legal framework and state policies regulating the implementation of AI-based technologies and protecting citizens' rights. Serbia has made progress by implementing the AI Strategy and related policies and laws. However, this is only the beginning and regulation in this area will have to develop in parallel with new technological advancements.

Serbia lags behind more developed countries in implementing AI and introducing the concept of Industry 4.0. However, this also could be regarded as an advantage. Namely, in the most developed countries, the ideas of Industry 5.0 and Society 5.0 emerged to tackle the issues brought by Industry 4.0 that had significant drawbacks. While Industry 4.0 focuses on transforming factories into AI-supported infrastructures, Industry 5.0 focuses on the humanistic aspect. The Industry / Society 5.0 initiative's main principles include preserving environmental, economic, and social ecosystems by applying advanced technologies (Ćoćkalo et al., 2021). Starting from these principles and the ethical recommendations for Trustworthy AI, we propose the following recommendations for the introduction of AI-based technologies in Serbia:

1. Fears and popular misconceptions about AI should be addressed to reduce technological anxiety among the general population. This can be done by informing citizens of the changes that can be expected in the future and available coping mechanisms and strategies. Since there is no available data, this action should be preceded by thorough research on public attitudes regarding AI.
2. Technological transitions inevitably create winners and losers. Given that AI-based social transformation is inevitable, a policy framework should be implemented to appropriately buffer changes in the job market and prepare workers and society for

the expected changes through the education system, lifelong learning, prequalification, etc.

3. Timely preparation of all relevant actors and state institutions for expected market transformation, for creating and implementing measures aimed at avoiding/reducing job polarisation and unemployment is considered very important.
4. Companies employing AI should seriously consider their impact on local communities and the environment. They should work to respect ethical standards and sustainable development goals.
5. Scientific and research institutions should be further supported to work on AI projects. Interdisciplinary work should be encouraged. Cooperation with the business sector should be stimulated but should not follow the logic of profit at the expense of the public interest.
6. By the Guidelines for Trustworthy AI, the development, deployment and use of AI systems should adhere to the following ethical principles: respect for human autonomy, prevention of harm, fairness and explicability. Particular attention should be devoted to situations that involve vulnerable groups such as children, persons with disabilities etc.
7. Deployment and use of AI systems should meet the seven critical requirements for Trustworthy AI: (1) human agency and oversight, (2) technical robustness and safety, (3) privacy and data governance, (4) transparency, (5) diversity, non-discrimination and fairness, (6) environmental and societal well-being and (7) accountability.

6. REFERENCES

- Arntz, M., Gregory, T., Zierahn, U. (2017) Revisiting the Risk of Automation, *Economics Letters*, Vol. 159.
- Autor, D. H. (2015) Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives*, 29(3): 3–30.
- Autor, D. H., Dorn, D. (2013) The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5): 1553–1597.
- Barr, A. & Feigenbaum, E. (1982) *The Handbook of Artificial Intelligence* (Volume 2). Stanford, Calif.: HeurisTech Press.
- Baum, S. (2017) A Survey of Artificial General Intelligence, *Projects for Ethics, Risk, and Policy Global Catastrophic Risk Institute Working Paper*, 17(1). Available at SSRN: https://ssrn.com/abstract=3070741 or http://dx.doi.org/10.2139/ssrn.3070741
- Brynjolfsson, E., & McAfee, A. (2014) *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*, New York, NY: Norton.
- Ćoćkalo, D., Bakator, M., Đorđević, D., Vorkapić, M., Stanisavljev, S. (2021) *Industry 5.0: A New Paradigm in Manufacturing, Industry 5.0: A New Paradigm in Manufacturing*, PaKSoM 2021, 3rd Virtual International Conference Path to a Knowledge Society-Managing Risks and Innovation Proceedings, pp. 245-250.
- Crawford, K. (2021) *Atlas of AI,* New Haven: Yale University Press.
- Cséfalvay, Z. (2020) Robotization in Central and Eastern Europe: catching up or dependence?, *European Planning Studies*, 28(8): 1534-1553.
- European Commission (2018) Employment and social developments in Europe. *Annual review 2018*. Directorate-General for Employment, Social Affairs and Inclusion, Publications Office of the European Union, Luxembourg. Retrieved from https://ec.europa.eu/social/BlobServlet? docId=19719&langId=en (20/08/2022)
- Ford, M. (2015) *Rise of the robots: Technology and the threat of a jobless future*. New York, NY: Basic Books.

- Frey, C. and M.A. Osborne (2017) The Future of Employment: How Susceptible are Jobs to Computerization? *Technological Forecasting and Social Change*, 114: 254–280.
- Frey, C. B., Berger, T., & Chen, C. (2018) Political machinery: Did robots swing the 2016 US presidential election? *Oxford Review of Economic Policy*, 34(3): 418–442.
- Goertzel, T. (2014) The path to more general Artificial Intelligence, *Journal of Experimental & Theoretical Artificial Intelligence,* 26(3): 343-354.
- Haenlein, M., Kaplan, A. (2019) A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence, *California Management Review,* 1–10.
- Iphofen, R., & Kritikos, M. (2019) Regulating artificial intelligence and robotics: ethics by design in a digital society, *Contemporary Social Science*, 1–15.
- Kelly Joyce, K, Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T. et al. (2021) Toward a Sociology of Artificial Intelligence: A Call for Research on Inequalities and Structural Change, *Socius: Sociological Research for a Dynamic World*, 7: 1–11.
- Kurzweil, R. (1990) *The Age of Intelligent Machines*, Massachusetts: MIT Press.
- Madridakis, S. (2017) The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms, *Futures,* 90: 46–60.
- Nedelkoska, L., & Quintini, G. (2018) Automation, Skills Use and Training. OECD Social, *Employment and Migration Working Papers* 202, OECD Publishing, Paris.
- Noble, S. U. (2018) *Algorithms of oppression: How search engines reinforce racism,* New York: NYU Press.
- OECD (2017) *The next production revolution: Implications for governments and business*, Paris: OECD Publishing.
- Radun, V. (2019) Mogućnosti i izazovi veštačke inteligencije u transformaciji ekonomije i društva u Srbiji, *Limes* 3: 55-76.
- Searle, J. (1980) Minds, Brains and Programs, *Behavioral and Brain Sciences*, 3: 417–57.
- Searle, J. (1999) The Chinese Room, in R.A. Wilson and F. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences*, Cambridge, MA: MIT Press.

- Stanić, M. (2019) *The birth of humanoid robotics*, 2nd Edition, Belgrade: Museum of Science and Technology.
- Stuart J. R. & Norvig, P. (2003) *Artificial Intelligence: A Modern Approach* (3rd. ed.), London: Pearson Education.
- Takayama L, Ju W, Nass C (2008) Beyond dirty, dangerous and dull: what everyday people think robots should do. In HRI 2008-Proceedings of the 3rd ACM/ IEEE *International Conference on Human-Robot Interaction: Living with Robots*. https://doi.org/10.1145/1349822.1349827.