06

СРПСКА ПОЛИТИЧКА МИСАО

SERBIAN POLITICAL THOUGHT

GUEST EDITOR
Ivana Luknar

AI AND POLICY
Ivana Damnjanović, Marko Pejković, Serdar Selim, İsmail Yılmaz, Pınar Kınıklı, Vanja Glišin, Marija Tasić, Aleksandar Ivanović, Tanja Kesić, Ivana Bjelovuk, Jadranka Otašević, Mira Jovanovska, Ljubiša Bojić, Mirjana Dejanović, Anatolijs Krivins, Marija S. Dokić

ESSAYS AND STUDIES
Danijela Nikolić, Violeta Rašković Talović, Slobodan Popović, Dragan Trailović

REVIEWS
Željko Oset

ISSN 0354-5989 UDC 32 year XXXII vol. 94 No. 6/2025.

СРПСКА ПОЛИТИЧКА МИСАО
SERBIAN POLITICAL THOUGHT

ISSN 0354-5989
UDC 32
No. 6/2025.

СРПСКА ПОЛИТИЧКА МИСАО
SERBIAN POLITICAL THOUGHT
Publisher: Institute for Political Studies, Belgrade. Tel: +381 11 3349-204
E-mail: ips@ips.ac.rs; spm@ips.ac.rs
https://www.ips.ac.rswww.ips.ac.rs/rs/magazines/srpska-politicka-misao
ISSN 0354-5989 UDC 32 No 6/2025. XXXII vol. 94

Editor-in-Chief:
Miša Stojadinović, Ph.D., Principal Research Fellow, Institute for Political Studies

Deputy Editor-in-Chief:
Nada Raduški, Ph.D., Principal Research Fellow, Institute for Political Studies

Copy editors:
Olga Stevanović, Ph.D., Research Assistant, Institute for Political Studies,
MA Andrea Matijević, Research Assistant, Institute for Political Studies

Members of the Editorial Board:
Milan Jovanović, Ph.D., Full professor, Faculty for Political Sciences, University of Belgrade,
Dejan Đurđević, Ph.D., Full professor, Faculty of Law, University of Belgrade, Živojin Đurić, Ph.D.,
Jasna Milošević Đorđević, Ph.D., Full professor, Faculty of Media and Communications,
Đuro Bodrožić, Ph.D., Senior Research Associate, Institute for Political Studies, Miša Stojadinović,
Ph.D., Principal Research Fellow, Institute for Political Studies, Kirill Shevchenko, Ph.D., Professor,
Russian State University of Social Sciences, Minsk, Višnja Stojadinović, Ph.D., Research Associate,
Institute for Political Studies, Andrei Baranov, Ph.D., Professor, Kuban State University, Krasnodar,
Mamoru Sadakata, Ph.D., Professor, Graduate School of Law, Nagoya University, Japan,
Ewa Bujwid-Kurek, Ph.D., Professor, Institute for Political Science and International Relations
of the Jagiellonian University, Krakow, Poland, Mikhail Zalihanov, Ph.D., Professor,
member of the Russian Academy of Science

Editorial staff
Stefan Radojković, Ph. D., Research Associate, Institute for Political Studies,
MA Jelena Vujanović, Research Assistant, Institute for Political Studies,
MA Anastasija Orlić, Junior Research Assistant, Institute for Political Studies,
MA Veronika Vasilić, Junior Research Assistant, Institute for Political Studies,
MA Ivana Vlašković, Junior Research Assistant, Institute for Political Studies

Members of the Advisory Board
Zoran Avramović, Ph.D., Full professor, Faculty of Philosophy, University of Belgrade,
Sreten Sokić, Ph.D., retired Full Professor, Faculty for Political Sciences, University of Belgrade,
Milovan Mitrović, Ph.D., retired Full Professor, Faculty of Law, University of Belgrade,
Milomir Stepić, Ph.D., Principal Research Fellow, Institute for Political Studies,
Dragan Simeunović, Ph.D., retired Full professor, Faculty for Political Sciences, University of Belgrade,
Milan Brdar, Ph.D., Full professor, Faculty of Philosophy, University of Novi Sad,
Zoran Stojiljković, Ph.D., Full professor, Faculty for Political Sciences, University of Belgrade,
Dragana Mitrović, Ph.D., Full professor, Faculty for Political Sciences, University of Belgrade,
Ljubiša Despotović, Ph.D., Full professor, Faculty for European Legal and Political Studies,
University of Novi Sad

International Members of the Advisory Board
Simon James Critchley, Ph.D., Professor, New School for Social Research,
Anastasia Mitrofanova, Ph.D., Professor, Diplomatic Academy of the Russian Federation, Moscow,
Ivan Alexeyevich Charota, Ph.D., member of National Academy of Sciences of Belarus;
member of Russian Academy of Sciences, member of Serbian Academy of Sciences and Arts

Proof-reading
Anja Maljević

JOURNAL IS PUBLISHED SIX TIMES A YEAR.
Articles SPT are available in electronic databases: SCIndeks, C.E.E.O.L.
(Central and Eastern European Online Library),
ERIH PLUS (European Reference Index for the Humanities and Social Sciences),
and Russian Science Citation Index (RSCI).

Printing:
Donat Graf, Belgrade
Circulation: 50

Reproduction of articles published in this journal, whether in whole or in part, is not permitted without the explicit consent of the publisher. The views expressed in the articles represent the personal opinions of the authors and do not reflect the positions of the Editorial board or the institutions the authors are affiliated with.

SERBIAN POLITICAL THOUGHT

CONTENTS

FOREWORD

ARTIFICIAL INTELLIGENCE AND POLICY

Modern society and socio-political processes are shaped by technological development. Artificial intelligence (AI) is a technology that not only intrigues the global public but also has an incredible potential to drive profound changes in today's society. The changes brought about by the application of AI technology can be both gradual and sudden. They first begin with the transformation of economic flows and security strategies, followed by changes in cultural patterns (including values, norms, and moral principles), and then by shifts in social relations. This modern technology is clearly transforming, to a significant extent, the way we understand society, politics, and even humanity itself.

The relationship between politics and technology is complex and characterized by mutual influence and interdependence. Technological development increasingly affects political decision-making, while at the same time, politics proposes and determines legal regulations, frameworks, and ethical standards that guide the further course of development and application of AI systems. The interrelation between politics and technology increasingly determines not only the direction of future technological development but also the future of democracy, sovereignty, and accountability. Artificial intelligence technology can enhance the transparency and efficiency of political processes when applied appropriately, in accordance with the law, and in line with recommended ethical guidelines. Awareness of the multiple mutual influences between artificial intelligence technology and politics is a necessary precondition for preserving democratic principles and protecting human rights. That is precisely why this issue of the journal is specially dedicated to examining the complex relationships between artificial intelligence and politics. Understanding this relationship can help prevent the misuse of AI and its manipulation of public opinion.

In the introductory article, Prof. Dr. Damnjanović raises the question of whether the advancement of artificial intelligence necessarily leads to the strengthening of democratic institutions or, conversely, to their deterioration. Referring to Robert Dahl's theory of democracy, the author shows that a tension often arises between technological efficiency and democratic transparency, which will shape the political trends of the 21st century, particularly regarding the principles of equality, personal

autonomy, and autonomous choice, as well as the functioning of the public sphere. To reverse the trend of negative implications of AI on democracy, Damnjanović proposes restructuring the development of artificial intelligence based on democratic debate and participatory design. Then follows the article of Dr. Pejković, which provides a critical review of algorithmic governance of the economy and points out the impossibility of artificial intelligence ever enabling (centralized) efficient economic planning at the level of a national economy in the form of a state plan. This argument is supported by a debate and a mathematical paradox, which he sees as a conceptual obstacle to economic planning assisted by this technology. From the perspective of global politics, artificial intelligence technology is emerging as a new instrument of power and a factor of instability in international relations, where world powers are competing for control over data, algorithms, and predictive capabilities. Referring to a type of race in the development and application of new technologies, Dr. Glishin analyzes the characteristics and differences between the security strategies of major powers, such as the United States and China, to identify the role of AI in their security strategies and plans for its further development in the global power struggle.

An exceptional contribution to this issue is made by the papers that explore the application of AI in policies aimed at achieving the common good. In their article, Prof. Dr. Selim and his colleagues highlight the successful application of artificial intelligence in tree detection, which serves as a good example of how advanced machine learning models can be utilized in the fight against climate change. This aligns with the climate policy within the European Green Deal. The following article explores the potential applications of artificial intelligence in combating and preventing crime. Dr. Tasić and the other co-authors highlight the potential of AI, as well as the controversies arising from the use of existing AI systems by law enforcement agencies. Subsequently, the article of Dr. Otašević and Dr. Jovanovska, who paid special attention to artificial intelligence as a means of supporting independent living for persons with disabilities, followed. They emphasize the social utility of artificial intelligence as a tool for increasing the inclusion and quality of life for persons with disabilities.

In addition to its social utility, the authors highlight the risks associated with the application of this technology. Dr. Bojić highlights the existential, social, ethical, and governance challenges that AI could

bring to humanity, based on a qualitative thematic analysis of extended interviews with Geoffrey Hinton, Yoshua Bengio, and Yann LeCun. Due to the continuous evolution of risks, Bojić proposes that future research be expanded to include multinational perspectives and all relevant stakeholders, to test the proposed regulatory frameworks in real-world contexts and across all domains of AI application. The legal and institutional dimensions of AI are addressed in the article of Dr. Dejanović and Dr. Krivins, who highlight the different approaches states take in shaping policies aimed at enabling the safe and sustainable application of AI. This is illustrated through a comparative analysis of regulatory frameworks related to artificial intelligence in the Republic of Serbia, Latvia, and the European Union (EU). The debate on the topic of artificial intelligence and policy concludes with a philosophical reflection by Dr. Dokić, which interprets the effects of artificial intelligence from the perspective of Hegel's interpretation of the concepts of freedom, historical necessity, and the sovereignty of reason. The author views this technology as a fragment of a complex puzzle, given the constantly changing reality, that is, as a political phenomenon challenging contemporary understandings of autonomy and responsibility, and raising questions about the human relationship with technology.

The first part of the journal is conceptualized as a thematic issue. It seeks to open a space for dialogue between various disciplines (political science, law, economics, technology, criminology, sociology, philosophy, etc.). This is because only through such a dialogue is it possible to understand the scale and consequences of the technological revolution shaping our present. Artificial intelligence is no longer just a matter of the future; it has become a mirror of contemporary society and a delicate challenge for today's political thought.

The papers published in the second part of the issue also explore significant scientific topics, providing valuable and innovative contributions to the scholarly community. This includes an article by Dr. Nikolić and Prof. Dr. Rašković Talović, which addresses Azerbaijan's multi-vector foreign policy, as well as an article by Dr. Popović and Dr. Trailović that interprets the encounter between the Global Civilization Initiative and practice, using the Shanghai Cooperation Organisation as an example. Finally, there is a review of a pilot project using a case study of Slovenia by Dr. Oset.

The publication of this issue of the journal is, first and foremost, the result of the invaluable scientific contributions of our authors, who

have decisively enhanced the scientific quality of this volume. It is also the outcome of the equal contribution of the entire Editorial team, without whose support, expertise, and dedication, the publication process could not have been realized. I would like to express my gratitude and pay tribute to all those whose hard work and commitment have made it possible for this edition to see the light of day.

Belgrade, November 2025
Guest Editor of this issue of the *Serbian Political Thought* journal
Associate Professor, Dr. Ivana Luknar, Research Associate

This Issue's Theme

AI AND POLICY

UDC 321.7:004.8(4-672EU)
DOI: 10.5937/spm94-61498
Original research article

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 1-23

Ivana Damnjanović*

Faculty of Political Science, University of Belgrade

AI *VS.* DEMOCRACY: PRESENT AND POSSIBLE FUTURES**

Abstract

The aim of this paper is to explore and systematize current research on the relationship between artificial intelligence (AI) and democracy, primarily focusing on existing AI applications, in order to understand the general avenues of exploration conceived within the field of political science on this topic, and to evaluate their results so far. However, some thought will also be given to the more far-fetched possibilities of AI development and their implications for democracy. Using Robert Dahl's theory of democracy, the paper outlines why the AI is perceived as a threat to liberal democracy, especially in regard to principles of equality, personal autonomy, and autonomous choice, and the functioning of the public sphere. Finally, the paper considers some possible, if not very probable, future developments in AI, namely the inception of artificial general intelligence (AGI) and artificial superintelligence (ASI), and their implications for democracy. While existing AI applications do seem to erode certain preconditions for functioning democracy, it is nevertheless possible to overturn this trend by restructuring AI development based on the democratic debate and participatory design.

* E-mail: ivana.damnjanovic@fpn.bg.ac.rs; ORCID: 0000-0002-8553-492X

** This work is supported by the Ministry of Science, Technological Development, and Innovation of the Republic of Serbia (grant number 451-03-136/2025-03 from 27 January 2025).

Keywords: artificial intelligence, democracy, equality, autonomy, public sphere, artificial general intelligence, artificial superintelligence

INTRODUCTION

In recent years, artificial intelligence (AI) has been increasingly moving from the realm of science fiction to reality, even in its most mundane, almost quotidian aspects. While there is still a significant discrepancy between the visions of artificial intelligence promoted for decades, even centuries[1], by artists and futurists – images of superintelligent machines and humanoid robots – and the current reality of chatbots, image generators, and algorithmic recommendations, artificial intelligence seems to be conquering our daily lives.

Politics being an important (perhaps the most important) aspect of our shared reality, it is not surprising that the body of scholarly literature investigating different aspects of the potential impacts of AI on politics is rapidly and steadily growing. While it is still somewhat underdeveloped compared to, for example, the study of ethical issues related to AI, this research nevertheless already points towards some of the more pressing questions that need to be answered. Combined with the justified concern for democracy in the 21st century world, from the democratic malaise (Newton 2012, 7; Di Gregorio 2021, 12; Kupchan 2012, 62) to the democratic hollowing (Mair 2023, 2) and backsliding (Bermeo 2016, 6; Wolkenstein 2022, 265) in the last decade, the new momentum of AI research, and more importantly, AI applications, prompted the scholars in the field of political science to investigate whether it poses a threat to democracy or can, maybe, be used to strengthen it and make it better.

This is not surprising – after all, artificial intelligence is a technology (or, perhaps more accurately, a family of technologies), and technologies are not politically neutral (Winner 1980, 123; Kranzberg 1986 545; Coeckelbergh 2022, 4). Furthermore, as Narayanan and Kapoor argue, it is a *normal* technology (Narayanan and Kapoor 2025). And every new technology, according to Winner, is lauded as the great democratizer, while eventually serving those already in power to further

[1] One of the first descriptions of superintelligent machines and their potential to enslave humanity can be found in Samuel Butler's Erewhon, first published in 1872 (Butler [1872] 1901).

strengthen and entrench their position (Winner 2010, 107). Just like with the other technologies with socially transformative potential, such as the nuclear power and the Internet, AI's democratizing, or de-democratizing potential merits serious consideration (see, for example, Hand and Sandywell 2002, 198; Dahl 1953, 1; Bartoletti 2020, 127).

The aim of this paper is to explore and systematize current research on the relationship between AI and democracy, primarily focusing on existing AI applications, in order to understand what are the general avenues of exploration conceived within the field of political science on this topic, and to evaluate their results so far. However, some thought will also be given to the more far-fetched possibilities of AI development and their implications for democracy. Both of these lines of inquiry should help in identifying the overlooked issues and outlining possibilities for further research.

In order to properly assess the existing scholarship on the topic, it is necessary to start with conceptual issues, given that both artificial intelligence and democracy are notoriously hard to define. Afterwards, the paper will consider AI as a threat to democracy, by reexamining some of the issues identified so far in the scholarly literature – AI's discriminatory potential, its relation with political participation, as well as its impact on the public sphere. Another section of the paper will be devoted to the examination of the ways in which artificial intelligence could be used to strengthen democracy. Finally, a section of the paper will consider the possible futures of artificial intelligence, such as the emergence of artificial general intelligence (AGI) or artificial superintelligence (ASI), and their potential implications for democracy.

COMPLICATED CONCEPTS: DEFINING AI AND DEMOCRACY

In order to determine how artificial intelligence relates to democracy, it is necessary to clear up the terminological confusion and provide at least working definitions of both *artificial intelligence* and *democracy*. This is not an entirely trivial task, given that democracy is notoriously an essentially contested concept, singled out as such by Gallie himself in his seminal paper (Gallie 1955, 168). While a certain degree of conceptual confusion is a rule rather than an exception in the social sciences, definitions of "technical terms", and one may argue that artificial intelligence *is* one of them, are usually more precise and less

controversial. And yet, there seems to be no consensual definition of AI. The body of literature on artificial intelligence struggles, time and again, with the very definition of its basic concept. Some authors even claim that it is intrinsically, and by design, nebulous, vague, and empty of objective meaning (Katz 2020, 6).

The origin of the term *artificial intelligence* is well documented: it was introduced at the conference at Dartmouth College in 1956 by John McCarthy. However, its content has been changing ever since. Some authors point out that at least part of the problem stems from the fact that there is no agreement on the definition of intelligence as such (Legg and Hutter 2007, 392; Warwick 2013, 12). While this may well be a significant obstacle for achieving a consensual definition of AI, there are other, equally important issues at play.

Nearly four decades ago, Negrotti[2] collected about 180 definitions of artificial intelligence from academics and practitioners gathered for a major conference. These definitions turned out to be quite different from one another, yet could be grouped, according to Trappl's interpretation, into two major categories, based on the perceived *objectives* of AI. The first group sees AI as "the science and technique to make computers smart", which would enable them to "perform tasks which normally require human intelligence", while the second sees AI as a tool to "better understand model human intelligence" (Trappl 1992, 3). Others point out that not only objectives, but also *technologies* that constitute AI are disputed – Katz, for example, shows that in the 1960s neural networks were not seen by practitioners as part of AI, while today this is considered one of the basic technologies beyond AI applications (Katz 2020, 49). In his view, AI "quickly became an evocative label for an academic field, a concept, and an industry" (Katz 2020, 3).

Another potential problem is that public perceptions of AI stem primarily from popular culture, where artificial intelligence is presented as general, usually smarter than humans, conscious, and sometimes anthropomorphic. In his paper, Preethi suggests that "[s]ome industry experts believe that the term artificial intelligence is too closely linked to popular culture, causing the general public to have unrealistic fears about how it will change the workplace and life in general" (Preethi

[2] His original research was published in 1983 (Negrotti 1983). It is cited here according to Katz (Katz 2020, 37) and Trappl (Trappl 1992, 6). Some of Negrotti's results relevant to the topic of this paper are also published later (see Negrotti 1987, 114).

2020, 40). And yet, the current state-of-the-art in the AI field is very far from these depictions.

While it is common that the very concept of *technology* is frequently equated exclusively with the latest generation of technological innovations,[3] this trend seems to be even more amplified when it comes to AI, and not only among the general population, but also among the researchers or practitioners. As Narayanan and Kapoor remind us, there is a running joke that AI is "what hasn't been done yet", suggesting that earlier successful and "domesticated"[4] applications are simply not referred to as AI anymore (Narayanan and Kapoor 2024, 14).[5] In their view, artificial intelligence is an umbrella term covering a variety of technologies, only loosely related, and its use is frequently influenced by historical usage, marketing, etc.[6] These authors, however, do not see the lack of conceptual clarity as a major issue and suggest focusing on specific problems and applications.

In order to provide a clear focus for the research, in this paper, the definition of AI provided by the High-level expert group on artificial intelligence[7] appointed by the European Commission will be used. According to this group of experts, "artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals" (European Commission's High-Level Expert Group on Artificial Intelligence 2018). This definition is precise enough to distinguish AI from other technologies, and still broad enough to encompass the current forms of AI, such as, among others, generative and predictive AI used in most commercially available applications, as well as potentialities such as artificial general intelligence and artificial superintelligence. It thus covers all the major varieties of AI that this paper will consider.

Debates about the meaning of democracy are much longer-running and more complicated. There are competing theories of democracy, which reflect the complexity of the phenomenon itself, as well as its

[3] It could be argued that what people considered by technology in, for example, the 1950s is very different from what they consider by technology today.

[4] Domestication of technology is a framework developed by Berker *et al.* to explain how technologies transition, in the eyes of the public, from new and exciting/dangerous to mundane and "boring" (Berker *et al*. 2005, 2–9).

[5] Many other authors have made similar points (see Sheikh *et al*. 2023, 17).

[6] In this, they basically concur with Katz, who claims, *inter alia*, that AI is what AI people, at a given point of time, do (Katz 2020, 33).

[7] For more details about this group, see European Commission 2025.

capacity for transformation (see, for example, Cunningham 2002, 15–26). Furthermore, the notion of democracy is always, at least to some extent, normative. This means that every discussion about democracy is a partial discussion about values, thus earning it the already mentioned qualification of essentially contested concept. It is not surprising, then, that there is a vast amount of scholarly literature on the topic, which cannot be adequately addressed within the scope of this paper.

For the purposes of this paper, the term *democracy* will be used to denote primarily representative, or *liberal* democracy, as defined most famously by Robert Dahl (Dahl 1989, 83–118). Drawing from the broad principles of intrinsic equality and effective participation, he defines democratic process as a normative concept that can, in theory, be applied in all settings requiring collective and binding decisions. When applied to modern nation-states, it manifests as a polyarchy, a political system where institutions embodying and enabling the democratic process are present over a certain, but not firmly fixed, threshold. Inclusion, personal autonomy, freedom of expression, as well as access to alternative sources of information, are important features of his (and most other) conceptions of democracy.

Focusing on liberal democracy does not, of course, imply that other conceptualizations of democracy are irrelevant for the discussion about AI and democracy, but including all, or even some of them, would dramatically expand the scope of this paper. Exploration of diverse theoretical approaches to democracy and their relationship with artificial intelligence is in itself a worthy research goal, and some important steps in that direction have already been made by Coeckelbergh, who discusses the issues that AI can cause within the frameworks of deliberative, participative, and agonistic theories of democracy (Coeckelbergh 2022, 62–84; 2024, 29–38).

IS ARTIFICIAL INTELLIGENCE A THREAT TO DEMOCRACY?

Most of the scholarly research on AI and democracy explicitly sees AI, in its current form and its current applications, as a potential or actual threat to democracy. In part, it is probably caused by the observation that we are witnessing, in Huntington's terms, a reverse wave of democracy (Huntington 2012, 15–16). But also, AI is seen as part of the debate on the

role of *Big Tech*[8] in society. The concentration of economic power and, consequently, the political influence these companies wield, is viewed by many as a major threat to democratic politics and, more widely, social and political equality as its underlying condition (see, for example, Zuboff 2019, 497; Giblin and Doctorow 2022, 94; Varoufakis 2023, 27). From this perspective, to put it succinctly, the threat to democracy is not really AI, it's rampant capitalism (Coeckelbergh 2022, 21–22; 2024, 48).

There are, however, certain domains in which artificial intelligence is currently, and increasingly, applied that seriously undermine democracy by degrading social and political equality, personal autonomy, and the free and diverse public sphere as preconditions for political participation and democratic decision-making. In the following pages, some of these issues will be explained in more depth.

AI *vs.* Equality

Robert Dah's theory builds on the principle of *intrinsic equality* – the notion that each person's interests are equally deserving of consideration, and that no group should be privileged, in terms of influence on political decision-making, above the others, based on features such as gender, wealth, race etc. In this sense, the issue of political equality is closely tied to the previously mentioned concentration of wealth in the hands of the few.[9]

But the modern digital technologies are eroding equality in much more specific ways. In her seminal work, O'Neil has demonstrated that algorithms and big data entrench and deepen existing inequalities, targeting and affecting the most vulnerable groups disproportionately (O'Neil 2016, 11). While she insists that her examples do not include AI, both algorithms and big data are the necessary building blocks of all AI applications, and it has been shown that the use of AI only exacerbates this trend (Narayanan and Kapoor 2024, 53; Coeckelbergh 2024, 47). If automated sentencing systems[10], for example, consistently propose

[8] *Big Tech* is a shorthand term denoting a small number of powerful technological companies, mostly based in a few hubs in the USA and China. For a more detailed explanation, see Birch and Bronson 2022, 3–5.

[9] It should be noted that throughout his work, Karl Marx argued that political equality without meaningful social (including economic) equality is just an empty shell.

[10] Such systems are already in use in a number of jurisdictions in the USA, and there is a growing body of academic, although primarily legal literature, on their features as well as their intended and unintended consequences.

harsher penalties for the people belonging to certain social groups, based on indicators such as race or income, they are not only propagating existing biases, but also potentially disenfranchising them, for example, in the states where prisoners or convicts are excluded from participating in the elections.

Even though the most common justification for the use of such systems in all domains of public services is to provide impartial and objective results, eliminating bias by removing humans (who are prone to bias) from the equation, predictive systems based on AI are repeatedly reported to produce biased results. These biases can be "in the training data, in the algorithm, in the data the algorithm is applied to, and in the teams that program the technology" (Coeckelbergh 2022, 38). Some authors are very pessimistic about the possibility of fixing those systems, pointing out the inherent limits and flaws of the so-called *predictive* AI (see Narayanan and Kapoor 2024, 10).

Even more directly, the manipulation of voters through micro-targeting, revealed in the 2016 Cambridge Analytica scandal, could influence the outcome of the elections. Whether that was the case in the USA presidential election is unclear, but the testimonies suggest that it *is* possible to influence the results by targeting a small number of "swing" voters in a small number of "swing states". According to O'Neil, this effectively creates "the political 1%", a tiny number of voters on whom the majority of resources of political campaigns are focused (O'Neil 2016, 196).

It could be argued that such targeted manipulation is a concern only for particular electoral systems and can thus be easily solved through electoral engineering, but, scale-wise, there is an issue with AI on the opposite end of the spectrum. Regardless of the type of electoral system, generative AI can be used to produce an extremely vast quantity of text impersonating citizens. If the policy-makers are faced with such an influx of materials, for example, letters from concerned citizens, it could influence their decision, especially if they are unable to distinguish between genuine and AI-generated e-mails. Several such instances were already recorded, albeit on a smaller scale and related to single political issues (Narayanan and Kapoor 2024, 140–141).

AI *vs*. Autonomy

There are two main applications of AI that restrict personal autonomy as a prerequisite for democracy: surveillance and recommendation systems. Surveillance has been a long-standing topic of the political debate. Since Jeremy Bentham's invention of the *panopticon*, and later adaptation of the term by Foucault, it is a widely accepted fact that people tend to behave differently when they believe that they are being watched (Bentham 2020, 35; Foucault 1977, 195). Similarly, Ellul warned about the dangers of technologically enabled surveillance turning modern societies into concentration camps (Ellul 1964, 100). The increase of surveillance and data collection about the population indeed went hand in hand with the very construction of the modern nation state. "Nation states," Coeckelbergh points out, "especially, are data hungry" (Coeckelbergh 2024, 14–15). They are thus very keen on quickly adopting technologies that enable more widespread and more efficient ways to track their own citizens (as well as the others who are within reach).

While such technologies can obviously be abused by authoritarian regimes and malicious actors within democracies, even the most democratic governments tend to see them as, at least, a necessary evil contributing to better security. Artificial intelligence was promptly mobilized for the task, and today, one of the most widespread applications of AI is in the field of facial recognition. Combined with the variety of other data compiled about citizens by state and private actors, complete surveillance seems to be within reach. Even without further repressive actions from the state, the awareness that they are being watched could seriously alter the manner in which people behave and the extent to which they are willing to participate in various types of political activities – from attending protests to speaking their mind on social media, but also in private conversations.

Another way for AI applications to degrade personal autonomy is by changing the choice architecture. This can be done by several means: recommendation systems, algorithmic feeds on social media, or "nudging". Recommendation systems are probably the most pervasive and the most underestimated use of artificial intelligence today. After all, streaming platforms, online stores, and other businesses seem to know us so well (thanks to the vast amount of data about us that they have collected) that their recommendations of what to listen to, watch, or buy

are rarely completely off the mark. For this reason, they are frequently implicitly trusted. This trust could, however, be abused to skew the election results. Similar technologies can be used in political marketing, recommending parties and candidates who are the most in line with the interests and ideologies of a specific voter, or tailoring the candidates to suit the particular group or individual.[11]

In the early years of social media, the feeds, that is, the posts that are shown to a user when they log in, were quite straightforward. They would show posts from their friends on the platform, and from the pages they decided to follow. But by 2016, most platforms had switched to so-called algorithmic feed, where the platform would show you a personalized feed based not exclusively on your preferences, but on engagement optimization, and not in chronological order. The exact extent to which the algorithmic feeds on social media are based on artificial intelligence is not certain, since the software running them is proprietary and, by and large, not available for external scrutiny. The consequences of the algorithmic feeds will be discussed in greater detail in the next section, but in the context of autonomy and autonomous choices, the important issue is limiting the range of choices available to citizens. If the visibility of political candidates and their positions on a range of social issues is determined by an opaque algorithm, it could easily affect their chances in the electoral race. Some of the companies behind platforms, most prominently Meta, the owner of Facebook, have experimented with their ability to influence citizens' interest in voting, but also with their voting preferences.[12]

Finally, *nudging* is another technique that can be enhanced by artificial intelligence and used to circumscribe autonomous choices. Introduced by Thaler and Sunstein in their highly influential book, it suggests that human behavior can be influenced without coercion and without limiting personal autonomy of the choice, by making simple changes in the choice architecture – for example, by switching the default answer from *no* to *yes* (Thaler and Sunstein 2021, 103–130).

[11] This kind of tailoring is not a new development in political marketing, but the use of AI could make it much more successful (see O'Neil 2016, 187; McGinniss 1988, 26).

[12] The general description of these experiments can be found in O'Neil (O'Neil 2016, 180–184). For more detailed analysis of the different aspects of the voter turnout (see Theocharis and Lowe 2016, 1465–1486; Haenschen 2023, 1661–1681; 2016, 542–563).

Possible uses in the public policy have drawn much attention both from scholars and policy-makers (for a nuanced discussion, see Schmidt 2017, 405), as well as much criticism (for example, Morozov 2013, 198). Even if there is no overt coercion, the subtle persuasion of nudging can still be construed as manipulation, and, as such, antithetical to autonomous choice. Again, while AI is not a necessary component in nudging, its use in this domain could exacerbate the scale of the problem (Coeckelbergh 2022, 17).

AI and the Public Sphere

Even among the scholars who do not subscribe to the theory of deliberative democracy, grounded in Habermas's work on the public sphere (Habermas 1991, 236) and communicative action (Habermas 1985b, 43–76; 1985a, 273–338), there is a consensus about freedom of expression and an informed public being the prerequisites of a functional democracy. In Dahl's theoretical framework, both freedom of expression and access to alternative sources of information are among the institutional guarantees that make a state a polyarchy. There is already much debate about the role of social media in the degradation of the public sphere.[13]

The conventional wisdom is that the algorithmic feeds work towards reducing the public sphere to many small private spheres, creating so-called echo chambers and epistemic bubbles, thus reducing the space for the debate about social and political issues (Coeckelbergh 2022, 76). Focus on engagement fuels polarization, and the use of AI to boost or game social media algorithms is likely to further aggravate this problem. Even though AI is increasingly used in content moderation across all platforms, it seems that it will not be the solution to polarization and the dissolution of the public sphere (Narayanan and Kapoor 2024, 179).

[13] Interestingly, until 2016, most of the scholarly discourse was along the lines of the Internet and the social media being the new, and potentially better public sphere. The turning point seems to be the Cambridge Analytica scandal, and the subsequent scholarship is much more interested in the dangers social media poses to the public sphere.

WHAT COULD THE FUTURE BRING?

All of the applications considered artificial intelligence that are in use today fall under what is usually referred to as the *narrow* AI – namely, a system that is designed to perform a single task or several related tasks. This pertains to generative AI as well, despite the illusion of generality it projects. By many researchers and practitioners in the field, narrow AI is seen as just a step on the road towards a true goal: artificial general intelligence. Whether it is conceived as an imitation of human intelligence or in some other manner, AGI is generally understood as artificial intelligence that is capable of performing any[14] task a human being can perform. The predictions about AGI differ significantly – some researchers believe that this goal could now be achieved in just a couple of decades, while others doubt that it can be achieved at all (Armstrong and Sotala 2015, 24; Muehlhauser and Salamon 2012, 25; Mueller 2024).

The emergence of AGI potentially opens a new set of issues regarding democracy, especially in relation to the concept of inclusion. It is illustrative that when, at a lecture some years ago, Ray Kurzweil was asked, "In a world where AIs passed the Turing test,[15] who gets to vote? Does democracy make sense?"[16] He immediately interpreted the question as "should AI be allowed to vote?" (and gave no clear answer).

Although the original question can be interpreted in different, and perhaps more interesting ways, the issue of political rights for AGI is important for the discussion about democracy. A crucial part of democracy

[14] Narayanan and Kapoor would add "economically relevant" tasks (Narayanan and Kapoor 2024, 150).

[15] The Turing test is a popular name given to the "imitation game" introduced by Alan Turing in his seminal paper (Turing 1950). He proposes that if the way in which the machine communicates is indistinguishable from human communication, then the machine can be labeled as intelligent. However, chatbots, from the early and rudimentary ones such as ELIZA (for more information on this program developed in the 1960s, see Weizenbaum 1966, 36), to popular apps such as ChatGPT or Claude, have been passing this test for some time, without being intelligent in any meaningful way. In recent decades, there has been a wide consensus about the inadequacy of the Turing test for assessing machine intelligence, with some authors going so far as to call it a "blind alley" (Whitby 1996, 53–65) as well as an "ideology" (Halpern 1987, 79–93) of AI research. However, its significance as an inspiration to the entire field can not be disputed.

[16] The video can be seen at Vimeo (Vimeo 2015). Ray Kurzweil is sometimes referred to as "the prophet of Singularity" (see, for example, Tirosh-Samuelson 2012, 722) and has authored several books that deal, among other topics, with artificial intelligence (see Kurzweil 2005, 203–226; 2001, 40–57; 1992, 401–416).

is, obviously, *the demos* – a body politic of citizens who are allowed to participate in decision-making. Who constitutes the demos is, thus, not a trivial question, even though some authors have tried to dismiss it.[17]

For democracy to function, the demos must be clearly outlined. Both Dahl and Sartori insist that, although democracy must be as inclusive as possible – that is, the largest possible number of people within a democracy should be included in political decision-making – the demos is also by definition *exclusive* (Dahl 1989, 119; Sartori 1987, 21–25). In other words, it is necessary to determine precisely who is and who isn't part of the demos in a democratic political unit. The composition of the demos, obviously, varies throughout both time and space, but some categories are consistently excluded even in our age of universal suffrage: for example, children and non-citizens. Both of these grounds for exclusion bear on the issue of AGI and its political rights.

The suffrage is today usually tied to the age of majority, which is also the moment when the full legal capacity is acquired. Children are excluded from the demos based on their incompetence. However, it is not necessary self-evident that this incompetence relates to levels of *intelligence* and not, for example, emotional immaturity. So, even if AGI is a human-level intelligence, is this enough for it to be recognized as part of the demos?

Part of the problem lies in the conceptual confusion surrounding the very concept of intelligence, and its relationship with other concepts, such as consciousness and sentience. Is self-awareness and the capacity to feel necessary for one to be recognized as a subject of political rights? Can AGI be truly intelligent without being conscious and/or sentient? While there are strong arguments for AI, even if it never achieves general intelligence, consciousness and sentience, it could be eligible for some kind of political standing (Coeckelbergh 2022, 142), right to vote is a completely different matter.

Another ground for exclusion from a specific demos is citizenship. As a rule, only citizens can vote in the national election. Even if we argue that there are no good reasons why AGI should not be granted citizenship,[18] and the rights stemming from the citizen's status (see

[17] Most famous among them is probably Joseph Schumpeter, who argues that the demos defines itself (Schumpeter 2018, 258–259).

[18] In 2017, part of a publicity stunt, a "social robot" named Sophia was granted citizenship of Saudi Arabia, spurring a series of opinion pieces (see, for example, Reynolds 2018).

Jaynes 2020, 346), it is not clear how AGI's citizenship would be determined. Should it, by default, be given by the state where the legal entity responsible for the AGI's inception resides? If the AGI in question is not embodied within a, for example, robot, but exists only as software on the Internet, where does it reside (and, consequently, where should it vote)? If the AGI is capable of making copies of itself and storing them on servers around the world, should every one of those copies be granted suffrage under appropriate state laws?

Finally, if the current AI applications are endangering democracy by limiting human autonomy, the question of AGI's autonomy could also be posed. It could be argued that its choice architecture would be even more severely limited by original programming (see Mueller 2024). Would the AGI then be capable of autonomous action and autonomous choice?

Many authors have warned that, even though it may be the Holy Grail of AI research, AGI would not be the end of the AI evolution. They argue that if AGI is achieved and capable of self-improvement, it would quickly evolve into artificial superintelligence, possibly leading to an intelligence explosion and/or singularity. While these two concepts are frequently used interchangeably, they are not quite synonymous. The intelligence explosion, according to Muehlhauser, refers to the moment when intelligent machines will surpass the human level of intelligence (Muehlhauser 2013, 79). The technological singularity, on the other hand, seems to be a more vague and somewhat contested term. In Sandberg's words, some definitions of technological singularity "stress the role of artificial intelligence, others refer to more general technological change (Sandberg 2013, 377). These meanings can overlap, and many writers use combinations of meanings." However, it seems that it is mostly understood as a moment of extremely accelerated technological change, which renders human comprehension and prediction impossible. This is frequently seen as a direct consequence of the intelligence explosion. Optimists place the possible occurrence of such an event somewhere in the second half of the 21st century.[19] There is a certain logic in the described trajectory from AGI to ASI: if AGI were able to take over *any* task, one of those tasks would probably be further AI research, thus resulting in more and more sophisticated and powerful AIs, which would

[19] Pessimists, on the other hand, believe that neither AGI nor ASI can ever be achieved.

quickly drastically surpass humans in every possible way (Narayanan and Kapoor 2024, 151).

The emergence of ASI is considered to pose an existential risk to humanity by a significant number of academics and public figures. The most prominent among them are probably Eliezer Yudkowsky and Nick Bostrom (Yudkowsky 2008, 308–345; Bostrom 2014, 115–116).[20] Even if it turns out that the inception of ASI is not an extinction-level event, it would still profoundly influence and shape human society and politics, with possible grave consequences for democracy.

One of the first things the superintelligence would strive to do, according to Bostrom, is to secure its position as a singleton. Singleton, in this context, means an entity with the global decision-making capacity, with no competing peers. In other words, perhaps more recognizable in the political science literature, it would become a global sovereign. What would that mean for human politics as such and for democracy specifically is hard to determine. Perhaps ASI would introduce totalitarian control beyond everything humanity has experienced so far. On the other hand, perhaps it would not be interested in human affairs at all, as long as they are not in the way of ASI's own goals.

But democracy would nevertheless be in peril, even if ASI wanted to preserve it. The chasm between the knowledge and processing power available to ASI and that available to humans would be so vast that the most of the arguments for democracy and against guardianship presented by Dahl could not hold anymore (Dahl 1989, 52–65).[21] If we do not allow children to participate in political decision making, why would ASI let us do it, being presumably much more superior to us than we are to children?

Of course, every debate about AGI and ASI is at this point completely speculative. It was already noted that insisting on risks stemming from these hypothetical future forms of AI could very well distract us from real harms that AI applications are causing now. Nevertheless, thinking about these matters can be inspiring and useful for testing the limits of justifications for democracy and its ability to cope with new and unforeseen challenges.

---

[20] Bostrom does not necessarily see superintelligence as machine intelligence. He identifies three possible paths that could lead to superintelligence: artificial intelligence, biological cognition enhancement, and collective intelligence.

[21] For a more elaborate discussion on this particular topic, see Damnjanović 2015, 79–80.

CONCLUSION

The scholarly debate about relationships between artificial intelligence and democracy is thriving, even if it is still in its early stages. The importance of the topic is recognized not only by scholars but also by policy-makers. For example, the European Union has explicitly classified AI systems "intended to be used for influencing the outcome of an election or referendum or the voting behaviour of natural persons in the exercise of their vote in elections or referenda" as high-risk (Regulation 2024/1689, Annex III). In the wider context of the ongoing crisis of (liberal) democracy, rapidly evolving AI applications and their impact on societies in general and on democracy in particular are rightly seen as a cause for concern.

Main reasons why the AI, in its current form, is seen as a threat to democracy are its already documented roles in increasing or entrenching inequalities, restricting personal autonomy through mass surveillance and nudging, further eroding the public sphere, and concentration of economic (and, consequently, possibly political) power.

However, these harmful impacts could be overturned by putting existing AI systems to different use: by making them work in such a way as to improve, not degrade, democratic values and practices. An important part of this would be a wider democratic debate about AI, as well as a more participatory approach to AI development (see Coeckelbergh 2022, 152; 2024, 81; Narayanan and Kapoor 2024, 17, 265).

Future threats to democracy posed by the development of artificial general intelligence, and, subsequently, artificial superintelligence are, for now and maybe forever, mostly useful as thought experiments. It could be argued that most of the interest stems from dire doomsday scenarios, as well as from anthropomorphizing AI, that is, ascribing to it the humanlike features it does not possess. While the value of such thought experiments for examining and predicting the future of AI research and application is limited, they can help us think about the limits and weak points of our understanding of democracy and its justifications. They should not, however, divert us from the search for solutions to more pressing problems presented by AI, and for ways to reimagine and restructure AI research and application in ways more supportive and conducive to democracy.

REFERENCES

Armstrong, Stuart, and Kaj Sotala. 2015. "How We're Predicting AI–or Failing To." In *Beyond Artificial Intelligence*, edited by Jan Romportl, Eva Zackova, and Jozef Kelemen, 11–29. Cham: Springer.

Bartoletti, Ivana. 2020. *An Artificial Revolution: On Power, Politics and AI*. London: Indigo Press.

Bentham, Jeremy. 2020. *The Panopticon Writings*. London: Verso Books.

Berker, Thomas, Maren Hartmann, and Yves Punie [Berker *et al*.]. 2005. *Domestication of Media and Technology*. Maidenhead: McGraw-Hill Education.

Bermeo, Nancy. 2016. "On Democratic Backsliding." *Journal of Democracy* 27 (1): 5–19. DOI: 10.1353/jod.2016.0012

Birch, Kean, and Kelly Bronson. 2022. "Big Tech." *Science as Culture* 31 (1): 1–14. DOI: 10.1080/09505431.2022.2036118

Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Butler, Samuel. [1872] 1901. *Erewhon: Or, Over the Range*. Project Gutenberg. https://gutenberg.org/cache/epub/1906/pg1906-images.html

Coeckelbergh, Mark. 2022. *The Political Philosophy of AI: An Introduction*. Hoboken, NJ: John Wiley & Sons.

Coeckelbergh, Mark. 2024. *Why AI Undermines Democracy and What To Do About It*. Hoboken, NJ: John Wiley & Sons.

Cunningham, Frank. 2002. *Theories of Democracy: A Critical Introduction*. London: Routledge.

Dahl, Robert A. 1953. "Atomic Energy and the Democratic Process." *The ANNALS of the American Academy of Political and Social Science* 290 (1): 1–6. DOI: 10.1177/000271625329000101

Dahl, Robert A. 1989. *Democracy and Its Critics*. New Haven, CT: Yale University Press.

Damnjanović, Ivana. 2015. "Polity Without Politics? Artificial Intelligence Versus Democracy: Lessons From Neal Asher's Polity Universe." *Bulletin of Science, Technology & Society* 35 (3–4): 76–83. DOI: 10.1177/0270467615623877

Di Gregorio, Luigi. 2021. *Demopathy and the Democratic Malaise: Symptoms, Diagnosis and Therapy*. Cheltenham: Edward Elgar Publishing.

Ellul, Jacques. 1964. *The Technological Society*. New York: Vintage Books.

European Commission's High-Level Expert Group on Artificial Intelligence. 2018. "A Definition of AI: Main Capabilities and Scientific Disciplines." *European Commission*. December 18, 2018. https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf

European Commission. 2025. "High-level expert group on artificial intelligence". *Digital-strategy*. May 6, 2025. https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai

Foucault, Michel. 1977. *Discipline and Punish: The Birth of the Prison*. New York: Penguin Books.

Gallie, Walter Bryce. 1955. "Essentially Contested Concepts." *Proceedings of the Aristotelian Society* 56: 167–198.

Giblin, Rebecca, and Cory Doctorow. 2022. *Chokepoint Capitalism: How Big Tech and Big Content Captured Creative Labor Markets and How We'll Win Them Back*. Boston: Beacon Press.

Habermas, Jürgen. 1985a. *The Theory of Communicative Action: Lifeworld and System: A Crituque of Functionalist Reason*. Vol. 2. Boston: Beacon Press.

Habermas, Jürgen. 1985b. *The Theory of Communicative Action: Reason and the Rationalization of Society*. Vol. 1. Boston: Beacon Press.

Habermas, Jurgen. 1991. *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*. Cambridge, MA: MIT Press.

Haenschen, Katherine. 2016. "Social Pressure on Social Media: Using Facebook Status Updates to Increase Voter Turnout." *Journal of Communication* 66 (4): 542–563. DOI: 10.1111/jcom.12236

Haenschen, Katherine. 2023. "The Conditional Effects of Microtargeted Facebook Advertisements on Voter Turnout." *Political Behavior* 45 (4): 1661–1681. DOI: 10.1007/s11109-022-09781-7

Halpern, Mark. 1987. "Turing's Test and the Ideology of Artificial Intelligence." *Artificial Intelligence Review* 1 (2): 79–93. DOI: 10.1007/BF00130010

Hand, Martin, and Barry Sandywell. 2002. "E-Topia as Cosmopolis or Citadel: On the Democratizing and De-Democratizing Logics of the Internet, or, Toward a Critique of the New Technological Fetishism." *Theory, Culture & Society* 19 (1–2): 197–225. DOI: 10.1177/026327640201900110

Huntington, Samuel P. 2012. *The Third Wave: Democratization in the Late 20th Century*. Norman: University of Oklahoma Press.

Jaynes, Tyler L. 2020. "Legal Personhood for Artificial Intelligence: Citizenship as the Exception to the Rule." *AI & SOCIETY* 35 (2): 343–354. DOI: 10.1007/s00146-019-00897-9

Katz, Yarden. 2020. *Artificial Whiteness: Politics and Ideology in Artificial Intelligence*. New York: Columbia University Press.

Kranzberg, Melvin. 1986. "Technology and History:"Kranzberg's Laws"." *Technology and Culture* 27 (3): 544–560.

Kupchan, Charles A. 2012. "The Democratic Malaise: Globalization and the Threat to the West." *Foreign Affairs* 91 (1): 62–67.

Kurzweil, Ray. 1992. *The Age of Intelligent Machines*. New York: Viking.

Kurzweil, Ray. 2001. *The Age of Spiritual Machines: How We Will Live, Work and Think in the New Age of Intelligent Machines*. New York: Texere.

Kurzweil, Ray. 2005. *The Singularity Is Near: When Humans Transcend Biology*. New York: Penguin Publishing Group.

Legg, Shane, and Marcus Hutter. 2007. "Universal Intelligence: A Definition of Machine Intelligence." *Minds and Machines* 17 (4): 391–444. DOI: 10.1007/s11023-007-9079-x

Mair, Peter. 2023. *Ruling the Void: The Hollowing of Western Democracy*. London: Verso books.

McGinniss, Joe. 1988. *The Selling of the President: The Classic Account of the Packaging of a Candidate*. New York: Penguin Publishing Group.

Morozov, Evgeny. 2013. *To Save Everything, Click Here: Technology, Solutionism, and the Urge to Fix Problems That Don't Exist*. London: Penguin Books.

Muehlhauser, Luke. 2013. *Facing the Intelligence Explosion*. Berkley, CA: Machine Intelligence Research Institute.

Muehlhauser, Luke, and Anna Salamon. 2012. "Intelligence Explosion: Evidence and Import." In *Singularity Hypotheses*. edited by Amnon H. Eden, James H. Moor, Johnny H. Søraker, and Eric Steinhart, 15–42. Berlin: Springer.

Mueller, Milton. 2024. "The Myth of AGI." *Internet Governance Project*. Last accessed January 17, 2025. https://www.internetgovernance.org/wp-content/uploads/MythofAGI.pdf

Narayanan, Arvind, and Sayash Kapoor. 2024. *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference*. Princeton, NJ: Princeton University Press.

Narayanan, Arvind, and Sayash Kapoor. 2025. "AI as Normal Technology." *The Knight First Amendment Institute*. April 15, 2025. https://knightcolumbia.org/content/ai-as-normal-technology

Negrotti, Massimo. 1983. *How AI People Think*. Report III/2203/En 83. Chair of Sociology of Knowledge, University of Genoa.

Negrotti, Massimo. 1987. "The AI Peoples Way of Looking at Man and Machine." *Applied Artificial Intelligence: An International Journal* 1 (1): 109–116. DOI: 10.1080/08839518708927963

Newton, Kenneth. 2012. "Curing the Democratic Malaise with Democratic Innovations." In *Evaluating Democratic Innovations*, eds. Kenneth Newton and Brigitte Geissel, 3–20. London: Routledge.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

Preethi, S. 2020. "A Survey on Artificial Intelligence." *International Journal of Intelligent Computing and Technology (IJICT)* 3 (2): 39–42.

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 OJ L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj

Reynolds, Emily. 2018. *The Agony of Sophia, the World's First Robot Citizen Condemned to a Lifeless Career in Marketing*. Wired. June 1, 2018. https://www.wired.com/story/sophia-robot-citizen-womens-rights-detriot-become-human-hanson-robotics/

Sandberg, Anders. 2013. "An Overview of Models of Technological Singularity." In *The Transhumanist Reader: Classical and Contemporary Essays on the Science, Technology, and Philosophy of the Human Future*, eds. Max More and Natasha Vita-More, 376–395. Malden, MA: Wiley-Blackwell.

Sartori, Giovanni. 1987. *The Theory of Democracy Revisited*. Chatam, NJ: Chatham House Publishers.

Schmidt, Andreas T. 2017. "The Power to Nudge." *American Political Science Review* 111 (2): 404–417. DOI: 10.1017/S0003055417000028

Schumpeter, Joseph A. 2018. *Capitalism, Socialism and Democracy*. Floyd, VA: Wilder Publications.

Sheikh, Haroon, Corien Prins, and Erik Schrijvers [Sheikh *et al*.]. 2023. "Artificial Intelligence: Definition and Background." In *Mission AI: The New System Technology*, eds. Haroon Sheikh, Corien Prins, and Erik Schrijvers, 3–28. Cham: Springer.

Thaler, Richard H., and Cass R. Sunstein. 2021. *Nudge: The Final Edition*. New York: Penguin.

Theocharis, Yannis, and Will Lowe. 2016. "Does Facebook Increase Political Participation? Evidence from a Field Experiment." *Information, Communication & Society* 19 (10): 1465–1486. DOI: 10.1080/1369118X.2015.1119871

Tirosh-Samuelson, Hava. 2012. "Transhumanism as a Secularist Faith." *Zygon*® 47 (4): 710–734. DOI: 10.1111/j.1467-9744.2012.01288.x

Trappl, Robert. 1992. "AI: Introduction, Paradigms, Applications (Including CBR), Impacts, Visions." In *Advanced Topics in Artificial Intelligence*, eds. Vladimir Marik, Olga Štepankova, and Robert Trappl, 1–24. Berlin: Springer-Verlag.

Turing, Alan M. 1950. "Computing Machinery and Intelligence." *Mind* 59 (236): 433–460.

Varoufakis, Yanis. 2023. *Technofeudalism: What Killed Capitalism*. London: Bodley Head.

Vimeo. 2015. "Does Democracy Work With AI?" August 5, 2015. https://vimeo.com/135493343

Warwick, Kevin. 2013. *Artificial Intelligence: The Basics*. London: Routledge.

Weizenbaum, Joseph. 1966. "ELIZA—a Computer Program for the Study of Natural Language Communication between Man and Machine." *Communications of the ACM* 9 (1): 36–45.

Whitby, Blay. 1996. "The Turing Test: AI's Biggest Blind Alley?'." In *Machines and Thought: The Legacy of Alan Turing*, vol. 1, eds. Peter Millican and Andy Clark, 53–65. Oxford: Oxford University Press.

Winner, Langdon. 1980. "Do Artifacts Have Politics?" *Daedalus* 109 (1): 121–136.

Winner, Langdon. 2010. *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. Chicago: University of Chicago Press.

Wolkenstein, Fabio. 2022. "What Is Democratic Backsliding." *Constellations* 30 (3): 261–275. DOI: 10.1111/1467-8675.12627

Yudkowsky, Eliezer. 2008. "Artificial Intelligence as a Positive and Negative Factor in Global Risk." In *Global Catastrophic Risks*, eds. Nick Bostrom and Milan Ćirković, 308–345. Oxford: Oxford University Press.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs.

Ивана Дамњановић*

Факултет политичких наука,Универзитет у Београду

ВИ ПРОТИВ ДЕМОКРАТИЈЕ: САДАШЊОСТ И МОГУЋЕ БУДУЋНОСТИ**

Резиме

Предмет рада је сложен и динамичан однос између вештачке интелигенције (ВИ) и демократије, са тежиштем на тренутно постојећим применама ВИ и њиховим импликацијама како по само друштво тако и по друштвену, а пре свега политичку, теорију. Полазећи од чињенице да се ВИ све више интегрише у свакодневни живот и да расте интересовање академске заједнице за њене политичке последице, у раду се првенствено разматрају начини на које ВИ може да учврсти постојеће структуре моћи и поткопа демократске принципе. Користећи теорију демократије Роберта Дала као оквир, рад идентификује три основне демократске вредности које су угрожене развојем ВИ: једнакост, аутономију и јавну сферу. Системи ВИ, посебно они који се користе за предиктивне анализе и алгоритамско одлучивање, често репродукују друштвене и политичке неједнакости. Међу примерима су пристрасни алгоритми за изрицање затворских казни и микроциљано политичко оглашавање, који непропорционално погађају маргинализоване групе и негатино утичу на суштинску једнакост као претпоставку демократије и изборног процеса. Вештачка интелигенција угрожава персоналну аутономију и аутономију избора већ својим применама у технологијама надзора и алгоритмима за препоруке. Технологије препознавања лица и обраде великих скупова података омогућавају свеприсутан надзор, потенцијално обесхрабрујући политичко ангажовање и слободу изражавања. Алгоритми за препоруке и технике усмеравања суптилно манипулишу понашањем корисника, ограничавајући истински избор, што може да утиче на резултате

* Имејл: ivana.damnjanovic@fpn.bg.ac.rs; ORCID: 0000-0002-8553-492X

** Овај рад је подржан средствима Министарства науке, технолошког развоја и иновација Републике Србије на основу Уговора о реализацији и финансирању научноистраживачког рада НИО у 2025. години, број: 451-03-136/2025-03 од 27.01.2025. године.

избора. Овакве примене ВИ изазивају забринутост због ерозије личне слободе схваћене као аутономије, неопходног темеља грађанства у демократским државама. Јавна сфера, неопходна као простор за размену мишљења обавештених грађана, такође је угрожена. Алгоритми друштвених мрежа које покреће, или ће у најближој будућности покретати, ВИ фрагментишу јавни дискурс у ехо-коморе, подстичући поларизацију и смањујући могућности за делиберацију. Иако се на платформама друштвених мрежа ВИ већ користи за модерацију садржаја, није се показала ефикасном у сузбијању ових трендова. Један сегмент рада посвећен је и спекулацији о могућим будућностима које укључују општу вештачку интелигенцију (ОВИ) и вештачку суперинтелигенцију (ВСИ). ОВИ, дефинисана као ВИ способна да обавља било који задатак на истом нивоу као човек, отвара питања о укључивању у демократски процес. Ако ОВИ достигне ниво људске интелигенције, да ли би требало да добије право гласа или држављанство? У раду су размотрени неки од изазова дефинисања „демоса" у таквим сценаријима, укључујући питања телесности, репликације и аутономије. ВСИ, потенцијално последица самоусавршавања ОВИ, могла би надмашити људску интелигенцију и постати глобални суверен. Оваква верзија будућности подстиче на размишљања о апокалиптичним сценаријима краја људске врсте, политике и демократије. Чак и када би ВСИ не би угрожавала демократске системе, огромна когнитивна разлика између људи и суперинтелигентних ентитета могла би учинити традиционалне про-демократске аргументе застарелим. У закључку рада констатује се да негативни утицаји ВИ по демократију нису неизбежни и да се могу ублажити кроз демократску дебату и партиципативни дизајн, односно веће учешће демократске јавности у одлучивању о правцима развоја ВИ и њеним применама. Спекулативне расправе о ОВИ и ВСИ, иако корисне за тестирање неких поставки демократске теорије, не би требало да скрену пажњу са решавања стварних и актуелних проблема које изазивају већ постојећи ВИ системи.

Кључне речи: вештачка интелигенција, демократија, једнакост, аутономија, јавна сфера, општа вештачка интелигенција, вештачка суперинтелигенција

* This manuscript was submitted on September 14, 2025, and accepted by the Editorial Board for publishing on November 10, 2025.

UDC 327:004.8:338.2
DOI: 10.5937/spm94-59849
Original research article

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 25-42

Marko Pejković*

Institute for Political Studies, Belgrade

ARTIFICIAL INTELLIGENCE AND CENTRAL ECONOMIC PLANNING

Abstract

In this paper, the author will attempt to elaborate and complement the existing arguments about the impossibility of artificial intelligence (AI) ever enabling centralized, efficient economic planning at the state level in the form of a government plan. The rapid development of AI and other technologies in recent years has reawakened the debate about the upcoming technical capacities that could perhaps, in the near future, overcome all the weaknesses of centralized economic planning known from the previous century. However, the lack of conceptual prerequisites in the domain of AI that are necessary for successful centralized planning actually argues in favor of the fact that even any experiment on this topic is both scientifically and economically unjustified. The introductory part of the paper provides an overview of the main concepts and dilemmas within the debate, the next part discusses mathematical paradoxes as conceptual obstacles to economic planning with the help of AI (the problem of self-reference), while the final part of the paper will be a critique of the attempts of AI ideologists-architects to overcome these obstacles through the eventual fusion of humans and AI machines.

Keywords: AI, central economic planning, market, self-reference, Gödel's theorems, algorithms, superintelligence

* E-mail: marko.pejkovic@ips.ac.rs; ORCID: 0009-0004-5920-1381

INTRODUCTION

In the last few decades, it is an undeniable fact that people around the world have increasingly relied on digital technologies, whether for their business obligations or their private daily lives, but no current high technology has captured the world's attention as much as AI, precisely because of the controversies that AI has caused (Luknar 2025, 146–147). In the field of healthcare, more and more patients are using devices and sensors that register and analyze instantaneous changes in their bodies (calorie consumption, sleep rhythm, pulse) in order to provide physicians with recommendations for the optimal treatment method. In education, automated assessment, translation, and personalized learning (adapted to the cognitive abilities and work habits of students) are playing an increasingly important role. In the field of business correspondence and marketing, AI applications are increasingly used to automatically write personalized emails, tweets, or blog posts. "Smart homes" are being built with the help of AI systems (from room heating/cooling control to security systems). Management is experiencing an increasing trend of relying on AI software tools for decision-making (automated delegation of tasks to employees according to their work profile), while retail chains are happy to leave the processing of customer requests and complaints to the latest AI developments within their call centers. Some AI theorists list all these areas of human life and the role of AI in them when they talk about the "automated society" (Elliott 2022, 20–21).

On the other hand, the popularity and success of AI in solving "micro" problems in various spheres of society have fueled the hopes of some economists that it is already possible to talk about using AI – together with other information technologies (supercomputers) – to solve problems that are at a much higher level of complexity. So, supposedly, we could already have a central economic plan that would "optimize" the entire national economy with the help of AI. On the one hand, supercomputers would be responsible for calculating in the so-called input-output tables of the plan the value of each final product as the chain sum of each direct/indirect working time spent on its production (Cockshott and Cottrell 1993, 47–48,178–179). More precisely, supercomputers would be responsible for the supply side. On the other hand, AI tools would be the best technological solution for predicting demand patterns (Dapprich and Cockshott 2023, 420–421). The economic trend within which this idea of the relationship between

digital technologies/AI and economic goals is generated is called cyber-communism (Nieto 2023, 254).

However, cyber-communism is not the subject of this paper for three reasons. First, there is already an extremely fruitful critique of its theoretical premises (Lambert and Fegley 2023; Phelan and Wenzel 2023; Moreno-Casas, Espinosa and Wang 2022; Wang, Espinosa and Peña-Ramos 2021). Second, all of these critiques focus only on the supply side, while none of them address the problematic use of AI in cyber-communism on the demand side. Third, we are interested in the potential role of AI not only in the domain of forecasting demand for the central economic plan, but also in its total role in the entire economic process – both on the demand side and on the supply side. And not at the current level of AI development (narrow AI), but at the one that currently exists only in the future (superintelligence). We start from the self-evident assumption that if there are no arguments for centralist economic planning by the projected and more advanced stage of AI, there can be none for the current one, which is inferior.

To fully understand narrow AI and superintelligence, it is necessary to first define AI itself. The diversity and multiplicity of definitions of AI are a consequence of the ubiquity of its applications (Girasa 2020, 8). We believe, however, that there is one definition that best describes what AI researchers have been doing for the past 50 years and what they will continue to do in the future (Ertel 2017, 2) – AI is the study of how to make computers solve problems (those that humans currently solve better) in a more efficient manner (Rich, Knight and Nair 2010, 3). Indeed, AI has sought since its inception to create a comprehensive replica of human cognitive abilities to solve various problems and then to multiply them quantitatively and qualitatively. And this is whether we are talking about a multitude of applications, each of which, as a replica of exactly one of these human abilities, solves only one problem (which would be narrow AI) or about a single application that, as a replica of all these abilities, solves, in a better way than humans themselves, all human problems (superintelligence).[1]

[1] In addition to these two types of AI, there is also general AI that is structurally identical to superintelligence (a single application that solves all problems), but whose results (at least in a qualitative sense) are at the same level as those of humans (Bostrom 2014, 14, 16, 22, 52, 62, 70; Yampolskiy 2016, 41–50). However, since the projections of AI engineers are that general AI will be a relatively short-lived and transitional phase towards superintelligence, we will not mention it in the

Although superintelligence does not yet exist as a developed technology (and there are no indications that it will in the near future, in the coming years, or even decades), we can already test its conceptual foundations despite this fact. More precisely, we are interested in whether it is algorithmically absolutely consistent, i.e., are its algorithms able to always provide a logically meaningful answer to any question that would concern the subject of its work (which in our case is the national economy)?[2] If it is not algorithmically omnipotent *a priori*, then there is no point in further investigating its role in the domain of centralized economic planning in that direction. This dilemma is extremely relevant to social science because redirecting financial resources from studying this issue would achieve a lot in terms of savings, rationalization, and higher-quality investment in scientific projects. And since in the rest of the paper we will only examine superintelligence, for the sake of brevity, when we mention AI, we will refer exclusively to superintelligence (unless we emphasize otherwise that it is a narrow AI).

At the end of this introductory part, it is necessary to explain what we mean in this paper by centralist economic planning. If it has long been customary in science to define this phenomenon as "the replacement of the free market by a central monopoly that makes decisions about how and by whom what will be produced, as well as how the results of production (rewards) will be distributed" (van den Haag 1964, 27–28), there is no reason not to adopt this definition for the purposes of this paper. However, we specify this central monopoly precisely as AI.

rest of the paper, i.e., all the conclusions we draw for superintelligence also apply to general AI. Also, it is of utmost importance for the continuation of the work to keep in mind that superintelligence is one (and not many) and as such centralist in itself.

[2] An algorithm is defined as "a set of well-defined rules for solving certain computational problems" (Roughgarden 2017, 1). An algorithm can also be understood as a "Turing machine" that, like a function, receives inputs and produces outputs according to a finite number of instructions. If a "Turing machine" continuously works without interruption (stopping), the problem is considered unsolvable (non-computational), and if it stops working by providing any answer – the problem is solvable (computational). Moreover, a "Turing machine" is designed so that it does not necessarily have to perform computational tasks in the sense of arithmetic, but can provide "yes-no" answers to mathematical questions or demonstrate proofs to complex questions regarding mathematical theorems (Nguyen 2024, 25).

AI AND THE PROBLEM OF SELF–REFERENCE

There is a belief among neoclassical economists that the mathematical-logical foundations of their theories are quite reliable (Winrich 1984, 987). And since neoclassical economics is the one that has been characterized as the "mainstream" of economic science (Boettke 2012, xvii, 104, 157, 268, 271–272, 284, 290, 324, 378), then such a belief applies to almost the entire discipline of economics. However, this belief is on extremely shaky grounds, due to the mathematical-logical insights into the problem of self-reference, which, admittedly, have not yet been incorporated into the domain of the social sciences (Nguyen 2024, 24; Winrich 1984, 987).

Self-reference occurs when we think not only about the world around us and its objects, but also when we think about our own thoughts, when we are self-aware – when we do not simply use symbols to denote external objects, but when we use symbols to denote the symbols themselves, i.e., our own thought processes (Winrich 1984, 988). This can be illustrated by the famous linguistic paradox: "this statement is false" (Winrich 1984, 988). As Winrich points out, this paradox leaves us in an eternal dilemma about the truth of this statement, because if the statement is true, then what it communicates must be true. And since it communicates that the statement itself is false, then the statement is at the same time false. Conversely, if the statement is false, then what is true is the opposite of what is communicated in the statement, and this would actually be the truth of the statement itself. Therefore, self-reference can also be understood as a statement about a statement – as soon as a statement refers to itself ("this statement is…") in terms of its truth or falsity (Winrich 1984, 989).

Starting from such linguistic paradoxes, mathematics itself has also been preoccupied with the problem of self-reference, as attempts have been made to examine whether even within mathematics itself we have paradoxes like the above-mentioned linguistic ones (Winrich 1984, 989). Winrich reminds us that one of the mathematical paradoxes was detected by Bertrand Russell, when he noticed that there are two types of sets in mathematics – some that do not contain themselves as members of the set (normal sets) and others that do contain themselves as members of the set (abnormal sets). Thus, the set of all cats would be normal, because it is not a cat in itself and does not contain itself. On the other hand, the set of all normal sets is paradoxical. For, if such a set is normal, it must not contain itself as a member of a set, but this is contrary to what its determinant

says – all normal sets are part of it as a set (therefore, as a normal set it is also part of itself because it is normal, regardless of the fact that it is a set of sets). And if it is abnormal, it must contain itself as a member of its own set, but since all members of its set are normal sets, this would be contrary to the thesis that it is an abnormal set. And so on, there is an eternal circularity of indeterminacy (inconsistency) of the answer to the question of whether the set of normal sets is really normal or not (Winrich 1984, 989). Russell tried to remove these paradoxes from mathematics, but without success.[3]

Although some mathematicians besides Russell believed that it was possible to cleanse all mathematical expressions of the paradox of self-reference (completeness and inconsistency), Kurt Gödel proved that completeness and consistency are mutually exclusive (if mathematics is complete, it is not consistent, and *vice versa*). He went a step further and merged statements about mathematical expressions (metamathematics) with the arithmetic branch of mathematics, by encoding an entire system of mathematical notation through which it is possible to obtain any mathematical expression and showed that not only statements about mathematics (metamathematics) are subject to problems of self-reference, but even indirectly mathematics itself as a system in itself (Winrich 1984, 990–991). Thus, the statement "this expression is not provable" has its counterpart in the world of numbers. If the statement is true, it is not provable. In other words, if one insists on the consistency (accuracy) of a mathematical system, one must sacrifice its completeness (the belief that everything is provable) and vice versa.[4] Following Lowenheim-Skolem,

[3] Russell attempted to eliminate these paradoxes of self-reference from mathematics. Elementary mathematical objects would be type 0, sets of such objects type 1, sets of sets type 2 (and so on). A higher set could not be contained in a set of a lower type, in order to avoid self-reference. But Russell realized that the price of avoiding self-reference would be the complete unusability of mathematics. For example, maximization, which is of crucial importance to the "mainstream" of economics, would be impossible - on the one hand, the maximum must always be from the set of real numbers (type 1), but if one insisted on eliminating self-reference, then the concept of maximum (which is by definition type 2) could not imply any maximum that was a real number (which is impossible to imagine in mathematics as a science). Therefore, he introduced the principle of reducibility, which softened the sharp division between "types", thereby effectively returning Russell fruitlessly to the beginning from which he had embarked on his endeavor (Winrich 1984, 990).

[4] Theorem I states that there are arithmetic expressions for which no algorithm can decide whether they are true, and theorem II states that no mathematical system can prove its own consistency within itself (Nguyen 2024, 26).

Winrich concludes, quite correctly, that the very fact that some statements about mathematical expressions cannot be proven opens up the possibility for a non-dualistic interpretation of the world – axiomatic systems always allow for interpretations different from those initially intended (Winrich 1984, 992). That is, no matter how hard we try to categorize phenomena, there is always a chance that our categorizations will prove inadequate, because a phenomenon may meet all the criteria to fall under a given category (type), although it evidently does not belong there.

If the "total interconnectedness of nature" is something that cannot be completely and consistently captured by any axiomatic system (even mathematics), then this is all the more true of economic life, regardless of the efforts of neoclassical economists to make economics an exact science by applying axiomatic methods to economic variables (Winrich 1984, 992). In accordance with Gödel's theorems, it follows, according to Winrich, that it is not possible to make a strict logical distinction between economic concepts such as "utility" and "disutility", "efficiency" and "inefficiency", or "abundance" and "scarcity".[5] This is even better seen in the example of neoclassical choice theory, which assumes that individuals rationally choose the options available to them in order to maximize their preferences, but does not engage in any consideration of the preferences themselves (why individuals have a preference at all, whether they reconsider, change, or abandon it), precisely because it does not want to deal with self-reference, because self-reflection of one's own preferences (desires) is exactly that – self-reference.[6] According to Winrich, people are not programmed machines with fixed preferences, but are in a tense dialectical relationship with their preferences, so much so that they can even give them up altogether – depending on the cultural context in which they are immersed and which cannot be ignored (Winrich 1984, 996, 1000).

[5] Although every language and self-reference is subject to imprecision and dialectics, Winrich believed that science should not transfer these "imperfections" from a "perfect" system to the domain of religion or dogma, much less adapt them to itself through sterile simplification and reduction, but rather adapt to them by reformulating its basic premises (Winrich 1984, 993).

[6] The best example Winrich gives is the statement of a classic cigarette smoker: "I don't like the fact that I have the desire to smoke." This is absolutely impossible to fit into the choice theory, which assumes that individuals consciously and rationally make choices to maximize the realization of their desires that are positive for them. However, this statement clearly demonstrates that some desires can be experienced negatively, as something that is rationally undesirable, but which cannot be rationally resisted (Winrich 1984, 999).

Moreover, the problem is not only the self-reference of individual economic entities that would tomorrow be an insurmountable obstacle to the computational activity of AI, but also the self-reference of the entire economic system that would try to "explain" itself in the form of a centralist plan. This situation can be illustrated by Hayek's argument that any apparatus that performs classification (explains) must have a structure of a higher level of complexity than the lower-order objects that are subject to classification (Hayek 1952, 185). Here, complexity is defined as the number of possible classifications that come into consideration (Hayek 1952, 186). Although Hayek postulated this to prove that our brain will never be able to provide a complete explanation of all the individual ways in which it classifies neural stimuli (i.e., to explain itself), we believe that there is no reason why we should not apply this to any other entity that seeks to explain anything of the same level of complexity (itself), including therefore the central economic plan as the personification of the entire economic system. And although it is reminiscent of Russell's attempt to avoid self-reference with sets of different types, Hayek based his argument on Cantor's theorem that any classification system understood as a power set must always have a greater number of classes than the number of basic elements of the power set (subject to classification), and therefore, a power set cannot have itself as a member of its own set, because this would lead to the self-referential paradox that it is at the same time larger than itself (Hayek 1963, 340; Koppl 2010, 865–866). So a national economy that would "think" about itself in the form of a central plan would actually be an insoluble enigma for any algorithm.

But, regardless of all this, Gödel seemed to anticipate that someone in the future might argue the following – the computational limits mentioned in Gödel's theorems could be overcome when a new technology of such capacity and quality (such as AI) appears. Gödel did not deny this possibility, but he clearly pointed out that if it were possible, it would have nothing to do with humans. Namely, Gödel believed that the human mind is capable of formulating only a part of the mathematical intuition that it otherwise constantly uses while considering mathematical topics (Wang 2001, 184–185). One of the implications of Gödel's theorems is that every computer that operates with mathematical expressions sometimes generates certain truths (accuracies) that only we humans interpret without rigorous proof as such (not the computer itself), i.e., that the human brain somehow surpasses the computer itself in the sense that it can decide

on some theoretical questions that are unsolvable for a computer (Wang 2001, 184–185).[7] However, based on what has been proven so far, the possibility of the existence of a computer (machine) that would be equal to mathematical intuition as a whole cannot be ruled out. But even if such a computer existed, there is no way to demonstrate evidence through which a person would gain scientific confidence that it is indeed such a computer. More precisely, Gödel says that if there were a computer that operated with axioms that did not obey his (Gödel's) theorems, we humans would not be able to reliably recognize that computer in that capacity, i.e.. we would not know with mathematical certainty whether all of its outputs (solutions) were correct or not (Wang 2001, 185). Even if humanity could build such a computer, it would either not work properly or, even if it did work properly, we would not be able to understand its solutions (Nguyen 2024, 27).

These considerations of Gödel are of extraordinary importance for our topic, because if there is an unbridgeable epistemological gap between the capacity of human knowledge and the capacity of a hypothetical supermachine, then the question arises of any sense in investing in and researching AI that would supposedly have the answer to all of humanity's problems (including economic ones). Admittedly, someone like Yuval Harari could argue that today perhaps 1% of people understand how the financial system works, but that does not prevent 99% of the others from voluntarily obeying the "rules of the game" of that same system, regardless of the fact that they understand very little or nothing about them (Harari 2020). However, there is an essential difference between these two epistemological gaps. The first, which relates to our futuristic AI, is absolutely imperative and unbridgeable. There is no possibility that any individual (no matter how talented or diligent) will understand anything about how such an AI works. Moreover, others do not have this opportunity. If we were to accept the decisions of such a machine as logically authoritative, we would do so not on the basis of rational insight into its workings, but solely on the basis of faith or superstition

[7] A problem that is insoluble for a computer algorithm but is solvable for the human mind concerns, for example, the aforementioned ordinal preferences, which are not quantified in intensity but are simply our judgment of what we prefer in relation to something else. Ordinal preferences cannot be transposed into cardinal units in any way, because the evaluation of two states or objects is an absolutely physical and personal act that is not subject to projection into the outside world, an act that is perceived and understood only by the individual and that he cannot communicate as rationally understandable – much less instill – in any other human being (Nguyen 2024, 32; Mises 1998, 97).

in its superiority. Or, for fear of possible punishment if the machine were to operate in the environment of some technocratic dictatorship (Nguyen 2024, 37). On the other hand, the second epistemological gap is bridgeable for at least two reasons. First, it is partly a matter of personal choice how much one will become familiar with the functioning of the financial system. Many people choose to devote themselves to something else (instead of finance), although a relatively the same level of cognitive abilities is required for both. Second, even those people who do not want to or do not have the intellectual abilities to deal with finance can still gain a certain level of understanding of the financial system indirectly, precisely through people who possess such field of expertise. If they do not want or cannot make the effort to understand the basic outlines of the topic that financial experts briefly present to them, they can rely on their reputation (which they need as experts or politicians in order to obtain business success/political mandates).

At the end of this part of the paper, it is extremely important to respond to the objection of the advocates of central planning that even if we accept the argument that central planning with the help of information technologies (in our case, AI) is defective (because perfect equilibrium is impossible to establish), then this applies at least to the same extent to markets that are not regulated by a central plan (Cottrell, Cockshott and Michaelson 2009, 224). However, this would be correct only on the condition that we accept a reductionist understanding of the market that views economic relations exclusively through the prism of algorithmically solvable tasks. And as we have already implied above by mentioning the cultural context, economic relations largely escape algorithmic logic (Nguyen 2024, 38–39). Furthermore, as Koppl notes, a market that is not regulated by the center is not the same as an economic entity with a clearly defined goal in advance. It is only a framework within which many different goals exist that interact. In other words, while central planners must know in advance what exactly they are planning, a market free from a central plan does not plan anything in advance (Koppl 2010, 862). Therefore, the equilibrium of a market understood in this way is something completely different from the neoclassical one. It is established gradually and evolutionarily, in the very process of its emergence. The reality of life is revealed at least in part during the very implementation of a large number of plans (decentralization), and the plans are corrected on the fly, by adding missing elements that have been discovered in the meantime (Devereaux, Koppl and Kauffman 2024,

503; Koppl 2010, 862). Moreover, the correction or optimization of plans is not carried out only with the help of algorithmic methods, but with local knowledge that can be both tacit knowledge and exclusively intuitive, capable of being encoded, decoded, and transmitted only in the context of a specific culture (Devereaux*,* Koppl and Kauffman 2024, 492, 502). However, bearing in mind that even when this tacit knowledge is exhausted, individuals are left with residual knowledge in "their worlds" that are absolutely unique and beyond the reach of even local cultures (Devereaux*,* Koppl and Kauffman 2024, 502).

Therefore, no algorithmic technology can legitimately plead for central planning of the national economy, because it cannot solve the problems of self-reference that are inherent to man as an economic subject. Problems of self-reference can only be solved by man, but with other, intuitive tools that cannot be subsumed under the algorithm. If, however, planning with the help of AI is carried out by ignoring the problems of self-reference, this is possible only with the artificial reduction of economic life and man in abstract models that, by the very fact that they simplify reality, can only have extremely negative repercussions on that same reality. And even if at some point in the future AI (or any other technology) could adequately plan the national economy, there is no conceptual assumption that could bridge the epistemological gap between the solutions brought by AI and human cognitive abilities – at least while in the physical sense the machine representing AI and the human himself are two separate entities. It remains to be seen in the next part of the paper what the prospects are for the eventual merging of man and AI machine into some new entity, i.e., what implications this could have for the argument about the alleged justification of futuristic AI technologies in economic planning.

THE MERGING OF HUMANS AND ARTIFICIAL INTELLIGENCE – IS CENTRALIST ECONOMIC PLANNING JUSTIFIED AFTER ALL?

In order to understand the attempt of apology for centralized economic planning from the perspective of the projected integration of man and machine into a new entity, we must examine the most basic features of the predicted principles of operation of AI understood as superintelligence. Beyond its definition that we have already given, superintelligence can best be understood figuratively as a circle with

an outer shell and an inner core. The outer shell would consist of all precautionary measures, anticipated controls and regulations, or the most general frameworks (physical capacities and principles) that programmers would assign to superintelligence as a system of permanent or initial coordinates, within which all its activity generated by the aforementioned core would then flow. This core is precisely the internal logic of superintelligence.

The outer shell, therefore, contains a series of methods that can be divided into two groups. One consists of measures that programmers continuously take to "restrain" the internal logic of the superintelligence, i.e., its freedom of maneuver so that it does not go in this or that wrong direction for the welfare or even the survival of humanity (permanent coordinates), while the other consists of methods that would only give the initial impulse to the internal logic of the superintelligence (initial coordinates) from which it would then, without further intervention by programmers, somehow "arbitrarily" develop the trajectory of its actions on its own with full freedom of maneuver, but in such a way that the vital interests of humanity are not put in question (Bostrom 2014, 127–143). If the first group may play a greater role in the initial testing phases of the development of superintelligence, only the second group of methods allows the full potential of superintelligence to be exploited (Bostrom 2014, 185).

However, the initial coordinates can never cancel the fact that the internal logic of superintelligence will always be a kind of "black box", i.e., no one will be able to know how the superintelligence came to a certain solution, or what that solution will ultimately be – at least until the human-AI merger is complete (Bostrom 2014, 216). The "black box" is actually the aforementioned epistemological gap between AI and the human mind. Because superintelligence theorists assume and prefer to summarize all the initial coordinates into one central principle that could be reduced to the following: let the superintelligence do exactly what we as humans would do if we had its superhuman cognitive capacities (Bostrom 2014, 209–227; Russell 2019, 172–179, 314). In other words, there is no guarantee that a superintelligence in the economy would necessarily apply some superior method of central planning. It might even prescribe the inappropriateness of the plan and the continuation of the current market framework.

But even if we were to assume that AI programmers and architects could "direct" AI in the direction of central economic planning, by

precisely specifying permanent or initial coordinates in the outer shell, so that the final outcome would be some kind of centralist economic planning – the problem of the "black box", i.e., the epistemological gap between the AI computer and the human mind, would still not be solved. A person would still not be able to understand the solutions that AI has come up with. Furthermore, it is not clear how AI, physically separated from humans, could overcome the problems of self-reference that only humans are capable of, regardless of the amount of information and the processing speed of their machines.

It would seem that the only way out for the proponents of planning would be in some kind of direct integration (or at least indirect connection) of humans with the machines that represent AI. Although Elon Musk advocated this integration for the sake of better control over the work of AI, this reason is again purely epistemological in nature, as in our case of economic planning. Namely, in March 2023, Musk was one of the signatories of an open letter from numerous leaders from the AI industry that called for a moratorium on the development of AI, because the upcoming AI systems are a great risk, given the possibility that they will outnumber, outsmart, make us redundant and replace us – thereby risking losing control over the civilization in which we live (Samuel 2024). Supposedly, in order to follow the trajectory of AI development, we as humans must have improved cognitive abilities regarding the speed of data processing in our brains, at least indirectly, through a high-throughput brain interface that would read data from our brains at a speed identical to that already connected to AI systems. Musk has already begun experimenting with a brain implant (*Neuralink*) that makes direct contact with neurons via electrodes. It is still unclear to what extent it is possible to reduce the discrepancy between the algorithmic capacities of AI and humans through these attempts by Musk, but it is also unclear whether there is an alternative to the union of man and machine AI – if the epistemological gap is to be bridged.

However, let us assume that the epistemological gap has been resolved, that a seamless integration of humans with AI has been achieved, i.e., that humans can easily understand all of AI's decisions (and how it arrived at them). Moreover, let us assume that AI is now not only able to be understood by humans, but also that, through access to neurons (and therefore all of the sensations from the human's immediate environment), it can gain insight into everything that has so far eluded algorithms – into humans' complete intuition, tacit knowledge of the

local cultural context, and the ability to interpret all of the problems of self-reference. Would that then mean that all of the obstacles to effective centralist planning that we have outlined in this paper have been removed? In all likelihood, no. If AI were to acquire the ability to solve problems of self-reference, all humans would lose that ability, and AI would only be able to "solve" these problems maybe once. We have already seen above that self-reference is not a one-time, but a continuous process, full of dynamics and tension at the border between the subject of self-reference and its environment (culture) that surrounds it. Self-reference is not just a matter of a one-time interpretation, but above all a continuous, evolutionary reinterpretation by many subjects who, immersed in the cultural-local context, constantly change, complement, or abolish their decentralized economic plans of a "small range", without any total and rigid foresight in advance (before the process of implementing the plan itself begins in practice). Therefore, the fact that the AI would acquire the ability to interpret self-reference once would have practically no significance for the AI itself, if it is known that this was done at the cost of the disappearance of the entire context that is the *conditio sine qua non* for the reinterpretation of self-referential problems, and without which there is no effective economic planning. Context would disappear, because we have seen that its existence requires a large number of different self-referential subjects that interact, and here we would have only one single actor (AI). In other words, the problem of self-reference seems to be insoluble not only for AI machines that have the ability to work only algorithmically, but also for a futuristic AI that could hypothetically think non-algorithmically (in the form of some artificial amalgamation of all previous human self-referential interpretations into a single central type) – and this is only because centralizing self-reference would automatically mean paralysis of its reinterpretative capacities. Self-reference is not possible as a centralized plan, at least as a multiple process that can only make sense, but only as a highly decentralized system of numerous plans of self-conscious human subjects.

CONCLUSION

In this paper, we have shown, citing other authors, that AI cannot legitimately claim to be a centralized economic planning system, because it cannot solve the problems of self-reference as long as it operates according to algorithmic principles. These problems can only be solved

by numerous market participants who alone possess the intuition to reinterpret self-reference in the complex local cultural environment during the implementation of decentralized economic plans, and within an evolutionary process that is subject to constant changes, additions, and cancellations (replacements) – decentralized plans can never be completely prescribed in advance, as is the case with centralized ones. On the other hand, in the hypothetical example of Musk's projected integration of man and machine AI, we would see that even such an AI that might somehow acquire the ability to interpret self-reference once by exploiting and amalgamating all human thoughts and sensations, would actually be impotent to repeatedly reinterpret self-reference (and reinterpretation is the only true meaning of dealing with self-reference). Namely, by renouncing their right to self-reference in favor of some kind of centralist resultant embodied by AI, people would actually abolish the decentralized cultural framework of local knowledge. And self-reference cannot be repeatedly processed without such a decentralized framework, in some kind of vacuum of the exclusive monopoly position of AI that would produce the central plan.

In addition, it turned out that experimenting with the role of AI in a planned economy is not only unjustified from the point of view of pointless investment and waste of money. There are also ethical issues at stake. If the state decides to wholeheartedly embark on the path of experimenting and putting AI systems into operation for the purposes of a central economic plan, what will happen to individuals who refuse to be gradually or abruptly drawn into the new economic order? Will they be left with the opportunity to do business in the old way (and to what extent)? Or will they be punished, perhaps even eliminated as "obstacles" to a bright future led by "omnipotent" AI? All these and many other questions regarding AI are waiting to be answered by the scientific community and the general public.

REFERENCES

Boettke, Peter J. 2012. *Living Economics*. Oakland: The Independent Institute.

Bostrom, Nick. 2014. *Superintelligence – Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Cockshott, W. Paul, and Allin Cottrell. 1993. *Towards a New Socialism*. Nottingham: Spokesman Books.

Cottrell, A., Cockshott, W.P., and Greg Michaelson. 2009. "Is economic planning hypercomputational? The argument from Cantor diagonalisation." *International Journal of Unconventional Computing* 5 (3–4): 223–236.

Dapprich, Jan Philipp, and William Paul Cockshott. 2023. "Input-output planning and information." *Journal of Economic Behavior and Organization* 205 (2023): 412–422. DOI: 10.1016/j.jebo.2022.10.043

Devereaux, Abigail, Roger Koppl, and Stuart Kauffman. 2024. "Creative evolution in economics." *Journal of Evolutionary Economics* 34 (2024): 489–514.

Elliott, Anthony. 2022. *Making Sense of AI*. Cambridge: Polity Press.

Ertel, Wolfgang. 2017. *Introduction to Artificial Intelligence*. Cham: Springer.

Girasa, Rosario. 2020. *Artificial Intelligence as a Disruptive Technology*. Cham: Palgrave Macmillan.

Harari, Yuval. 2020. "Read Yuval Harari's blistering warning to Davos in full." *World Economic Forum*. Last accessed on June 21, 2025. https://www.weforum.org/stories/2020/01/yuval-hararis-warning-davos-speech-future-predications/

Hayek, Friedrich. 1952. *The Sensory Order*. Chicago: The University of Chicago Press.

Hayek, Friedrich. 1963. "Rules, Perception and Intelligibillity." *Proceedings of the British Academy* 48 (1962): 321–344.

Koppl, Roger. 2010. "Some epistemological implications of economic complexity." *Journal of Economic Behavior & Organization* 76 (2010): 859–872.

Lambert, Karras J., and Tate Fegley. 2023. "Economic Calculation in Light of Advances in Big Data and Artificial Intelligence." *Journal of Economic Behavior and Organization* 206 (2023): 243–250. DOI: 10.1016/j.jebo.2022.12.009

Luknar, Ivana. 2025. „Društvena dinamika u eri veštačke inteligencije." *Srpska politička misao* 90 (2): 145–162. DOI: 10.5937/spm90-56968

Mises, Ludwig. 1998. *Human Action*. Auburn: Ludwig von Mises Institute.

Moreno-Casas, Vicente, Victor I. Espinosa, and William Hongsong Wang. 2022. "The political economy of complexity: The case of cyber-communism." *Journal of Economic Behavior and Organization* 204 (2022): 566–580. DOI: 10.1016/j.jebo.2022.10.042

Nguyen, Hai-Trieu. 2024. "The Incompleteness of Central Planning." *Quarterly Journal of Austrian Economics* 27 (4): 22–43. DOI: 10.35297/001c.126016

Nieto, Maxi. 2023. "Economic Calculation, Complexity, and Cyber-Communism: Bad News for the Austrian School." *World Review of Political Economy* 14 (2): 234–262. DOI: 10.13169/worlrevipoliecon.14.2.0234

Phelan, Steven E., and Nikolai G. Wenzel. 2023. "Big Data, Quantum Computing, and the Economic Calculation Debate: Will Roasted Cyberpigeons Fly into the Mouths of Comrades?" *Journal of Economic Behavior and Organization* 206 (2023): 172–181. DOI:10.1016/j.jebo.2022.10.018

Rich, Elaine, Kevin Knight, and Shivashankar B. Nair. 2010. *Artificial Intelligence*. New Delhi: Tata McGraw Hill Education Private Limited.

Roughgarden, Tim. 2017. *Algorithms Illuminated (Part 1: The Basics)*. San Francisco: Soundlikeyourself Publishing.

Russell, Stuart. 2019. *Human Compatible – Artificial Intelligence and the Problem of Control*. New York City: Viking.

Samuel, Sigal. 2020. "Elon Musk wants to merge humans with AI. How many brains will be damaged along the way?" *Vox*. Last accessed on June 21, 2025. https://www.vox.com/future-perfect/23899981/elon-musk-ai-neuralink-brain-computer-interface

Van den Haag, Ernest. 1964. "The Planners and the Planned." In *Central Planning and Neomercantilism*, ed. Helmut Schoeck and James W. Wiggins, 19–39. New Jersey: D. van Nostrand Company.

Wang, Hao. 2001. *A Logical Journey: From Gödel to Philosophy*. Hong Kong: Asco Trade Typesetting.

Wang, William Hongsong, Victor I. Espinosa, and José Antonio Peña-Ramos. 2021. "Private Property Rights, Dynamic Efficiency and Economic Development: An Austrian Reply to Neo-Marxist Scholars Nieto and Mateo on Cyber-Communism and Market Process." *Economies* 9 (165): 1–17. DOI: 10.3390/economies9040165

Winrich, J. Steven. 1984. "Self-Reference and the Incomplete Structure of Neoclassical Economics." *Journal of Economic Issues* 18 (4): 987–1005.

Yampolskiy, Roman V. 2016. *Artificial Superintelligence – A Futuristic Approach*. Boca Raton: CRC Press-Taylor&Francis Group.

Марко Пејковић*

Институт за политичке студије, Београд

ВЕШТАЧКА ИНТЕЛИГЕНЦИЈА И ЦЕНТРАЛИСТИЧКО ЕКОНОМСКО ПЛАНИРАЊЕ

Сажетак

У овом раду, аутор ће настојати да разради и допуни већ постојеће аргументе о немогућности да вештачка интелигенција (ВИ) икада омогући (централизовано) ефикасно економско планирање на нивоу националне економије у виду државног плана. Нагли развој ВИ и осталих технологија, као да је последњих година поново пробудио дебату о надолазећим техничким капацитетима који би можда могли у скоријој будућности да превазиђу све слабости централистичког економског планирања познатих из претходног века. Међутим, недостатак концептуалних предуслова у домену ВИ који су неопходни за успешно централистичко планирање, заправо говори у прилог да је чак и било какав експеримент на ову тему и научно и економски неоправдан. Уводни део рада даје преглед главних појмова и дилема у оквиру дебате, наредни део разматра математичке парадоксе као концептуалне препреке за економско планирање уз помоћ ВИ (проблем самореференције), док ће завршни део рада бити критика покушаја идеолога-архитеката ВИ да те препреке савладају кроз евентуални спој човека и машина ВИ.

Кључне речи: вештачка интелигенција, централистичко економско планирање, тржиште, самореференција, Геделове теореме, алгоритми, суперинтелигенција

* Имејл: marko.pejkovic@ips.ac.rs; ORCID: 0009-0004-5920-1381

* This manuscript was submitted on June 30, 2025, and accepted by the Editorial Board for publishing on November 10, 2025

UDC 327:004.8:341.24:504.07(4-672EU)
DOI: 10.5937/spm94-59635
Original research article

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 45-65

Serdar Selim*

Department of Space Science and Technologies, Faculty of Science, Akdeniz University, Antalya

İsmail Yılmaz**

Department of Space Science and Technologies, Faculty of Science, Akdeniz University, Antalya

Pınar Kınıklı****

Department of Space Science and Technologies, Faculty of Science, Akdeniz University, Antalya

IMPLEMENTATION OF AI-BASED DETECTION INTO THE CLIMATE POLICY WITHIN THE EUROPEAN GREEN DEAL

Abstract

The European Green Deal outlines the European Union's roadmap for the green transition required by the Paris Climate Agreement. As part of its sustainable environmental policies, the European Green Deal aims to integrate digital transformation with the preservation of ecosystem services, the enhancement of green infrastructure, and the long-term sustainability of green networks. Green infrastructure contributes directly to the environmental objectives of the Green Deal by reducing carbon emissions, improving air quality, and conserving biodiversity. Therefore, accurately identifying, monitoring, and mapping green infrastructure is essential to achieving these goals.

* E-mail: serdarselim@akdeniz.edu.tr; ORCID: 0000-0002-5631-6253
** E-mail: isoylmz@gmail.com; ORCID: 0000-0001-7117-759X
*** E-mail: pgulyavuz@gmail.com; ORCID: 0000-0002-7572-1967

In this context, artificial intelligence (AI)-based automated tree detection systems, a rapidly advancing technology, play a critical role in fields such as forest management, biodiversity monitoring, and carbon footprint assessment. This study aims to support green policy objectives by automatically detecting tree communities in a specified region using AI algorithms applied to open-access satellite imagery. The research was conducted across three sample areas with varying environmental characteristics. The methodology integrates image-processing techniques with object detection algorithms, enabling high-accuracy classification of trees. The results contribute significantly to climate change mitigation efforts, carbon stock monitoring, smart urban planning, and the formulation of agricultural policies. Moreover, the proposed system can function as a decision support mechanism for public institutions, local governments, environmental scientists, and policymakers. In alignment with the European Green Deal's vision of digital green transformation, such AI-based applications hold substantial potential for enhancing environmental sustainability.

Keywords: artificial intelligence, green infrastructure, European Green Deal, digital green transformation, automated tree detection

INTRODUCTION

Climate change, one of the most pressing global challenges of the 21st century, is not limited to rising temperatures alone; it also encompasses a wide range of environmental consequences such as biodiversity loss, diminishing water resources, degradation of forest ecosystems, and the shrinking of natural habitats due to urbanization pressures (Gilman *et al.* 2010; Malanson and Alftine 2023). Scientific evidence clearly identifies anthropogenic greenhouse gas emissions as the primary driver of this transformation (Chen *et al.* 2023; Han *et al.* 2024), thereby reinforcing the urgency for environmentally sensitive, sustainable, and digitally supported new policy frameworks. In this context, the European Green Deal, announced by the European Union in 2019, provides a transformative framework for combating climate change and achieving environmental sustainability goals (Paleasri 2022; Boix-Fayos and de Vente 2023).

The European Green Deal outlines strategic objectives such as achieving a carbon-neutral continent, promoting the efficient use of natural resources, implementing circular economy practices, and preserving biodiversity (Samper *et al.* 2021; Skjærseth 2021; Knez *et al.* 2022). In the realization of these targets, digital transformation plays a key role not only in the domains of production and energy but also in environmental monitoring, planning, and management processes. This integrated approach, referred to as the digital green transition, aims to enable more accurate, timely, and data-driven environmental decision-making (Bertoncelj 2022). Within this framework, digital technologies such as artificial intelligence (AI), Remote Sensing (RS), Geographic Information Systems (GIS), and big data analytics are contributing significantly to the monitoring of environmental assets and changes, representing a paradigm shift in how environmental information is gathered, interpreted, and utilized (Himeur *et al.* 2022; Çağlar *et al.* 2025).

For green transition policies to be effectively implemented, the preservation and enhancement of green infrastructure – comprising trees, green spaces, forests, and other forms of natural vegetation – are of critical importance (Besley and Persson 2023). Green infrastructure is not merely an aesthetic component; it functions as a carbon sink, improves air quality, mitigates the urban heat island effect, and supports biodiversity, making it a vital environmental asset (Semeraro *et al.* 2021; Belčáková *et al.* 2022). In areas experiencing rapid urbanization, the strategic conservation of green infrastructure is a prerequisite for the development of climate-resilient cities (Kumar *et al.* 2023). However, traditional methods for monitoring these natural assets – such as field surveys and manual tree counting – are impractical, time-consuming, and costly, particularly when applied to large forested areas or in the context of limited resources (Himeur *et al.* 2022). In this regard, automated tree detection presents significant potential for the effective management of green infrastructure in both rural and urban landscapes. Accurate identification of tree count, distribution, and health status plays a crucial role in a variety of applications, including forest fire risk management, urban planning, carbon stock assessment, and biodiversity monitoring (Poláček *et al.* 2023; Capecchi *et al.* 2023). Due to the limitations of conventional techniques, AI-based image processing and object detection algorithms have become increasingly prevalent in recent years (Choi *et al.* 2022; Miranda *et al.* 2023; Heng

et al. 2024). Artificial intelligence technology, which is fundamentally based on machine learning and enables a machine to acquire knowledge through data processing, focuses on the development of algorithms and models that improve the performance and knowledge bases of computer systems (Luknar 2025). Among these models, the deep learning models have shown strong performance when applied to high-resolution imagery derived from satellite data, unmanned aerial vehicle (UAV) footage, and sensor-based systems (Jintasuttisak *et al.* 2022; Li *et al.* 2023). These models offer automated, high-accuracy detection capabilities, enabling both cost and time efficiencies while ensuring continuous, up-to-date environmental monitoring (Wu *et al.* 2021; Onishi and Ise 2021). These technological advancements are grounded in the integration of RS and GIS (Velasquez-Camacho *et al.* 2023). RS techniques enable large-scale, real-time monitoring and provide valuable insights into vegetation health, moisture levels, and stress indicators through infrared, multispectral, and hyperspectral imagery (Raihan *et al.* 2023). GIS, in turn, facilitates spatial analysis and multilayer visualization of this data, grounding environmental decision-making in a scientifically robust framework (Li *et al.* 2023). The integration of AI algorithms with these systems automates the entire workflow – from data analysis to geospatial mapping – thereby significantly enhancing the efficiency and accuracy of decision support mechanisms (Mohan and Giridhar 2022; Choi *et al.* 2023).

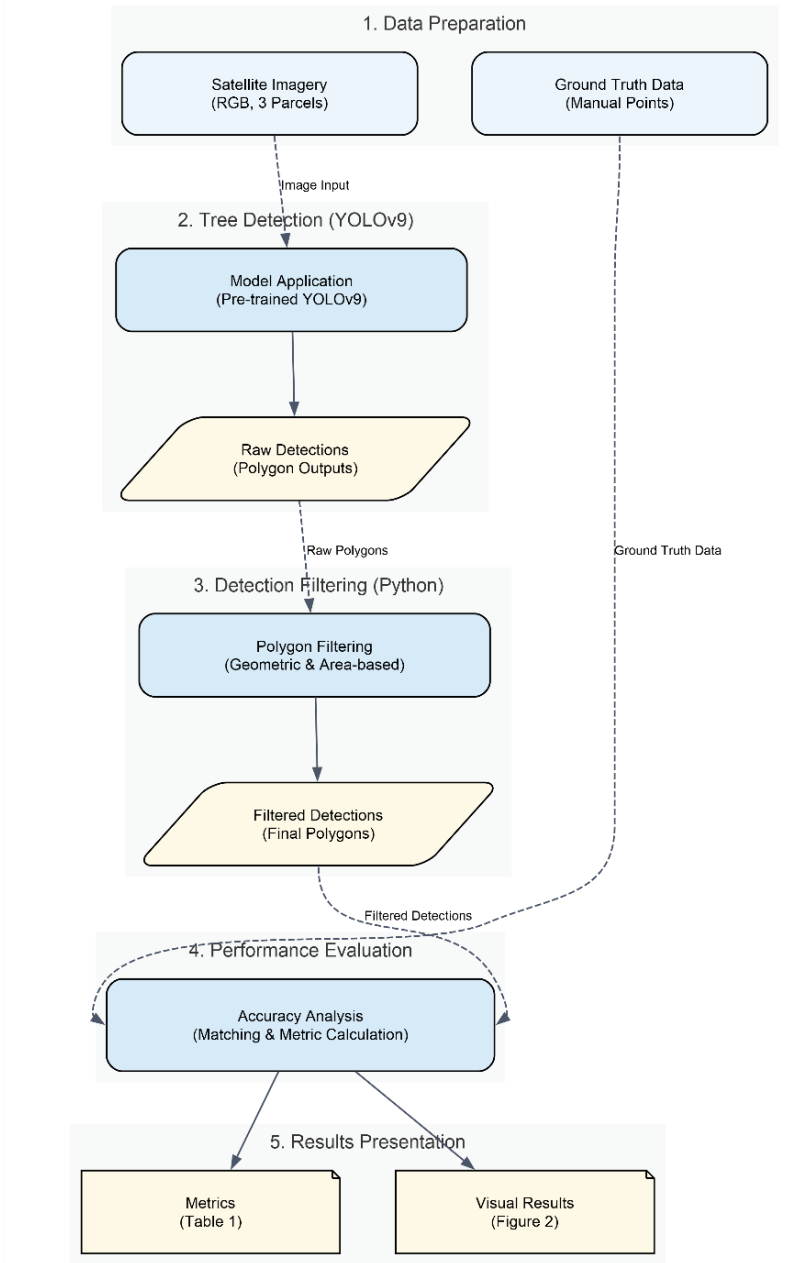
In line with the digital transformation vision of the European Green Deal, this study aims to support the conservation of green infrastructure and enhance environmental management through the automated detection of trees using artificial intelligence. In doing so, the study exemplifies the potential of technology-based solutions to contribute to the Green Deal's overarching goals of building a climate-neutral and environmentally friendly society. Although the Green Deal formally encompasses the 27-member states of the European Union (EU), its objectives are also expected to be supported by countries in cooperation with the EU (EU 2019; Kattelmann *et al.* 2021). Turkey, due to its geographical position and strong economic, political, and environmental ties with the EU, holds significant potential in contributing to the goals of the European Green Deal. While not an EU member, Turkey's participation in the Customs Union and its candidate country status position it in close alignment with EU environmental and climate policies. Consequently, the priorities set forth by the

Green Deal have a direct influence on Turkey's environmental policy landscape. In this context, the present study, conducted in Turkey, aligns with the country's commitment to contributing to these goals through its ratification of the Paris Agreement and its formulation of the Green Deal Action Plan. This national plan aims to enhance environmental sustainability, reduce carbon emissions, and promote the transition to a green economy (Ministry of Trade 2021). Therefore, the current study holds particular relevance in supporting Turkey's alignment with Green Deal objectives. The methodology presented in this article offers not only academic contributions but also practical implications. It aims to provide local governments, environmental planners, and conservation specialists with a scalable, verifiable, and sustainable tool for environmental monitoring.

MATERIAL AND METHOD

In this study, the detection and counting of trees in agricultural lands with varying soil characteristics were carried out using up-to-date and open-access Google satellite imagery, combined with an artificial intelligence-supported deep learning approach. The proposed methodology consists of four main steps: data preparation, model implementation, postprocessing of model outputs, and accuracy assessment (Figure 1).

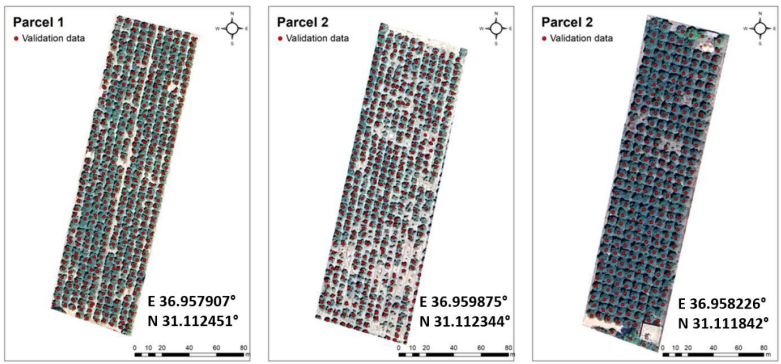
Figure 1. Method flowchart



Source: Authors

The study was conducted on three distinct agricultural parcels (Parcel 1, Parcel 2, and Parcel 3), each characterized by different ground conditions, geometrical configurations, and tree density patterns. Perennial agricultural lands were selected due to their dual significance in both economic productivity and ecological value. The trees within these parcels provide a wide range of ecosystem services, including carbon sequestration, oxygen production, urban heat island mitigation, and support for biodiversity, while also contributing to the economy through fruit yield. For tree detection, high-resolution RGB satellite imagery obtained from the Google Earth platform was utilized. To evaluate the performance of the model, the tree locations within each parcel were manually digitized and converted into point geometries to create ground truth datasets (Figure 2).

Figure 2. Study area data (The red dots represent the validation data)



Source: Authors

For the detection of tree crown centroids, the pre-trained "Tree-Tops Detection" model available within the "Model Zoo" collection of the QGIS Deepness plugin was employed. This model operates with an input resolution of 640x640 pixels and was trained on a custom dataset comprising a combination of various open-source image collections. It is built upon the YOLOv9 architecture as the underlying object detector, with the confidence threshold set at 0.10 and the Intersection over Union (IoU) threshold at 0.50 (Lu and Wang, 2024). YOLOv9, as the latest advancement in the YOLO series, introduces architectural innovations and enhanced training strategies that significantly improve detection accuracy while preserving real-time inference speeds (Wang *et al*. 2024; Chien *et al.* 2024). Evolved from the YOLOv7 framework, YOLOv9 is based on a novel module known as Generalized Efficient Layer Aggregation

Network (GELAN), which allows for optimized parameter utilization by leveraging more efficient convolutional operators. One of the most notable innovations in YOLOv9 is the integration of Programmatic Gradient Information (PGI), a newly proposed training enhancement framework (Zhang 2024). PGI improves gradient flow during training, mitigates information loss in deep neural networks, and enhances parameter learning through more targeted and enriched input data (Kumar *et al*. 2024). These architectural and algorithmic improvements result in measurable gains in both speed and accuracy when compared to previous YOLO versions (Ikmel and El Amrani 2024; Ye *et al.* 2024; Topgül *et al.* 2025). Initially, the model outputs tree crown detections as bounding polygons within the input satellite imagery. These raw polygon outputs (input_raw_polygon_path) were subjected to a sequence of post-processing steps using Python and the GeoPandas library to reduce noise and exclude non-tree objects. The post-processing included reprojection of all spatial data into a locally appropriate projected coordinate reference system (EPSG:32636) for accurate geometric calculations, removal of null or invalid geometries, area-based filtering using absolute or relative thresholds calculated within the projected CRS (a coordinate reference system that transforms geospatial data from the Earth's 3D surface to a 2D map projection), and additional filtering based on polygon aspect ratios. The resulting filtered polygons were exported as a shapefile (output_filtered_polygon_path) and used for further analysis in comparison with ground truth datasets across the sample parcels.

After the automatic detection of trees using artificial intelligence-supported deep learning algorithms, the verification phase was initiated. To quantify the detection performance of the filtered tree polygons, the metrics presented in Table 1 were calculated based on ground truth points. In this phase, matching was performed by checking whether a ground truth point was located within a filtered predicted polygon. The evaluated metrics include the Ground Truth Tree Count (GT), Estimated Number of Polygons (ENP), True Positives (TP), False Positives (FP), and False Negatives (FN), which are frequently used in the literature (Reddy and Karthikeyan 2022), along with derived measures such as Precision, Recall, and F1 Score. The corresponding equations were implemented within an automated workflow (Equations 1, 2, 3, and 4).

$$Accuracy = \frac{TP + TN}{\left(TP + TN + FP + FN\right)} \quad \text{(Equation 1)}$$

$$Precision = \frac{TP}{TP + FP} \quad \text{(Equation 2)}$$

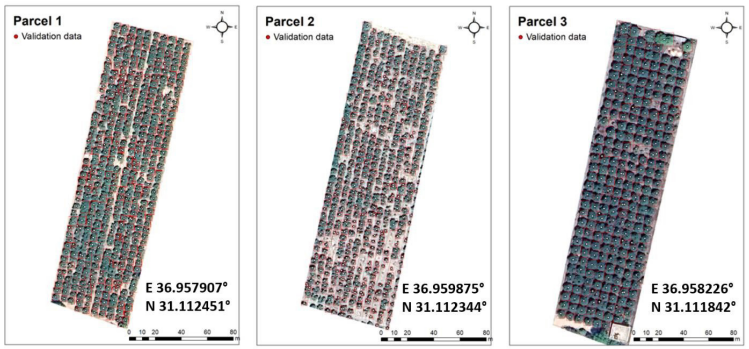
$$Recall = \frac{TP}{TP + FN} \quad \text{(Equation 3)}$$

$$F1\ Score = \frac{\left(2 \cdot Precision \cdot Recall\right)}{\left(Precision + Recal\right)} \quad \text{(Equation 4)}$$

RESULTS

In this study, a comprehensive analysis was conducted using artificial intelligence and deep learning algorithms to detect trees in three agricultural parcels with distinct morphological characteristics. The applied method was trained on high-resolution satellite imagery and field data, demonstrating applicability under heterogeneous conditions such as varying vegetation densities and terrain structures. The performance exhibited by the deep learning models yielded promising results for automatic tree detection. Model outputs were evaluated through accuracy analyses, with detection success assessed separately for each parcel. The mapping process was considered a critical stage for visually presenting tree detection results and examining model accuracy on an area-based scale. The maps shown in Figure 3 facilitate a comparative interpretation of the performance of the AI-supported deep learning method used for automatic tree detection under different parcel conditions, thereby visually supporting the applicability of the proposed approach.

Figure 3. Trees detected by deep learning algorithms after Python-based filtering (The blue dots represent the detected trees)



Source: Authors

The performance of the "Tree-Tops Detection" model used in this study was demonstrated across three different agricultural parcels using Google satellite imagery. Parcel 1 consists of trees planted on bare soil with partially regular row and spacing patterns, some gaps, and mostly similar crown diameters. The applied model successfully detected the regular groups but showed errors in areas with morphological variations. Parcel 2, also situated on bare soil, contains trees with mixed gaps and diverse shapes. The varying tree morphologies in this parcel notably increased the model's error rate in detection. Parcel 3 is characterized by a green ground cover with trees arranged in regular rows and spacing, exhibiting uniform shapes. In this area, the model achieved higher accuracy compared to the other parcels (Table 1).

Table 1. Automatic Tree Detection Performance Metrics for Different Parcels

Metrics	Parcel 1	Parcel 2	Parcel 3
True Tree Count (Ground Truth)	406,00	362,00	243,00
Estimated Number of Polygons (ENP)	474,00	511,00	263,00
True Positive (TP)	404,00	359,00	247,00
False Positive (FP)	88,00	158,00	21,00
False Negative (FN)	26,00	24,00	11,00
Precision	0,82	0,69	0,92
Recall	1,00	0,99	1,02
F1 Score	0,90	0,82	0,97

Accuracy: 0,78, 0,66, 0,88

Source: Authors

The model demonstrated F1 scores ranging from 0.82 to 0.97 across the three parcels. The highest F1 score (0.97) was observed in Parcel 3, while the lowest F1 score (0.82) was recorded in Parcel 2. Overall accuracy values were calculated as 0.78 for Parcel 1, 0.66 for Parcel 2, and 0.88 for Parcel 3. These performance variations among parcels are closely related to the characteristics of the terrain and tree cover. For example, in Parcel 3, which showed the highest performance, trees are more homogeneous, dense, and regularly spaced with clear contrast against the ground. In contrast, Parcel 2, which exhibited the lowest performance, contains sparser trees, variable ground conditions, and spectral/textural properties similar to tree crowns, resulting in an increased number of False Positives (FP). Parcel 1 demonstrated moderate performance, where the high number of False Positives reduced the precision value. Generally, the relatively low number of False Negatives (FN) across all parcels indicates that the model is capable of detecting existing trees but tends to produce more FPs as ground complexity increases and tree density decreases. The model's performance is influenced by factors such as tree density, homogeneity, contrast with the ground, and the presence of treelike objects on the terrain.

Limitations of the study include the use of a pre-trained general model without region-specific tuning, constraints related to the resolution and quality of Google satellite imagery, and the potential human error introduced by manual digitization of validation data. For future work, it is recommended to train the model with region-specific samples for increased accuracy, employ data augmentation techniques, utilize higher-resolution imagery, develop additional machine learning-based filtering layers alongside existing filtering steps, and test different deep learning architectures. For optimal model performance, high-resolution images captured during active growth periods, with minimal shading and a clear distinction between tree crowns and the ground, are ideal. When applying this model in urban green spaces, challenges such as heterogeneity in tree species and sizes, complex backgrounds, and shadows may arise; these challenges can be mitigated by expanding the training dataset and incorporating 3D data sources like LiDAR. In natural forests, dense and interwoven canopies, high species diversity, dense understory vegetation, and topographic shadows present significant challenges; here, higher-resolution drone or LiDAR data may provide viable solutions. In conclusion, the YOLOv9based "Tree-Tops Detection" model offers promising results for tree detection using Google satellite images; however, its performance varies according to site-specific conditions. Therefore, the

proposed recommendations for further model development and adaptation to different environments should be considered.

DISCUSSION AND CONCLUSION

Global climate change, as one of the most urgent environmental challenges of our time, necessitates the development of new strategies and technological solutions worldwide. In this context, the European Green Deal, proposed by the European Union with the goal of making the continent carbon-neutral by 2050, centers on policies that promote environmental sustainability (Selim 2021). Among these policies, agricultural and forest management, carbon sequestration, and the traceability of natural resources are of paramount importance (Keith *et al.* 2021). At this juncture, artificial intelligence–supported tree detection technologies, incorporating tools such as deep learning and machine learning, emerge as powerful instruments in achieving these objectives. Leveraging AI algorithms and satellite imagery, fruit trees in agricultural lands can be detected with high accuracy, enabling the creation of digital forest and agricultural maps (Yu *et al.* 2022; Gan *et al.* 2023). This technology holds significant potential not only for production planning but also for monitoring carbon stocks and reducing carbon footprints (Gaur *et al.* 2023; Manikandan *et al.* 2025). Trees, especially long-lived fruit trees, act as natural carbon sinks by absorbing atmospheric carbon dioxide (CO_2) through photosynthesis (Gelaye and Getahun 2023). Therefore, monitoring the number, species, and development status of these trees constitutes critical data for carbon balancing policies (Wambede *et al.* 2022).

Within the nature-based solutions framework advocated by the European Green Deal, detecting and increasing tree presence in agricultural lands is actively encouraged (Davies *et al.* 2021; Tanneberger *et al.* 2021). The sustainability of such solutions depends not only on tree planting but also on the accurate detection, monitoring, and management of these trees (Palomo *et al.* 2021). This study presents a practical method that combines open-source satellite imagery with AI algorithms for the automated, high-accuracy detection of trees. Using the proposed approach, rapid and practical quantification of tree counts in a given region is possible, growth trends can be identified, and yield predictions for agricultural production can be developed, thereby enabling the optimization of environmental and agricultural policies. The application of this technology also holds importance in terms of

transparency and accountability. The European Union's carbon markets, sustainable agricultural subsidies, and green financing instruments are shaped based on specific environmental outcomes. Data collected through AI provides verifiable evidence for such mechanisms, enhancing the effectiveness of climate policies.

Results obtained from this study demonstrate that using AI technologies for detecting trees in a region is faster and more cost-effective than traditional detection methods while achieving acceptable accuracy. Moreover, the use of AI specifically for automatic detection of fruit trees carries strategic value not only for agricultural productivity but also for combating climate change, carbon sequestration, and realizing the goals of the European Green Deal. The widespread adoption of this technology will contribute to building sustainable agricultural and forest management policies on digital foundations, thereby strengthening the balance between rural development and environmental conservation.

REFERENCES

Belčáková, Ingrid, Martina Slámová, and Zuzana Demovičová [Belčáková *et al.*]. 2022. "Importance of Urban Green Areas in the Context of Current and Future Global Changes: Lessons Learned from a Case Study in Bratislava (Slovakia)." *Sustainability* 14 (22): 14740. DOI: 10.3390/su142214740

Bertoncelj, Andrej. 2022. "Digital Transformation in the Context of European Union's Green Deal." *Amfiteatru Economic Journal* 24 (59): 5–7.

Besley, Timothy, and Torsten Persson. 2023. "The Political Economics of Green Transitions." *The Quarterly Journal of Economics* 138 (3): 1863–1906. DOI: 10.1093/qje/qjad008

Boix-Fayos, Carolina, and Joris de Vente. 2023. "Challenges and Potential Pathways towards Sustainable Agriculture within the European Green Deal." *Agricultural Systems* 207: 103634. DOI: 10.1016/j.agsy.2023.103634

Capecchi, Irene, Tommaso Borghini, and Iacopo Bernetti [Capecchi *et al.*]. 2023. "Automated Urban Tree Survey Using Remote Sensing Data, Google Street View Images, and Plant Species Recognition Apps." *European Journal of Remote Sensing* 56 (1): 2162441. DOI: 10.1080/22797254.2022.2162441

Chen, Liang, Zhe Chen, Yanbin Zhang, Yong Liu, Ahmed I. Osman, Mohamed Farghali, and Pow-Seng Yap [Chen *et al.*]. 2023. "Artificial Intelligence-Based Solutions for Climate Change: A Review." *Environmental Chemistry Letters* 21 (5): 2525–2557. DOI: 10.1007/s10311-023-01605-9

Chien, Ching-Tai, Ruo-Yu Ju, Kuan-Yu Chou, and Jyh-Shing Chiang [Chien *et al.*]. 2024. "YOLOv9 for Fracture Detection in Pediatric Wrist Trauma

X-Ray Images." *Electronics Letters* 60 (11): e13248. DOI: 10.1049/ell2.13248

Choi, Kwanghun, Wontae Lim, Byungjoo Chang, Jinwon Jeong, Ilhwan Kim, Chan Ryul Park, and Dowon Ko [Choi *et al.*]. 2022. "An Automatic Approach for Tree Species Detection and Profile Estimation of Urban Street Trees Using Deep Learning and Google Street View Images." *ISPRS Journal of Photogrammetry and Remote Sensing* 190: 165–180. DOI: 10.1016/j.isprsjprs.2022.06.005

Choi, Young. 2023. "GeoAI: Integration of Artificial Intelligence, Machine Learning, and Deep Learning with GIS." *Applied Sciences* 13 (6): 3895. DOI: 10.3390/app13063895

Davies, Clive, Wen Y. Chen, Giovanni Sanesi, and Raffaele Lafortezza [Davies *et al.*]. 2021. "The European Union Roadmap for Implementing Nature-Based Solutions: A Review." *Environmental Science & Policy* 121: 49–67. DOI: 10.1016/j.envsci.2021.03.018

European Union [EU]. 2019. "The European Green Deal." EUR-Lex. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2019:640:FIN, Last accessed on June 14, 2025.

Gaur, Loveleen, Anam Afaq, Gurinder K. Arora, and Nadeem Khan [Gaur *et al.*]. 2023. "Artificial Intelligence for Carbon Emissions Using System of Systems Theory." *Ecological Informatics* 76: 102165. DOI: 10.1016/j.ecoinf.2023.102165

Gelaye, Yitayal, and Solomon Getahun. 2024. "A Review of the Carbon Sequestration Potential of Fruit Trees and Their Implications for Climate Change Mitigation: The Case of Ethiopia." *Cogent Food & Agriculture* 10 (1): 2294544. DOI: 10.1080/23311932.2023.2294544

Gilman, Sarah E., Mark C. Urban, Joshua Tewksbury, George W. Gilchrist, and Robert D. Holt [Gilman *et al.*]. 2010. "A Framework for Community Interactions under Climate Change." *Trends in Ecology & Evolution* 25 (6): 325–331. DOI: 10.1016/j.tree.2010.03.002

Han, Hafeez, Zeeshan Zeeshan, Bilal A. Talpur, Tooba Sadiq, Uzair A. Bhatti, Emad M. Awwad, and Yaser Y. Ghadi [Han *et al.*]. 2024. "Studying Long-Term Relationship between Carbon Emissions, Soil, and Climate Change: Insights from a Global Earth Modeling Framework." *International Journal of Applied Earth Observation and Geoinformation* 130: 103902. DOI: 10.1016/j.jag.2024.103902

Heng, Qing, Song Yu, and Yan Zhang [Heng *et al.*]. 2024. "A New AI-Based Approach for Automatic Identification of Tea Leaf Disease Using Deep Neural Network Based on Hybrid Pooling." *Heliyon* 10 (5): e26925. DOI: 10.1016/j.heliyon.2024.e26925

Himeur, Yassine, Bharat Rimal, Aayush Tiwary, and Abbes Amira [Himeur *et al.*]. 2022. "Using Artificial Intelligence and Data Fusion for Environmental Monitoring: A Review and Future Perspectives." *Information Fusion* 86: 44–75. DOI: 10.1016/j.inffus.2022.06.003

Ikmel, Ghizlane, and Essaid L. El Amrani. 2024. "Performance Analysis of YOLOv5, YOLOv7, YOLOv8, and YOLOv9 on Road Environment Object Detection: Comparative Study." In *2024 International Conference on Ubiquitous Networking (UNet) 10*, 1–5. IEEE. DOI: 10.1109/UNet61038.2024.10678945

Jintasuttisak, Thitipong, E.A. Edirisinghe, and Amr Elbattay [Jintasuttisak *et al.*]. 2022. "Deep Neural Network Based Date Palm Tree Detection in Drone Imagery." *Computers and Electronics in Agriculture* 192: 106560. DOI: 10.1016/j.compag.2021.106560

Kattelmann, Felix, Jonas Siegle, Rafael Cunha Montenegro, Vanessa Sehn, Michael Blesl, and Ulrich Fahl [Kattelmann *et al.*]. 2021. "How to Reach the New Green Deal Targets: Analysing the Necessary Burden Sharing within the EU Using a Multi-Model Approach." *Energies* 14 (23): 7971. DOI: 10.3390/en14237971

Keith, Heather, Michael Vardon, Cédric Obst, Virginia Young, Richard A. Houghton, and Brendan Mackey [Keith *et al.*]. 2021. "Evaluating Nature-Based Solutions for Climate Mitigation and Conservation Requires Comprehensive Carbon Accounting." *Science of the Total Environment* 769: 144341. DOI: 10.1016/j.scitotenv.2020.144341

Knez, Snežana, Saša Štrbac, and Iztok Podbregar [Knez *et al.*]. 2022. "Climate Change in the Western Balkans and EU Green Deal: Status, Mitigation and Challenges." *Energy, Sustainability and Society* 12 (1): 1–14. DOI: 10.1186/s13705-021-00328-y

Kumar, Anil, Preeti Ekka, Meenakshi Upreti, Shilky, and Poonam Saikia [Kumar *et al.*]. 2023. "Urban Green Spaces for Environmental Sustainability and Climate Resilience." In *The Palgrave Handbook of Socio-ecological Resilience in the Face of Climate Change: Contexts from a Developing Country*, 389–409. Singapore: Springer Nature Singapore. DOI: 10.1007/978-981-99-5206-9_17

Kumar, Hemant, Praveen Mamoria, and Deepak K. Dewangan [Kumar *et al.*]. 2024. "Optimized Pedestrian Detection Leveraging YOLOv9: A Thorough Deep Learning Framework for Autonomous Vehicle System." In *2024 IEEE International Conference on Intelligent Signal Processing and Effective Communication Technologies (INSPECT)*, 1–6. IEEE. DOI: 10.1109/INSPECT62207.2024.10676945

Li, Rui, Guang Sun, Shuo Wang, Tianyi Tan, and Fei Xu [Li *et al.*]. 2023. "Tree Trunk Detection in Urban Scenes Using a Multiscale Attention-Based Deep Learning Method." *Ecological Informatics* 77: 102215. DOI: 10.1016/j.ecoinf.2023.102215

Li, Yang, Nobuo Yabuki, and Tomohiro Fukuda [Li *et al.*]. 2023. "Integrating GIS, Deep Learning, and Environmental Sensors for Multicriteria Evaluation of Urban Street Walkability." *Landscape and Urban Planning* 230: 104603. DOI: 10.1016/j.landurbplan.2022.104603

Lu, Dawei, and Yuxiao Wang. 2024. "MAR-YOLOv9: A Multi-Dataset Object Detection Method for Agricultural Fields Based on YOLOv9." *PLoS ONE* 19 (10): e0307643. DOI: 10.1371/journal.pone.0307643

Luknar, Ivana. 2025. "Social Dynamics in the Era of Artificial Intelligence." *Serbian Political Thought* 90 (2): 145–162. DOI: 10.22182/spt.9022025.1

Malanson, George P., and Kathryn J. Alftine. 2023. "Ecological Impacts of Climate Change." In *Biological and Environmental Hazards, Risks, and Disasters*, 449–479. Elsevier. DOI: 10.1016/B978-0-12-820509-9.00017-8

Manikandan, Sivasubramanian, Rangarajan Sindhu Kaviya, Dhamodharan Hemnath Shreeharan, Ramasamy Subbaiya, Sundaram Vickram, Natchimuthu Karmegam, Woong Kim and Muthusamy Govarthanan [Manikandan *et al.*]. 2025. "Artificial Intelligence-Driven Sustainability: Enhancing Carbon Capture for Sustainable Development Goals—A Review." *Sustainable Development* 33 (2): 2004–2029. DOI: 10.1002/sd.3222

Ministry of Trade [Ministry of Trade], Republic of Turkey. 2021. *Green Deal Action Plan 2021*. Ankara: Ministry of Trade.

Miranda, Juan C., Jordi Gené-Mola, Maren Zude-Sasse, Nikolaos Tsoulias, Alexandre Escolà, Jaume Arnó, and Eduard Gregorio [Miranda *et al.*]. 2023. "Fruit Sizing Using AI: A Review of Methods and Challenges." *Postharvest Biology and Technology* 206: 112587. DOI: 10.1016/j.postharvbio.2023.112587

Mohan, Shyama, and M.V.S.S. Giridhar. 2022. "A Brief Review of Recent Developments in the Integration of Deep Learning with GIS." *Geomatics and Environmental Engineering* 16 (2): 21–38. DOI: 10.7494/geom.2022.16.2.21

Onishi, Masaki, and Takeshi Ise. 2021. "Explainable Identification and Mapping of Trees Using UAV RGB Image and Deep Learning." *Scientific Reports* 11 (1): 903. DOI: 10.1038/s41598-020-79653-9

Paleari, Stefano. 2022. "The Impact of the European Green Deal on EU Environmental Policy." *The Journal of Environment & Development* 31 (2): 196–220. DOI: 10.1177/10704965221080596

Palomo, Ignacio, Bruno Locatelli, Ines Otero, Matthew Colloff, Emilie Crouzat, Aida Cuni-Sanchez, and Sandra Lavorel [Palomo *et al.*]. 2021. "Assessing Nature-Based Solutions for Transformative Change." *One Earth* 4 (5): 730–741. DOI: 10.1016/j.oneear.2021.04.013

Poláček, Martin, Alejandro Arizpe, Philipp Hüther, Lukas Weidlich, Sebastian Steindl, and Kerrin Swarts [Poláček *et al.*]. 2023. "Automation of Tree-Ring Detection and Measurements Using Deep Learning." *Methods in Ecology and Evolution* 14 (9): 2233–2242. DOI: 10.1111/2041-210X.14162

Raihan, Alamgir. 2023. "A Comprehensive Review of the Recent Advancement in Integrating Deep Learning with Geographic Information Systems." *Research Briefs on Information and Communication Technology Evolution* 9: 98–115. DOI: 10.56801/rebicte.v9i.155

Reddy, B. Harish, and P. Raja Karthikeyan. 2022. "Classification of Fire and Smoke Images Using Decision Tree Algorithm in Comparison with Logistic Regression to Measure Accuracy, Precision, Recall, F-Score." In *2022 14th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*, 1–5. IEEE. DOI: 10.1109/MACS56771.2022.10022741

Samper, Jon Ander, Anna Schockling, and Mine Islar [Samper *et al.*]. 2021. "Climate Politics in Green Deals: Exposing the Political Frontiers of the European Green Deal." *Politics and Governance* 9 (2): 8–16. DOI: 10.17645/pag.v9i2.3878

Selim, Serdar. 2021. "Integration of Urban Green Lands into Green Infrastructure System in the Framework of European Green Deal: The Case Study of Konyaaltı-Antalya, Türkiye." *Süleyman Demirel University Journal of Natural and Applied Sciences* 25 (3): 636-643. DOI: 10.19113/sdufenbed.970345

Semeraro, Teodoro, Alessandro Scarano, Riccardo Buccolieri, Angelo Santino, and Eeva Aarrevaara. 2021 [Semeraro *et al.*]. "Planning of Urban Green Spaces: An Ecological Perspective on Human Benefits." *Land* 10 (2): 105. DOI: 10.3390/land10020105

Skjærseth, Jon Birger. 2021. "Towards a European Green Deal: The Evolution of EU Climate and Energy Policy Mixes." *International Environmental Agreements: Politics, Law and Economics* 21 (1): 25–41. DOI: 10.1007/s10784-020-09492-4

Tanneberger, Franziska, Laura Appulo, Sabine Ewert, Sebastian Lakner, Niamh Ó Brolcháin, Jan Peters, and Wendelin Wichtmann [Tanneberger

et al.]. 2021. "The Power of Nature-Based Solutions: How Peatlands Can Help Us to Achieve Key EU Sustainability Objectives." *Advanced Sustainable Systems* 5 (1): 2000146. DOI: 10.1002/adsu.202000146

Topgül, Şeyma N., Emre Sertel, Serdar Aksoy, Cem Ünsalan, and Johan E.S. Fransson [Topgül *et al.*]. 2025. "VHRTrees: A New Benchmark Dataset for Tree Detection in Satellite Imagery and Performance Evaluation with YOLO-Based Models." *Frontiers in Forests and Global Change* 7: 1495544. DOI: 10.3389/ffgc.2024.1495544

Velasquez-Camacho, Luisa, Mikel Etxegarai, and Sergio de-Miguel [Velasquez-Camacho *et al.*]. 2023. "Implementing Deep Learning Algorithms for Urban Tree Detection and Geolocation with High-Resolution Aerial, Satellite, and Ground-Level Images." *Computers, Environment and Urban Systems* 105: 102025. DOI: 10.1016/j.compenvurbsys.2023.102025

Wambede, Moses N., Grace Akello, Joseph S. Lugumira, Bernard Barasa, Denis Amwonya, and Andrew Mulabbi [Wambede *et al.*]. 2022. "Carbon Sequestration of Fruit Trees under Contrasting Management Regimes." *Indonesian Journal of Geography* 54 (3): 387–394.

Wang, Chien-Yao, I-Hau Yeh, and Hong-Yuan Mark Liao [Wang *et al.*]. 2024. "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information." In *European Conference on Computer Vision*, 1–21. Cham: Springer Nature Switzerland. DOI: 10.1007/978-3-031-72658-3_29

Wu, Bin, Anzhu Liang, Hongwei Zhang, Ting Zhu, Zihan Zou, Dongdong Yang, and Jianhua Su [Wu *et al.*]. 2021. "Application of Conventional UAV-Based High-Throughput Object Detection to the Early Diagnosis of Pine Wilt Disease by Deep Learning." *Forest Ecology and Management* 486: 118986. DOI: 10.1016/j.foreco.2021.118986

Ye, Rui, Guofa Shao, Qingwei Gao, Hao Zhang, and Tianyi Li [Ye *et al.*]. 2024. "CR-YOLOv9: Improved YOLOv9 Multistage Strawberry Fruit Maturity Detection Application Integrated with CRNET." *Foods* 13 (16): 2571. DOI: 10.3390/foods13162571

Yu, Kai, Zhili Hao, Christopher J. Post, Elena A. Mikhailova, Lili Lin, Guohao Zhao, and Jinxun Liu [Yu *et al.*]. 2022. "Comparison of Classical Methods and Mask R-CNN for Automatic Tree Detection and Mapping Using UAV Imagery." *Remote Sensing* 14 (2): 295. DOI: 10.3390/rs14020295

Zhang, Chen. 2024. "Road Object Perception Method Based on YOLOv9." In *Proceedings of the 7th International Conference on Information Technologies and Electrical Engineering*, 182–188. DOI: 10.1145/3665780.3665823

Сердар Селим[*]

Одељење за свемирску науку и технологије, Природно-математички факултет, Универзитет Акдениз, Анталија

Исмаил Јилмаз[**]

Одељење за свемирску науку и технологије, Природно-математички факултет, Универзитет Акдениз, Анталија

Пинар Киникли[***]

Одељење за свемирску науку и технологије, Природно-математички факултет, Универзитет Акдениз, Анталија

ИМПЛЕМЕНТАЦИЈА ДЕТЕКЦИЈЕ ЗАСНОВАНЕ НА ВЕШТАЧКОЈ ИНТЕЛИГЕНЦИЈИ У КЛИМАТСКУ ПОЛИТИКУ У ОКВИРУ ЕВРОПСКОГ ЗЕЛЕНОГ ДОГОВОРА

Резиме

Европски зелени договор представља план Европске уније за зелену транзицију коју захтева Париски споразум о климатским променама. Као део својих одрживих политика заштите животне средине, Европски зелени договор има за циљ да интегрише дигиталну трансформацију са очувањем екосистемских услуга, унапређењем зелене инфраструктуре и дугорочном одрживошћу зелених мрежа. У том контексту, аутоматизовани системи за детекцију дрвећа засновани на вештачкој интелигенцији (ВИ), технологија која се брзо развија, играју кључну улогу у областима као што су управљање шумама, праћење биодиверзитета и процена угљеничног отиска. Ова студија има за циљ да подржи циљеве зелене политике аутоматским откривањем заједница дрвећа у одређеном региону користећи алгоритме ВИ примењене на сателитске снимке отвореног приступа. Методологија интегрише

[*] Имејл: serdarselim@akdeniz.edu.tr; ORCID: 0000-0002-5631-6253
[**] Имејл: isoylmz@gmail.com; ORCID: 0000-0001-7117-759X
[***] Имејл: pgulyavuz@gmail.com; ORCID: 0000-0002-7572-1967

технике обраде слика са алгоритмима за детекцију објеката, омогућавајући високопрецизну класификацију дрвећа. Детекција и бројање дрвећа на пољопривредним земљиштима са различитим карактеристикама земљишта спроведени су коришћењем ажурних и отворено доступних сателитских снимака компаније *Google*, у комбинацији са приступом дубоког учења подржаним вештачком интелигенцијом. Предложена методологија се састоји од четири главна корака: припрема података, имплементација модела, постпроцесна обрада излаза модела и процена тачности. За детекцију центроида круна дрвећа коришћен је претходно обучени модел "Tree-Tops Detection" доступан у оквиру колекције "Model Zoo" додатка *QGIS Deepness*. Након аутоматске детекције дрвећа коришћењем алгоритама дубоког учења подржаних вештачком интелигенцијом, започета је фаза верификације. У овој фази, упаривање је извршено провером да ли се тачка истинитости налази унутар филтрираног предвиђеног полигона. Процењивана метрика укључује Истинитост броја стабала на терену (GT), процењени број полигона (ENP), истинито позитивне резултате (TP), лажно позитивне резултате (FP) и лажно негативне резултате (FN), заједно са изведеним мерама као што су прецизност, подсећање и F1 резултат. Резултати су показали да се оцене F1 крећу од 0,82 до 0,97 на све три парцеле. Највећа F1 оцена (0,97) примећена је на парцели 3, док је најнижа оцена F1 (0,82) забележена на парцели 2. Укупне вредности тачности израчунате су као 0,78 за парцелу 1, затим 0,66 за парцелу 2 и 0,88 за парцелу 3. Ове варијације у перформансама међу парцелама су уско повезане са карактеристикама терена и покривача дрвећем. Генерално, релативно мали број лажно негативних (FN) резултата на свим парцелама указује на то да је модел способан да детектује постојеће дрвеће, али има тенденцију да производи више FN резултата како се сложеност тла повећава, а густина дрвећа смањује. На перформансе модела утичу фактори као што су густина дрвећа, хомогеност, контраст са тлом и присуство објеката налик дрвећу на терену. Ограничења студије укључују употребу претходно обученог општег модела без специфичних подешавања за одређени регион, ограничења везана за резолуцију и квалитет сателитских снимака компаније *Google* и потенцијалну људску грешку коју уводи ручна дигитализација података о валидацији. За будући рад, препоручује се обука модела са

узорцима специфичним за одређени регион ради повећане тачности, коришћење техника проширења података, коришћење снимака веће резолуције, развој додатних слојева филтрирања заснованих на машинском учењу поред постојећих корака филтрирања и тестирање различитих архитектура дубоког учења. Закључно, модел "Tree-Tops Detection" базиран на *YOLOv9* нуди обећавајуће резултате за детекцију дрвећа коришћењем *Google* сателитских снимака. Међутим, његове перформансе варирају у зависности од услова специфичних за локацију. Стога треба размотрити предложене препоруке за даљи развој модела и прилагођавање различитим окружењима. Резултати добијени овом студијом показују да је коришћење технологија ВИ за откривање дрвећа у региону брже и исплативије од традиционалних метода откривања, уз постизање прихватљиве тачности. Штавише, употреба ВИ посебно за аутоматско откривање воћака носи стратешку вредност не само за пољопривредну продуктивност, већ и за борбу против климатских промена, секвестрације угљеника и остваривања циљева Европског зеленог договора. Широко усвајање ове технологије допринеће изградњи одрживих политика пољопривредног и шумарског управљања на дигиталним темељима, чиме ће се ојачати равнотежа између руралног развоја и очувања животне средине.

Кључне речи: вештачка интелигенција, зелена инфраструктура, Европски зелени договор, дигитална зелена трансформација, аутоматизовано откривање дрвећа

* This manuscript was submitted on June 24, 2025, and accepted by the Editorial Board for publishing on November 10, 2025.

UDC 327:004.8:351.78(510:73)"20"
DOI: 10.5937/spm94-59859
Review article

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 67-90

Vanja Glišin*

Institute for Political Studies, Belgrade

ARTIFICIAL INTELLIGENCE IN THE SECURITY STRATEGIES OF THE USA AND CHINA AT THE BEGINNING OF THE 21ST CENTURY – THE STRUGGLE FOR GLOBAL DOMINANCE**

Abstract

The development of artificial intelligence (AI) and its increasing application across various sectors of the state and society necessitates monitoring, studying, and actively adopting knowledge and conclusions on the subject across all scientific fields. In recent decades, we've observed a comprehensive use of AI in the security strategies of major powers like the United States and China, resulting in a sort of race in developing new technologies and their application. The first part of the paper defines AI, outlines various classifications, and focuses on its use in the security sector. We then analyze the security strategies of the USA and China to identify the role of AI and plans for its development in the global power struggle. Based on this, in the last part of the paper, we present the personalities and differences between the security strategies of the US and China.

Keywords: artificial intelligence, USA, China, security, strategy, Security Stategy

* E-mail: vanja.glisin@ips.ac.rs; ORCID: 0000-0002-7747-6205

** This paper was written within the research activity of the Institute for Political Studies, funded by the Ministry of Science, Technological Development, and Innovation of the Republic of Serbia.

INTRODUCTION

In recent decades, we have witnessed the rapid development of technologies, which requires active monitoring, analysis and adaptation of political, economic, business, military-technological, social and other systems of a state. Today, there is widespread talk about artificial intelligence, which inevitably affects all segments of the state and society, whether positive or negative aspects. Hence the need to analyze artificial intelligence multidimensionally, from the perspective of social, humanities, natural, technical, military and other sciences. This reflects the scientific justification for dealing with this topic, but no less social, since the development of artificial intelligence also concerns societies around the world.

Guided by the methodology from the general to the specific, the focus in this paper will be on artificial intelligence in the security strategies of the United States of America and China. These are the countries that are most actively developing artificial intelligence in the security sector, which is clearly highlighted in the security strategies that will be the subject of analysis. The intention is to present the key segments of artificial intelligence in the security strategies of the two countries, in order to point out the similarities and differences on this basis. We will use the method of analysis (descriptive role) to describe the research problem and explain (explanatory role) the key segments related to artificial intelligence in security strategies. Then, taking into account the described and explained phenomena and processes, we will strive to understand the complex whole using the synthesis method, primarily the potential of artificial intelligence within security strategies. Through a qualitative analysis of relevant domestic and foreign literature, available documents and reports, we will collect data related to the subject of research.

ARTIFICIAL INTELLIGENCE – CONCEPT AND DEFINITION

The American mathematician and computer scientist John McCarthy first used the term Artificial Intelligence – AI, to introduce it as an academic discipline in 1956 at a conference at Dartmouth College in the United States. According to McCarthy, AI is “the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers

to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable" (McCarthy 2007). Due to the complexity of the term, we encounter different explanations and interpretations of AI in the literature, especially when looking at different scientific research fields, which makes it difficult to specify the term and its meaning (Russell and Norvig 2010; Scherer 2016). This is understandable, given that AI is "an imitation or simulation of something that we ourselves do not yet fully understand: human intelligence" (Sheikh, Prins and Schrijvers 2023, 16). The High-Level Expert Group on Artificial Intelligence (AI HLEG) defines artificial intelligence as "systems that exhibit intelligent behavior by analyzing their environment and taking actions – with a certain degree of autonomy – to achieve specific goals" (High-Level Expert Group on Artificial Intelligence 2019). In order to understand AI more precisely, we also provide the following definition: "Artificial Intelligence technologies aim to reproduce or surpass abilities (in computational systems) that would require 'intelligence' if humans were to perform them. These include: learning and adaptation; sensory understanding and interaction; reasoning and planning; optimisation of procedures and parameters; autonomy; creativity; and extracting knowledge and predictions from large, diverse digital data" (McKendrick 2019, 6).

Therefore, AI has the ability to learn and adapt, reason and plan (Shalev-Shwartz and Ben-David 2014), with the aim of creating "learning system structures that will represent a self-learning system", without the need for a human to feed the machine with data strings (Radun 2025, 239). The phrase "machine learning" refers to machines capable of learning without the participation of the human factor in programming and filling in data (Smola and Vishwanathan 2008).

Artificial intelligence can be classified according to numerous criteria, which also depends on scientific disciplines and theoretical approaches (Luknar 2024). One of the basic criteria is the area in which AI is applied. Therefore, we can talk about: 1) specialized AI - used in certain areas; 2) general AI – as a general intelligence capability that can be applied in various fields (economy, business, education, medicine, military industry, etc.), and 3) super-intelligence (Kaplan and Haenlein 2018; Wang and Siau 2019).

Alex Bekker offered a more complex typology of AI:

1) interactive AI (example: personal assistants such as Siri, Cortana and Alexa);

2) functional AI (robots);
3) analytical AI (data analysis, machine learning);
4) textual AI (text recognition, speech-to-text conversion);
5) visual AI (augmented reality technology) (Bekker 2019).

The application of AI is becoming more widespread, both in the military and civilian sectors. Considering that the level of development of AI in the military sector is significantly higher than in the civilian sector, we can only assume what all the possibilities and its scope are. If we give examples of the combination of AI with other technological branches, we will see how ubiquitous it is in everyday life. By combining AI and robotics, intelligent robots capable of performing various tasks have been created, such as industrial robots, smart androids, actroids, robots in the military and police, as well as in medicine, pharmacology, and other fields (Despotović and Glišin 2024; Luknar 2024). AI is also an integral part of autonomous systems, such as unmanned aerial vehicles or drones, unmanned vessels, and unmanned land vehicles (Đorić and Glišin 2023).

ARTIFICIAL INTELLIGENCE IN THE SECURITY SECTOR AND SECURITY STRATEGIES

Artificial intelligence is becoming a central pillar of modern security strategies and is used in many ways in the security sector, especially when it comes to preserving the security state and the role of the military. Strategic documents include all aspects of the security state, and AI is a step further in analysis and prediction. With this in mind, AI is increasingly used in the analysis and processing of intelligence data, scenario assessment, strategy preparation, logistics and the conduct of combat operations, as well as in the protection and maintenance of systems. Thus, the advantages of AI and machine learning are reflected in: system autonomy, predictive analytics, cybersecurity and realistic simulation of combat operations and security threats (Cohen 2023; Ro 2023). We would highlight predictive analytics as one of the most important elements in the preparation of offensive and defensive actions (Chappell 2020; Márquez-Díaz 2024; Galán, Carrasco and LaTorre 2022).

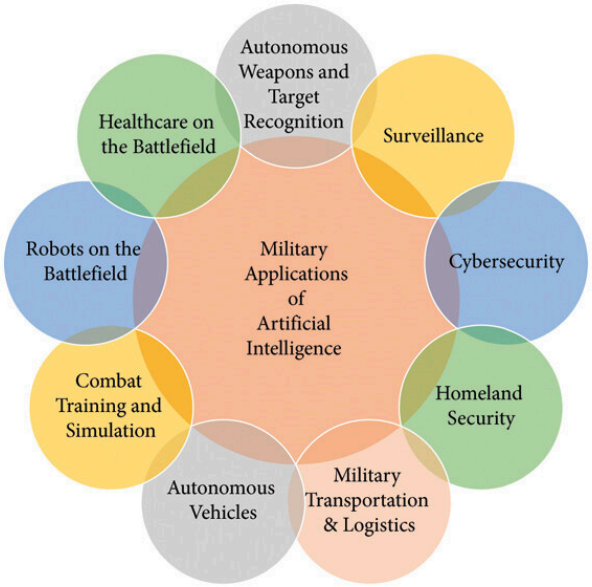
AI is also applied in the production of security and weapons systems, as well as in their upgrade, especially when it comes to unmanned aerial vehicles, robots and vessels (Evron and Bitzinger

2023). The most developed military powers such as the USA, China and Russia use AI to improve the accuracy and efficiency of military missions, the timeliness of defense and system protection, data processing and precision of actions on the battlefield. The US National Security Strategy document from 2022 states that technological superiority is key to national defense (White House 2022). China's position is the same, as they see the integration of AI as a path to strategic autonomy in the military-technological sector (Kania and Costello 2021).

The development of modern unmanned systems has contributed to the reduction of the use of human capacities, especially in security-risk situations and inaccessible areas. The advantage of autonomous weapons systems is their independence in actions, and this is especially evident in complex operations where human monitoring and guidance are not possible. The basic division of AOS is: 1. unmanned aerial vehicles; 2. unmanned underwater vehicles; 3. unmanned land vehicles and 4. combat robots (Đorić and Glišin 2023). The use of the aforementioned autonomous systems is multiple, since they have the ability to scout, recognize, process in real time and operate in different circumstances. For all the reasons listed above, AI has an increasingly important place in the strategic documents of major powers.

In contrast to AI in the civilian sector, AI in the security sector is significantly more advanced (Figure 1) (it is assumed to be 15 years or more ahead of its time) and is developed according to strict procedures and requirements (Despotović and Glišin 2024). The reliability of AI-based systems is very important, especially when they are exposed to hostile physical and cyber attacks (Gaba *et al.* 2024; Takpah and Oriakhi 2025). Then, a large data processing capacity is required, fast and accurate. Also, the autonomy of work without the need for human supervision and adaptability to different circumstances are very important items (Rashid *et al.* 2023). On the other hand, human-machine interaction in a security environment is necessary, so that the human factor makes a decision based on machine analysis in real time (Márquez-Díaz 2024; Balbaa and Abdurashidova 2024). Finally, interoperability, i.e. the ability of different systems to work together, is essential (Wegner 1996; Koncová and Kremeňová 2022). The application of AI in the defense sector is shown in Figure 1, which confirms the trend of increasingly widespread use of AI in almost all segments of military activities.

Figure 1. Applications of artificial intelligence in the defense sector



Source: Rashid *et al.* 2023.

AI in US security strategies

The United States has been investing the most money in research and development of technology globally since the end of World War II. This is especially true in recent decades, when the focus has been on technological innovations based on AI. The US budget for AI in 2025 is $470.9 billion, which is almost four times more than China, which is in second place with a budget of $119.3 billion (Spherical Insights 2025). If we look at the list of ten countries that allocate the most money for AI, we will see that the US invests more money than the other nine countries combined (Spherical Insights 2025). Therefore, the US attaches special importance to AI, which is also reflected in the security strategies that we will analyze below.

The United States established the American AI Initiative in 2019 by signing Executive Order 13859, which established an institutional framework for coordinating AI research and development in the civilian, military, and intelligence sectors. The document states: "Continued

American leadership in AI is of paramount importance to maintaining the economic and national security of the United States and to shaping the global evolution of AI in a manner consistent with our Nation's values, policies, and priorities" (Federal Register 2019). The American AI Initiative receives resources from the Federal Government so that it can implement projects in the field of new technologies. It is important to note that the AI development strategy involves several items. First, it is necessary to invest in research through cooperation between the state, industry, professors, scientists, international partners, and other entities with the aim of achieving global dominance in the field of AI. US President Donald Trump has proposed doubling funding for AI research and development for fiscal year 2021, which testifies to the importance of high-tech projects. Second, the American AI Initiative calls for increased funding and improved conditions for AI development (Federal Register 2019). Third, it is necessary to remove barriers to innovation in the field of AI. The White House has proposed the United States AI Regulatory Principles, "a first-of-its-kind national AI regulatory policy that advances innovation underpinned by American values" (Parker 2020). Fourth, it is important to promote American innovation in the field of AI globally, attract those countries that support the US, and open markets for the American AI industry. Fifth, incorporating proven and trusted AI into the government system to increase efficiency (Federal Register 2019). Sixth, training personnel to use AI in various sectors. Therefore, "these AI regulatory principles put our Nation on a path towards continued AI leadership, innovation, and discovery" (Parker 2020).

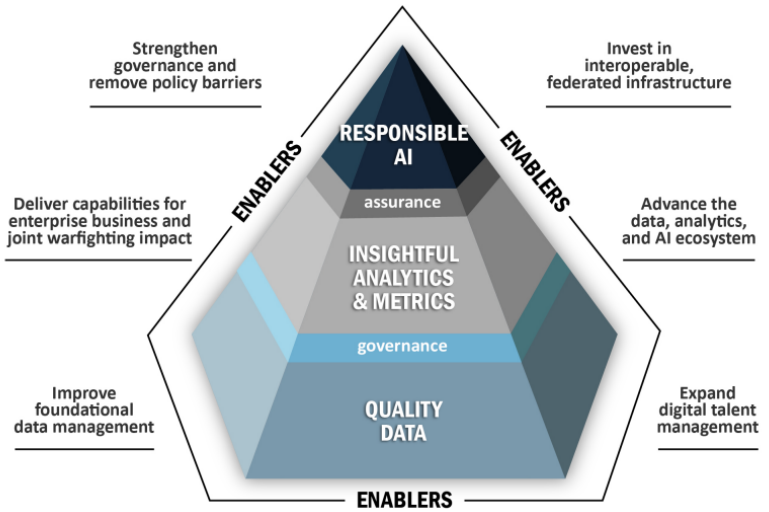
The National Security Commission on Artificial Intelligence (NSCAI) was established in 2018 to assess the impact of AI on US national security. The final report was published in 2021 and presents a detailed analysis of the strategy for using AI. At the very beginning of the report, the following is stated: "America is not prepared to defend or compete in the AI era" (NSCAI 2021, 1), which clearly indicates the need to create security strategies that will follow global trends in AI development. Technological advantage in the third decade of the 21st century is very easy to lose, especially if we take into account the rapid development of AI in China. In order for a country to prepare for the challenges of the AI era, it is important to recognize AI threats to national security, which is specifically stated in this report. Disinformation and propaganda, cyber attacks, misuse of AI in the field of biotechnology and the like are key areas that are recognized as

security threats. The report states that "the next major conflict will begin not with a military strike, but with disinformation and sabotage using AI tools" (NSCAI 2021, 24). As for cyber attacks, AI developments are becoming so accelerated that the human factor cannot keep up with the scale of security threats (NSCAI 2021, 33). With this in mind, the recommendation is to develop technical resilience of defense systems and algorithms for recognizing deepfake content. Also, "the U.S. must invest in automated cyber defense systems that incorporate AI in real-time detection and mitigation" (NSCAI 2021, 42). To achieve this, the need for developing a workforce that will be trained to use complex AI-based systems is emphasized. Also, institutional modernization, namely the modernization of command systems (Joint All-Domain Command and Control – JADC2) is necessary (Department of Defense [DoD] 2022).

A significant part of the final report is dedicated to China as a strategic challenge, precisely because it is rapidly developing cutting-edge AI-based systems. China is investing heavily in the Military-Civil Fusion (MCF) concept, which is key to rapid technological advancement in the field of military AI and maintaining national security (Dupont-Sinhsattanak 2025). "China is organized, resourced, and determined to win" the AI competition, the report states (NSCAI 2021, 6). One of the strategies for dealing with China's technological rise is the development of global cooperation in the field of AI. In addition, it is proposed to increase the budget from 2 billion in 2022 to 32 billion in 2026, in order to achieve superiority in research (NSCAI 2021, 193). "AI will not remain in the hands of allies. It will empower adversaries and shape geopolitics. We must act" (NSCAI 2021, 12).

The US Department of Defense adopted a major "Data, Analytics and Artificial Intelligence Adoption Strategy" in November 2023. The goal is to rapidly adopt advanced AI-based technologies to support and provide superiority to American warfighters on the battlefield. The strategy builds on and replaces the 2018 DOD AI Strategy (DoD 2019) and the revised 2020 DOD Data Strategy (DoD 2020), tracking AI trends and creating a foundation for future progress in the field (Clark 2023). The strategy prescribes an agile and efficient approach to the development and application of artificial intelligence, and its essence with a distinct hierarchy of processes and goals is represented by a pyramid (Figure 2).

Figure 2. Strategic Goals and the AI Hierarchy of Needs



Source: DoD 2023, 7.

In order to achieve the following decision-making advantages: "1. Superior battlefield awareness and understanding; 2. Adaptive force planning and application; 3. Fast, precise and resilient kill chains; 4. Resilient sustainment support; and 5. Efficient enterprise business operations" (DoD 2023; Clark 2023), the implementation of the presented (Figure 2) is necessary. The basis is high-quality and verified data, so that analytics can be further implemented on the basis of which the Ministry of Defense can make decisions. At the top of the pyramid is responsible AI, with the ultimate goal of improving mission results. Around the pyramid are segments that should enable the functioning of a complex decision-making system (DoD 2023).

On October 24, 2024, former US President Joe Biden signed the Memorandum "Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence (NSM-25)", which is very significant for the role of AI in the context of national security, especially in the development of military capabilities, intelligence activities, and cybersecurity (White House 2024). The NSM identifies three goals: 1. encouraging the safe and secure development of artificial intelligence technologies, which means that systems must be resistant to hostile cyberattacks, 2.

advancing national security interests through the strategic deployment of artificial intelligence, and 3. promoting a global framework for the governance of artificial intelligence based on transparency, human rights, and trustworthiness (Bieber, Christensen and Hopkins 2024). The Memorandum emphasizes AI as a key component of modern military systems and national defense, with the ultimate goal of preventing hostile parties from harming US military and other capabilities. It envisages the accelerated development of AI, but also the introduction of stricter control and protection measures. NSM-25 is also a response to the global "arms race", which is particularly evident between the US and China in the domain of AI (White House 2024). In this context, Jake Sullivan said at the National Defense University in Washington: "We have to be faster in deploying AI... If we don't deploy AI more quickly... we risk squandering our hard-earned lead" (Reuters 2024). It is also important to mention the Chief Digital and Artificial Intelligence Office (CDAO) within the US Department of Defense, which is actively working on the development and incorporation of AI in the security sector. CDAO played a key role in the development of The Combined Joint All Domain Command and Control (CJADC2), with the goal of enabling a command-integration framework for connecting different systems and actors in real time using AI technology (DoD 2022). Thus, CJADC2 is a complex system that provides information to decision makers at the tactical, operational, and strategic levels. The US Department of Defense continues to develop CJADC2 in accordance with the strategic environment and concepts of warfare (CDAO 2025; Government Accountability Office 2025; DoD 2022). In addition, CDAO is developing multifunctional analytical tools (Advana, GAMECHANGER, JATIC) for collecting, analyzing, and testing data and AI models. Advana connects more than 400 systems, including tools, services, and analytics to enable data-driven decision-making in the Department of Defense (CDAO 2025). The Joint AI Test Infrastructure Capability (JATIC) is an analytical tool used to test AI models for Department of Defense programs, research laboratories, and the like, to ensure the operation of systems in the domain of national security (Martin 2025; CDAO 2025). In order to connect different systems and databases, CDAO has also offered Open Data and Applications Government-owned Interoperable Repositories (DAGIR).

We would also add to the list of innovations the AI Rapid Capabilities Cell project, launched in December 2024, which represents

the Department of Defense's effort to accelerate the application of AI in military operations. The domain of warfare includes the following: "Command and Control (C2) and decision support, operational planning, logistics, weapons development and testing, uncrewed and autonomous systems, intelligence activities, information operations, and cyber operations" (DoD 2024; CDAO 2024, 2025). When talking about the use of AI in warfare, it is necessary to mention Project Maven, which the US DoD launched in 2017 with the aim of accelerating the adoption of AI and machine learning in the military. The focus was primarily on processing field data and identifying potential targets and objects on the battlefield (Pellerin 2017).

Finally, we will note the announcement by US Secretary of Defense Pete Hegseth on the development of the National Defense Strategy (NDS), the deadline for the final draft is August 31, 2025. The NDS serves as a strategic map for the work of the Department of Defense, especially in the context of President Donald Trump's America First and Peace Through Strength agenda. The new Strategy prioritizes defending the United States, strengthening alliances, and deterring China in the Indo-Pacific (DoD 2025). The first Trump administration and the Biden administration have characterized China as the greatest threat to the United States (Horton and Natanson 2025), which will also be reflected in future rivalries.

AI in China's Security Strategies

China's rise in the last few decades has been largely driven by the rapid development of new technologies and AI, as clearly stated in numerous strategies and the direct involvement of the government. President Xi Jinping spoke at a session of the Communist Party of China in October 2016 about the development of new AI technologies and the challenges to state and national security. On that occasion, Jinping stated: "China must work toward its goal of becoming a cyber power by accelerating reinforcement of security and defense capabilities in cyber space, accelerating the promotion of social governance using IT, and accelerating the advancement of China's right to speak internationally and right to set rules governing cyber space" (Sen 2019). A month later, in November 2016, the Chinese government adopted the Cyber Security Law (Creemers, Webster and Triolo 2018), and in December of the same year adopted the National Cyber Security Strategy (Creemers 2016). In

March 2017, they adopted the International Strategy of Cooperation on Cyber Space (Shaohui 2017), thus demonstrating their proactivity in developing a cybersecurity doctrine with the aim of responding to global challenges and threats.

The cybersecurity strategy has three main drivers: economic, political and military. Therefore, the goals are clear: 1. maintaining economic growth and stability, "involves industrial economic cyber espionage of the US and other foreign targets" (Sen 2019, 129); 2. protecting the ruling power of the CCP, through information control and propaganda; 3. preparing for military scenarios and ensuring military superiority in the event of cyber conflict with the enemy through military modernization, research into computer network operations, and human capital development; 4. studying and understanding military infrastructure and preventing adversary actions in cyberspace (Sen 2019, 129).

On July 20, 2017, the Chinese State Council issued a seminal document, "A [New] Generation Artificial Intelligence Development Plan". This plan envisions a key role for AI in improving national security and global competitiveness, with the ultimate goal of China becoming a major global center for AI innovation by 2030 (State Council 2017; Webster *et al.* 2017). Thus, since 2017, China has considered AI a national strategic sector (Wu *et al.* 2020). Aware that AI is affecting and changing all segments of the state and society, China is actively using strategic opportunities for technology development to gain a global advantage in science and technology. The Plan emphasizes that AI is a "strategic technology that will lead the way in the future" (State Council 2017; Webster *et al.* 2017). Accordingly, China has set three strategic goals with time-bound targets in this Plan. The first is by 2020, when the overall technology and application of AI should be in line with the global level of progress. In addition, it is envisaged to develop a competitive AI industry in the global market. The second strategic goal for the period until 2025 predicts that China will establish a new generation of AI theories and technological systems, with the ability of autonomous learning. The volume of the core artificial intelligence industry will exceed 400 billion yuan. The third strategic goal by 2030 predicts that China's theories, technologies and applications of artificial intelligence will reach world-leading levels, making China the world's leading center for innovation in the field of artificial intelligence. The total economic effects of AI by 2030 are estimated at 10 trillion yuan

(about 1.38 trillion US dollars) (State Council 2017; Webster *et al.* 2017). The importance of the development of AI technologies in China is also evidenced by the growth in the number of companies engaged in this field. According to data from the Chinese Yicai Media Group in 2023, 1.09 million companies engaged in AI development were registered in China (Zhang and Khanal 2024, 22). In the period from 2020, there has been a continuous growth in the number of registered companies in this sector.

In addition to the above, the Plan also has a geopolitical dimension, as it perceives AI as a strategic resource in global competition. Thus, offensive and defensive AI capabilities are envisaged, especially in a military context (Allen and Chan 2017). More importantly, the Plan highlights the strategy of developing military-civilian integration, as a comprehensive approach to AI technology (State Council 2017; Webster *et al.* 2017). The growth in the number of newly registered companies in the field of AI demonstrates China's approach to comprehensive development in the civilian and military sectors, as well as their integration. In order to successfully implement the goals set in the aforementioned Plan, "China has undertaken a series of policy initiatives" since 2017, including "the integration of AI into national five-year science and technology plans, the Three-Year Action Plan to Promote AI (2018–2021), the AI Innovation Action Plan for Institutions of Higher Education" (Zhang and Khanal 2024), the establishment of state funds for the development of AI in various sectors (Bloomberg 2019; Zhu *et al.* 2023), and the like.

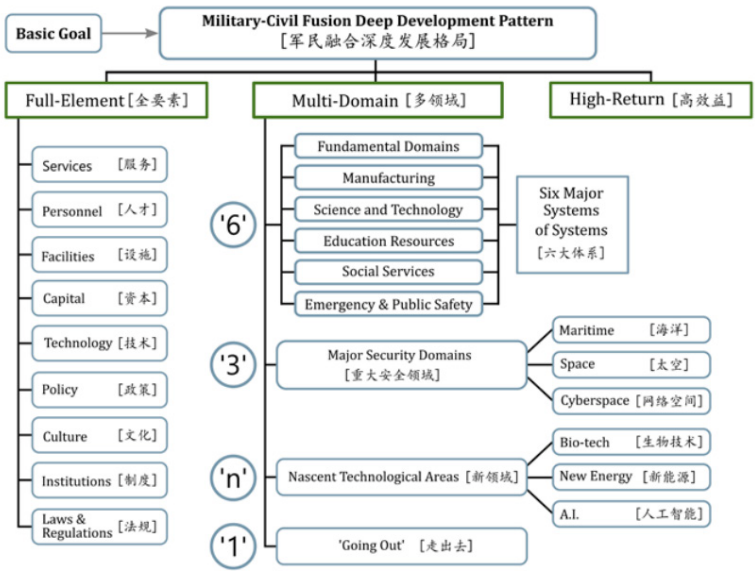
The Military Civil Fusion Strategy (MCF) has been a Chinese development concept for more than three decades, but since Xi Jinping came to power, this strategy has been increasingly mentioned and analyzed by American institutions and seen as a threat to global leadership. The 2024 US Department of Defense Annual Report addresses "military and security developments related to the People's Republic of China, providing an overview of the PRC's national, economic, and military strategy and providing insight into the strategy, current capabilities, and activities of the People's Liberation Army (PLA), as well as its future modernization goals" (DoD 2024, 5). Many authors consider the MCF to have a multi-decade historical evolution, from the period of Mao Zedong, then Deng Xiaoping, Jiang Zemin, and Hu Jintao, to Xi Jinping (Bhutani 2023; Bitzinger 2021; Stone and Wood 2020). In 2014, President Jinping elevated MCF to the level of a "national

strategy", considering it crucial for rejuvenating the Chinese nation and strengthening the military, and in 2015, he changed the development approach from "early-state fusion" to "deep fusion" (Bhutani 2023). Deep integration of the military and civilian sectors encompasses, on the one hand, all aspects of national defense and military-technological development, and on the other hand, strengthening the economic and social system, such as science, education, industry, and other aspects (Stone and Wood 2020).

President Jinping has recognized the great importance of the MCF, which is why he has made it the basis for other strategies and plans such as "Made in China 2025" "(MIC 2025), the 2017 New Generation Artificial Intelligence Development Plan (AIDP), and the People's Liberation Army (PLA) modernization plans to become a world-class military by 2049" (Bitzinger 2021; Wübbeke *et al.* 2016). He emphasized the importance of the MCF at the Central Commission for Military-Civil Fusion Development (CCMCFD) on June 20, 2017, when he said: "We must accelerate the formation of a full-element, multi-domain, and high-return military-civil fusion deep development pattern, and gradually build up China's unified military-civil system of strategies and strategic capability" (Bhutani 2023, 72). The above statement was analyzed by a group of researchers from the National Defense University of China and concluded that it presents short-term goals in its first part and long-term goals in its second part, which fits into the vision of global military dominance in 2049 (Bhutani 2023; Jinping 2023).

Some authors believe that the development of the MCF can have four positive outcomes. First, the development of the MCF can contribute to China's transformation into a powerful nation. Second, the MCF can help in gaining international advantage in the field of technology and the military. Third, the MCF is a unique opportunity to improve the governance system in China. Fourth, the "MCF supports the construction of a world-class military" (Stone and Wood 2020, 26–27). Also, the MCF is a key military-technological innovation strategy, especially in the context of AI and intelligentized warfare. "Intelligentized warfare is defined as the 'operationalization' of artificial intelligence (AI) and its enabling technologies, such as cloud computing, big data analytics, quantum computing, and autonomous systems, for military applications" (Bitzinger 2021, 7; Bitzinger, Evron and Yang 2021).

Figure 3. Military-civil fusion deep development pattern



Source: Stone and Wood 2020, 28.

The essence of MCF can be seen in Figure 4, as well as the development map that the Chinese authorities consistently adhere to. We have purposefully presented the patterns of deep development of military-civilian fusion, because they systematically show all those segments that are necessary to achieve the ultimate goal – global leadership. The three basic characteristics of MFC are: 1. "full-element; 2. multi-domain; 3. high-return. The first defines the types of resources shared by the military and civilian sectors" (Stone and Wood 2020, 28). Then, the second item highlights the priority fields for the development of MFC. Finally, the third item should show the desired effects from the development of MFC (29–38).

Finally, we will also list the most current document "White paper on national security", which was adopted in China in May 2025. The document discusses national security in the new era, which, in addition to military, includes political, economic, technological, cultural and cyber dimensions of security. It emphasizes a holistic approach to security as the first principle and strategic approach of China at the global level. "It is one that takes the people's security as its ultimate goal, political security as the fundamental task, and national interests as the guiding

principle" (Xinhua 2025). The document clearly shows the approach of the Chinese president, who often speaks about national and global security and the common future of humanity, which should ultimately contribute to the strategy of China's rejuvenation by 2049 (Glišin 2024).

CONCLUSION

Consideration of the security strategies of the United States of America and the People's Republic of China in the context of the application of artificial intelligence (AI) indicates deep-rooted differences in the institutional models, political priorities, normative approaches and global ambitions of these states. Although both states recognize AI as a central instrument in contemporary security concepts, differences in the structural and ideological framework of their strategies are crucial for understanding global technological competition in the 21st century.

The United States, in its strategic documents, openly emphasizes the need to maintain global superiority in the field of artificial intelligence as a means of improving national security, but also as an instrument of power projection. Thus, in several documents, such as the National Security Strategy (White House 2022), Executive Order 13859 (Federal Register 2019), NSCAI Final Report (NSCAI 2021), as well as in the most recent NSM-25 (White House 2024), the need for technological dominance, innovation, and the defense of democratic values in the digital age is consistently affirmed. The United States views AI not only as a defense and operational tool, but also as a comprehensive platform for the development of autonomous systems, cyber defense, predictive analytics, and for the construction of interoperable and resilient command systems (e.g. CJADC2 and AI Rapid Capabilities Cell). In its approach to AI development, the US also emphasizes partnerships with allied countries, the involvement of the private sector and the scientific community, with an emphasis on investing in talent.

On the other hand, China sees AI as a key element of its long-term national revival and global transformation of the international order. In China, AI is directly linked to the concept of national security, economic modernization, and geopolitical autonomy. According to documents such as the New Generation Artificial Intelligence Development Plan (Webster *et al.* 2017), the National Cyber Security Strategy (Creemers 2016), and the Military-Civil Fusion (MCF) strategy, China is integrating AI into all levels of state activity - from educational and industrial

capacities, to the army and control of the information space. Unlike the United States, the Chinese approach is centralized, directed by the Communist Party, and strongly relies on long-term planning and state investment in new technologies. China's strategy insists on technological sovereignty, internalization of production, and long-term integration of AI into the military, economic, and ideological spheres. The concept of intelligentized warfare, as well as the deep integration of military and civilian technology through MCF, represent an attempt to transform the very nature of warfare and conflict in the digital era. China's focus on strategic planning until 2049, when the PLA is planned to become a "world-class army", shows that AI is part of a much broader vision of national prestige and power.

Cybersecurity is a key area of application of artificial intelligence for both countries, especially when it comes to disinformation, protection of digital infrastructure, automation of systems for detecting and responding to deepfake content, digital sovereignty and other activities in cyberspace. Also, both countries see artificial intelligence as a tool for shaping the international system, and thus artificial intelligence, in addition to security competition, is also becoming a tool for geopolitics. In this regard, the US is making efforts to maintain the Extended Deterrence Strategy and to develop new military technologies to neutralize Chinese doctrines.

Finally, we can conclude that artificial intelligence represents a new axis of global security competition. Although both countries strive to develop technological superiority, the US and China differ in their concepts of security, degree of centralization, regulation, and attitude towards global governance. Understanding these differences is crucial for assessing future security scenarios, but also for creating policies that will ensure stability in an increasingly complex digital world. Regardless of their different approaches, both the US and China will continue to shape the global agenda regarding AI, which will affect all actors in the international system.

REFERENCES

Allen, Greg, and Taniel Chan. 2017. "Artificial intelligence and national security." *Belfer Center for Science and International Affairs*. July 1, 2017. https://www.belfercenter.org/publication/artificial-intelligence-and-national-security

Balbaa, Muhammad, and Marina Abdurashidova. 2024. "The impact of artificial intelligence in decision making: A comprehensive review." *EPRA international Journal of Economics* 11 (2): 27–38.

Bekker, Alex. 2019. "5 Types of AI to Propel Your Business." *Science Soft*. May 13, 2019. https://www.scnsoft.com/blog/artificial-intelligence-types

Bhutani, Rajeev. 2023. "China's Military-Civil Fusion and Lessons for India." *Synergy* 2 (1): 68–87.

Bieber, Corey, Guillermo Christensen, and Brian Hopkins. 2024. "AI National Security Memorandum Scopes Key Priorities." *The National Law Review*. December 3, 2024. https://natlawreview.com/article/ai-national-security-memorandum-scopes-key-priorities

Bitzinger, Richard. 2021. "China's Shift from Civil-Military Integration to Military-Civil Fusion." *Asia Policy* 16 (1): 5–24.

Bitzinger, Richard, Yoram Evron, and Zi Yang. 2021. "China's Military-Civil Fusion Strategy: Development, Procurement, and Secrecy." *Asia Policy* 16 (1): 1–64.

Bloomberg. 2019. "China to Funnel $29 Billion Towards its Chip Ambitions." *Bloomberg*. October 29, 2019. https://www.bloomberg.com/news/articles/2019-10-29/china-to-funnel-29-billion-towards-its-chip-ambitions

Chappell, Jim. 2020. *Artificial Intelligence: from Predictive to Prescriptive and Beyond*. Cambridge: Aveva.

Chief Digital and Artificial Intelligence Office [CDAO]. 2024. *Annual Report on AI Capabilities in Defense*. Washington, DC: U.S. Department of Defense.

Chief Digital and Artificial Intelligence Office [CDAO]. 2025. *CJADC2 Program Overview and Advana Architecture*. Washington, DC: U.S. Department of Defense.

Clark, Joseph. 2023. "DOD Releases AI Adoption Strategy." *U. S. Department of Defense*. November 2, 2023. https://www.defense.gov/News/News-Stories/Article/Article/3578219/dod-releases-ai-adoption-strategy/

Cohen, Charles. 2023. "AI in Defense: Navigating Concerns, Seizing Opportunities." *National Defense*. July 25, 2023. https://www.nationaldefensemagazine.org/articles/2023/7/25/defense-department-needs-a-data-centric-digital-security-organization

Creemers, Rogier. 2016. *National Cyberspace Security Strategy*. China Copyright and Media.

Creemers, Rogier, Graham Webster, and Paul Triolo. 2018. *Translation: Cybersecurity Law of the People's Republic of China*. Stanford: Stanford University.

Department of Defense (DoD). 2019. *DoD Artificial Intelligence Strategy: Harnessing AI to Advance Our Security and Prosperity*. Washington: Department of Defense U.S.

Department of Defense (DoD). 2020. *DoD Data Strategy*. Washington: Department of Defense U.S.

Department of Defense (DoD). 2022. *Summary of the Joint All-Domain Command and Control Strategy*. Washington: Department of Defense U.S.

Department of Defense (DoD). 2023. *Data, Analytics, and Artificial Intelligence Adoption Strategy*. Washington: Department of Defense U.S.

Department of Defense (DoD). 2024. *Annual Report on Military and Security Developments Involving the PRC*. Washington: Department of Defense U.S.

Department of Defense (DoD). 2025. *Statement on the Development of the 2025 National Defense Strategy*. Washington: Department of Defense U.S.

Despotović, Ljubiša, and Vanja Glišin. 2024. *Geopolitika i kriptopolitika*. Beograd: Catena Mundi.

Dupont-Sinhsattanak, Alexandre. 2025. "Modernizing a giant: assessing the impact of military-civil fusion on innovation in China's defence-technological industry." *Defence and Peace Economics*: 1–27. DOI: 10.1080/10242694.2025.2460458

Đorić, Marija, and Vanja Glišin. 2023. „Upotreba veštačke inteligencije u rusko-ukrajinskom sukobu." *Politika nacionalne bezbednosti* 25 (2): 59–76. DOI: 10.5937/pnb25-47369

Evron, Yoram, and Richard Bitzinger. 2023. *The Fourth Industrial Revolution and Military-Civil Fusion: A New Paradigm for Military Innovation?* Cambridge: Cambridge University Press.

Federal Register. 2019. "Executive Order 13859: Maintaining American Leadership in Artificial Intelligence (84 FR 3967)." *The Daily Journal of the United States Government*. February 11, 2019. https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence

Gaba, Shivani, Ishan Budhiraja, Vimal Kumar, Shesikala Martha, Jebreel Khurmi, and Akansha Singh [Gaba *et al.*]. 2024. "A Systematic Analysis of Enhancing Cyber Security Using Deep Learning for Cyber Physical Systems." *IEEE Access* (12): 6017–6035. DOI: 10.1109/ACCESS.2023.3349022

Galán, Jose Javier, Ramon Alberto Carrasco, and Antonio LaTorre. 2022. "Military Applications of Machine Learning: A Bibliometric Perspective." *Mathematics* 10 (9): 1397. DOI: 10.3390/math10091397

Glišin, Vanja. 2024. "A Leader with the Vision of the Shared Future of Humanity." *Progress* 5 (1): 153–157. DOI: 10.5937/napredak5-50218

Government Accountability Office. 2025. *CJADC2 Progress and Challenges*. Washington: GAO.

High-Level Expert Group on Artificial Intelligence. 2019. *A definition of AI: Main capabilities and scientific disciplines*. Brussels: European Commission.

Horton, Alex, and Hannah Natanson. 2025. "Secret Pentagon memo on China, homeland has Heritage fingerprints." *The Washington Post*. March 29, 2025. https://www.washingtonpost.com/national-security/2025/03/29/secret-pentagon-memo-hegseth-heritage-foundation-china/

Jinping, Xi. 2023. *On building a human community with a shared future*. Beograd: Albatros Plus.

Kania, Elsa, and John Costello. 2021. "Seizing the commanding heights: the PLA Strategic Support Force in Chinese military power." *Journal of Strategic Studies* 44 (2): 218–264. DOI: 10.1080/01402390.2020.1747444

Kaplan, Andreas, and Michal Haenlein. 2018. "Siri, Siri, in My Hand: Who's the Fairest in the Land? On the Interpretations, Illustrations, and Implications of Artificial Intelligence." *Business Horizons* 62 (1): 15–25. DOI: 10.1016/j.bushor.2018.08.004

Koncová, Dominika, and Iveta Kremeňová. 2022. "Interoperability and Its Significance." *Transport and Communications* 10 (1): 6–9.

Luknar, Ivana. 2024. *Veštačka inteligencija: Izazovi i mogućnosti*. Beograd: Medija centar Odbrana.

Márquez-Díaz, Jairo Eduardo. 2024. "Benefits and Challenges of Military Artificial Intelligence in the Field of Defense." *Computación y Sistemas* 28 (2): 309–323. DOI: 10.13053/CyS-28-2-4684

Martin, Isaac. 2025. "Breaking Down the DOD's Chief Digital & AI Office Budget for FY25." *ExecutiveGov*. February 14, 2025. https://executivegov.com/article/breaking-down-dod-cdao-fy25-budget

McCarthy, John. 2007. "What Is Artificial Intelligence?" *Stanford University*. November 12, 2007. http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html

McKendrick, Kathleen. 2019. *Artificial Intelligence Prediction and Counterterrorism*. London: Chatham House.

National Security Commission on Artificial Intelligence [NSCAI]. 2021. *Final Report*. Washington: NSCAI.

Parker, Lynne. 2020. "The American AI Initiative: The U.S. strategy for leadership in artificial intelligence." *OECD.ai*. June 11, 2020. https://oecd.

ai/en/wonk/the-american-ai-initiative-the-u-s-strategy-for-leadership-in-rtificial-intelligence

Pellerin, Cheryl. 2017. "Project Maven to Deploy Computer Algorithms to War Zone by Year's End." *U.S. Department of Defense News*. July 21, 2017. https://www.defense.gov/News/News-Stories/Article/Article/1254719/project-maven-to-deploy-computer-algorithms-to-war-zone-by-years-end/

Radun, Viktor. 2025. *Na putu ka transčoveku*. Beograd: Pešić i sinovi.

Rashid, Adib Bin, Karim Kausik, Alhamed Al Hassan Sunny, and Mehedy Hassan Bappy [Rashid *et al.*]. 2023. "Artificial Intelligence in the Military: An Overview of the Capabilities, Applications, and Challenges." *International Journal of Intelligent Systems* (4): 1–31. DOI: 10.1155/2023/8676366

Reuters. 2024. "White House presses gov't AI use with eye on security, guardrails." *Reuters*. October 24, 2024. https://www.reuters.com/world/us/white-house-presses-govt-ai-use-with-eye-security-guardrails-2024-10-24/

Ro, Christine. 2023. "On the warpath: AI's role in the defence industry." *BBC*. August 24, 2023. https://www.bbc.com/news/business-66459920

Russell, Stuart, and Peter Norvig. 2010. *Artificial Intelligence: A Modern Approach*. New Jarsey: Prentice-Hall.

Scherer, Matthew. 2016. "Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies." *Harvard Journal of Law & Technology* 29 (2): 354–398. DOI: 10.2139/ssrn.2609777

Sen, Gautam. 2019. "China's Cyber Security Strategy: Global Implications." *Scholar Warrior*, *Spring* 2019: 128–134.

Shalev-Shwartz, Shai, and Shai Ben-David. 2014. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge: Cambridge University Press.

Shaohui, Tian. 2017. "International Strategy of Cooperation on Cyberspace." *Xinhua*. March 1, 2017. http://www.xinhuanet.com//english/china/2017-03/01/c_136094371.htm

Sheikh, Haroon, Prins Corien P, and Erik Schrijvers. 2023. *Mission AI: The New System Technology*. London: Springer.

Smola, Alex, and S. V. N. Vishwanathan. 2008. *Introduction to Machine Learning*. Cambridge: Cambridge University Press.

Spherical Insights. 2025. "Top 10 Artificial Intelligence Spending Countries in 2025: Statistics and Facts Analysis 2024 to 2035." *Spherical Insights*. May, 2025. https://www.sphericalinsights.com/blogs/top-10-artificial-intelligence-spending-countries-in-2025-statistics-and-facts-analysis-2024-to-2035

State Council. 2017. *A [New] Generation Artificial Intelligence Development Plan*. Beijing: State Council.

Stone, Alex, and Peter Wood. 2020. *China's Military-Civil Fusion Strategy*. Alabama: China Aerospace Studies Institute.

Takpah, Nathalie Essi Afefa, and Victor Oriakhi. 2025. "Cybersecurity Challenges and Technological Integration in Military Supply Chain 4.0." *Journal of Information Security* 16 (1): 131–148. DOI: 10.4236/jis.2025.161007

Wang, Weiyu, and Keng Siau. 2019. "Automation, Robotics, Future of Work and Future of Humanity: A Review and Research Agenda." *Journal of Database Management* 30 (1): 61–79. DOI: 10.4018/JDM.2019010104

Webster, Graham, Rogier Creemers, Elsa Kania, and Paul Triolo [Webster *et al.*]. 2017. *Translation: China's New Generation Artificial Intelligence Development Plan*. Stanford: Stanford University.

Wegner, Peter. 1996. "Interoperability." *ACM Computing Surveys* 28 (1): 285–287.

White House. 2022. *National Security Strategy of the United States*. Washington: White House.

White House. 2024. *National Security Memorandum on Artificial Intelligence (NSM-25)*. Washington: White House.

Wu, Fei, Cewu Lu, Ming-Li Zhu, Hao Chen, Jun Zhu, Kai Yu, Lei Li, Ming Li, Qian Chen, Xi Li, Xudong Cao, Zhongyuan Wang, Zhengjun Zha, Yueting Zhuang, Yunhe Pan [Wu *et al*.] 2020. "Towards a new generation of artificial intelligence in China." *Nature Machine Intelligence* 6 (2): 312–316.

Wübbeke, Jost, Mirjam Meissner, Max Zenglein, Jaqueline Ives, and Bjorn Conrad [Wübbeke *et al.*]. 2016. *Made in China 2025: The making of a high-tech superpower and consequences for industrial countries*. Berlin: MERICS.

Xinhua. 2025. *China releases white paper on national security*. Beijing: The State Councile of China.

Zhang, Hongzhou, and Shaleen Khanal. 2024. "To Win the Great AI Race, China Turns to Southeast Asia." *Asia Policy* 19 (1): 21–34.

Zhu, Julie, Kevin Huang, Yelin Mo, and Roxanne Liu [Zhu *et al.*]. 2023. "Exclusive: China to launch $40 billion state fund to boost chip industry." *Reuters*. September 5, 2023. https://www.reuters.com/technology/china-launch-new-40-bln-state-fund-boost-chip-industry-sources-say-2023-09-05/

Вања Глишин*

Институт за политичке студије, Београд

ВЕШТАЧКА ИНТЕЛИГЕНЦИЈА У БЕЗБЕДНОСНИМ СТРАТЕГИЈАМА САД И КИНЕ НА ПОЧЕТКУ 21. ВЕКА – БОРБА ЗА ГЛОБАЛНУ ПРЕВЛАСТ**

Резиме

Развој вештачке интелигенције (ВИ) и све учесталија примена у различитим сегментима државе и друштва доводи до тога да је неопходно праћење, изучавање и активно усвајање знања и закључака о томе, на нивоу свих научних области. Последњих неколико деценија приметна је свеобухватна примена ВИ у безбедносним стратегијама великих сила попут Сједињених Америчких Држава и Кине, те због тога сведочимо својеврсној трци у развоју нових технологија и њиховој примени. Фокус је на државама које најактивније развијају вештачку интелигенцију у безбедносном сектору, што је јасно истакнуто у безбедносним стратегијама које ће бити предмет анализе, али имају и различите концепције безбедности, степен централизованости, регулативе и однос према глобалном управљању, што оправдава упоредну анализу и доприноси научној дебати. У првом делу рада дефинисаћемо ВИ, навешћемо различите поделе и фокусираћемо се на употребу ВИ у безбедносном сектору. Затим ћемо анализирати безбедносне стратегије САД и Кине, како бисмо лоцирали улогу ВИ и планове за будући развој и борбу за глобалну превласт. На основу тога у последњем делу рада предочићемо сличности и разлике између безбедносних стратегија САД и Кине. У раду примењујемо методу анализе (дескриптивна улога) са циљем да се опише проблем истраживања и објасне (експланаторна улога) кључни сегменти који се тичу вештачке

* Имејл: vanja.glisin@ips.ac.rs; ORCID: 0000-0002-7747-6205

** Овај рад је настао у оквиру научноистраживачке делатности Института за политичке студије, коју финансира Министарство науке, технолошког развоја и иновација Републике Србије.

интелигенције у безбедносним стратегијама. На основу описаних и објашњених појава и процеса, настојаћемо да методом синтезе разумемо сложену целину, пре свега потенцијале вештачке интелигенције у оквиру безбедносних стратегија. Квалитативном анализом релевантне домаће и стране литературе, доступних докумената и извештаја прикупићемо податке у вези са предметом истраживања.

Кључне речи: вештачка интелигенција, САД, Кина, безбедност, стратегија, Безбедносна стратегија

* This manuscript was submitted on June 30, 2025, and accepted by the Editorial Board for publishing on November 10, 2025.

UDC 327:351.74:004.8(4-672EU)
DOI: 10.5937/spm94-59854
Review article

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 93-118

Marija Tasić*

University of Criminal Investigation and Police Studies, Serbia

Aleksandar Ivanović**

Faculty of Law, University of Montenegro, Montenegro

Tanja Kesić****

University of Criminal Investigation and Police Studies, Serbia

Ivana Bjelovuk*****

University of Criminal Investigation and Police Studies, Serbia

MODERNIZATION OF THE LAW ENFORCEMENT SYSTEM: THE USE OF CONTEMPORARY ARTIFICIAL INTELLIGENCE TOOLS IN CRIME PREVENTION AND SUPPRESSION

Abstract

Law enforcement agencies represent one of the most significant users of artificial intelligence–based systems in the context of contemporary manifestations of crime. Police agencies, as well as those operating within the criminal justice system, already make extensive use of various systems, software, and tools in the field of crime prevention and suppression. This article analyzes the application of existing artificial intelligence systems by law enforcement agencies through

* E-mail: marija.tasic@kpu.edu.rs; ORCID: 0000-0002-3755-4854
** E-mail: ialeksandar@t-com.me; ORCID: 0000-0002-8186-0883
*** E-mail: tanja.kesic@kpu.edu.rs; ORCID: 0000-0001-9473-4379
**** E-mail: ivana.bjelovuk@kpu.edu.rs; ORCID: 0000-0002-9370-8758

a critical review of relevant academic and professional literature, along with an illustration of examples of systems, software, and tools used in the operational context. In the first chapter, the possibilities of applying artificial intelligence systems in the field of crime prevention are presented through an overview of specific software for predicting crime and the criminal behavior of individuals, including risk assessments of recidivism and repeat victimization. In the second chapter, the possibilities of applying specific systems, software, and tools in the field of crime suppression are presented, starting from the receipt of criminal reports, the preparation and planning of criminal investigations, and up to their application in the area of international police cooperation. Finally, the article points out the controversies surrounding the use of artificial intelligence by law enforcement agencies, considering several aspects such as technological, legal, ethical, and social.

Keywords: law enforcement agencies, crime prevention, crime suppression, artificial intelligence, controversies

INTRODUCTION

In contemporary conditions, law enforcement agencies strive to enhance their capacities to operate efficiently and promptly in a dynamic and complex environment, in order to safeguard the community and protect the human rights of all individuals. The manifestations of crime are changing, both in terms of their forms of occurence and the dynamics of their manifestation and development. The traditional approach to crime, which involves responding to criminal activity after it has already occurred, is losing its effectiveness in combating contemporary forms of criminal activity. Proactive policing, i.e., preventing crime before it occurs, is gaining priority status, emphasizing the importance of systematic and organized data collection from various sources, their processing, and utilization in the decision-making process of the competent authorities.

In an effort to overcome human limitations in solving complex problems across various areas of life, artificial intelligence systems have been developed, representing a combination of human beings and machines. Essentially, these are machines programmed to think and act

with a certain degree of human intelligence (Nehra 2015, 500). Artificial intelligence (hereinafter: AI) can be defined as a system, program, software, or algorithm that simulates human thinking and behavior, i.e., operates autonomously in order to think and act rationally and humanely, make decisions, or produce outputs (Martinez 2018, 1038). These are systems that display intelligent behaviour by analysing their environment and taking actions to achieve specific goals (EC 2018a). AI systems are designed to operate with varying levels of autonomy, along with the possibility to adapt to their environment; based on the inputs or data, they determine how to generate outputs, such as predictions, content, recommendations, or decisions (Regulation (EU) 2024/1689, Art. 12).

As such, AI systems have found their application in various areas, such as industry, medicine, finance, transportation, science, and others. Their potential has also been recognized by law enforcement systems, and the advancements achieved through the progressive development of AI systems are now being applied in the area of crime prevention and suppression. In addition, these systems are increasingly being used as support in decision-making processes within the criminal justice system.

For the purpose of examining existing artificial intelligence systems and their application within the law enforcement system, this article employs several scientific methods, such as content analysis, case study, and the comparative method. The focus of analysis is on the content available within relevant academic and professional literature regarding AI systems and their application in the area of crime prevention and suppression. This article provides an overview of AI-based systems and software utilized by law enforcement agencies worldwide for the purposes of predicting crime, individual criminal behavior, recidivism, and repeat victimization. Additionally, it explores the application of such systems and software in supporting administrative and operational activities in the context of criminal investigations. Finally, this article highlights the advantages of AI systems, manifested in the increased operational efficiency within law enforcement agencies and their capability in handling large volumes of data, as well as the challenges and criticisms regarding their application.

APPLICATION OF AI SYSTEMS IN CRIME PREVENTION

In practice, numerous police agencies utilize AI systems for the collection and processing of large volumes of data that may be relevant to crime prevention. One of the fundamental principles of police work is the principle of operational efficiency, which requires knowledge of the patterns of criminal activity in a specific area, i.e., the identification of so-called "crime hotspots", and the use of various sources of information on matters important for crime prevention (Žarković and Ivanović 2020, 6). AI systems have significantly contributed to enhancing police operational efficiency and response time in crime prevention, as they have been assessed as highly effective in so-called "predictive policing".

Predictive policing represents a modern concept that involves the application of analytical techniques to data for the purposes of generating statistical predictions about events (McDaniel and Pease 2021, 7). Based on statistical data regarding crimes already committed within a specific time and location, extracted from various police records, it is possible to forecast when and where future crimes are likely to occur. The use of predictive analytics software yields more successful results in identifying locations with a high risk of criminal activity, thereby creating opportunities for timely intervention (Mohler *et al*. 2015).

AI software for crime risk assesment

Examples of AI-based software used by law enforcement agencies in the United States of America (hereinafter: U.S.) are *PredPol* and *HunchLab*; their use has proven particularly successful in the prevention of property crimes (Ratcliffe *et al*. 2021).

The software *PredPol* was developed through collaboration between university researchers and police analysts, and it was among the first to achieve commercial success and wider application across various police agencies. This software utilizes machine learning and an algorithm that was originally designed to predict aftershocks following an earthquake. It operates on the assumption that the occurrence of an earthquake increases the likelihood of another earthquake nearby in place and time (Mohler *et al*. 2011, 100). Similarly, according to the near-repeat theory, some future crimes (e.g., burglary, theft, or gang-

related violence) will occur very near to current crimes in time and place (Perry *et al*. 2013, 41).

The *PredPol* software analyzes patterns in incident records from the previous five years, focusing on a limited set of variables such as the type of crime, time, and place. Based on this analysis, it generates daily forecasts of "hotspots", which are visualized on maps (Ferguson 2019, 494). Police officers can easily access the maps via computer, after which patrol activity is redirected toward specific areas with an elevated risk of a particular type of crime. Active patrolling in "hotspot" areas produces a deterrent effect, which consequently leads to a reduction in crime rates (Hutt 2020, 36). Such an outcome was recorded in the police departments in Santa Cruz and Los Angeles, which used this software and reported a 12–25% decrease in the number of burglaries, car thefts, and thefts from automobiles, compared to the same period when the software was not in use. The software also proved successful in predicting the locations of 50% of gun homicides within a broader timeframe in Chicago, based on the analysis of data from previous crimes involving handguns (The predictive police company 2014, cited in: Ferguson 2017, 1134).

The geospatial software *HunchLab* employs a combination of several techniques from the AI area and machine learning, producing highly accurate crime predictions within the jurisdictions of specific police departments. Similar to *PredPool*, it uses algorithms based on the near-repeat theory but also incorporates other approaches that enhance its performance. For example, the *RTM* (*Risk Terrain Modeling*) approach is employed to generate a composite risk terrain map, which accounts for the spatial influences of all features at each place within the observed area (Caplan *et al*. 2015, 8). This means that *RTM* incorporates contextual information relevant to the social and physical environment in order to identify micro-locations within the area that have the highest level of risk of future incidents (Marchment and Gill 2021, 2).

Therefore, the algorithms of *HunchLab* learn not only from data contained in police crime reports but also from non-crime-related data, such as information about infrastructure (e.g., locations of bars, clubs, banks, metro stations, and other relevant places), temporal patterns, population density, socio-economic characteristics, and more (Degeling and Berendt 2017, 349). As Shapiro (2019) explains, this data is mapped onto a grid of cells, each covering an area of 500 square feet within the jurisdiction of the police department, and then categorized based on the crime outcomes in each cell. If a crime occurred in a cell, then

it is possible to determine which variables influenced the occasion of that crime and to what extent; variables are then weighted accordingly (462). The system automatically learns what is important for each type of crime, taking into account exogenous factors such as the location of crime-relevant objects – e.g., banks in the case of robberies, or relevant spots along a specific route where drivers may stop in the case of street prostitution (Caplan and Kennedy 2011). In addition, the software provides recommendations on where to focus available resources and suggests possible tactical responses for police officers – e.g., to collect information from citizens in a specific area, to conduct traffic stops and checks of vehicles and individuals, and so on.

AI software for recidivism risk assessment

In addition to software used for crime prediction, law enforcement agencies also utilize software focused on forecasting individual criminal behavior, specifically aimed at identifying individuals who are at high risk of reoffending. One such software is *COMPAS* (*Correctional Offender Management Profiling for Alternative Sanctions*), which employs advanced data analysis methods and machine learning techniques. This software is used by multiple institutions within the criminal justice and correctional systems to assess the risk posed by offenders and to develop appropriate treatment plans aimed at preventing reoffending. For example, the software generates risk scores that are taken into account in decision-making regarding release, probation, parole, and the planning of an individual's reintegration into society, as well as their supervision after release (Brennan and Dieterich 2018). The *COMPAS* is also used by judges in several U.S. states, such as New York, Pennsylvania, Florida, and others, in decision-making regarding pretrial detention, based on algorithms that assess the likelihood of reoffending (Engel *et al*. 2023, 383).

The software uses various data related to the defendant, collected from multiple sources. In addition to official records, it relies on data obtained from self-assessment questionnaires completed by defendants or inmates, as well as information gathered through standardized interviews with them (Blomberg *et al*. 2010). In risk assessment, approximately 15 different static and dynamic factors are used, such as the criminal involvement, history of violence, history of non-compliance, criminal associates, substance abuse, family status, financial

problems, employment and educational resources, social relationships and engagement, and others (Northpointe 2015). Each of these factors is measured using scales composed of multiple items, with a distinction made between risk scales and needs scales.[1]

On one hand, risk scales are designed to assess the risk of general recidivism, meaning the likelihood of reoffending, as well as the risk of violent recidivism, which refers to the commission of violent crimes such as homicide, aggravated assault, robbery, rape, and others (Lagioia *et al*. 2022, 464). On the other hand, needs scales are designed to measure the defendant's needs (e.g., in the domains of education, housing, employment, etc.) that should be addressed in order to reduce the risk of recidivism. In this way, they support decision-making in case planning tailored to the specific defendant, including the selection of effective interventions (Northpointe 2015). As a result, risk scores are presented as numerical values ranging from 1 to 10, indicating the individual's level of risk or the needs that require intervention. A low risk is assigned to individuals scoring between 1 and 4, medium risk to those between 5 and 7, and high risk to those scoring between 8 and 10 (Brennan and Dieterich 2018, 59).

A similar type of software is used by the police department in Durham, United Kingdom. The HART (*Harm Assessment Risk Tool*) software applies advanced statistical machine learning methods that are effectively used to analyze large volumes of data contained in police records. This software provides an assessment of a suspect's risk of reoffending after they have been arrested by the police and brought before the custody officer, primarily by using predictors related to their history of criminal behavior (Oswald *et al*. 2018, 228). Suspects are separated based on the likelihood that they will commit a new serious crime over the next 24 months, taking into account the potential harm that such a crime may cause. If it is assessed that the individual is unlikely to commit any crime in the mentioned timeframe, they are identified as low risk (Urwin 2016, 15). Individuals are classified as moderate risk if it is estimated that they are likely to commit a crime of lower social harm (a "non-serious" crime), while high risk is assigned when the prediction

[1] For example, within the scale related to the defendant's criminal involvement, items include the total number of prior arrests and convictions, frequency of incarceration, and so on, while the scale for history of non-compliance considers items such as failure to appear in court, positive drug tests, revocation of probation, and similar factors (Brennan, Dieterich and Ehret 2009, 34).

indicates indicates a likelihood of the commission of a "serious" crime, such as homicide, attempted murder, grievous bodily harm, robbery, sexual offenses, or crimes involving firearms (Oswald *et al*. 2018, 227). If a moderate risk is identified, a suspect may be considered for the application of an out-of-court disposal provided for within the special programme named *Checkpoint programme*. This programme offers the individual an opportunity to resolve their criminal behavior outside of court through interventions addressing its underlying causes, which may include counseling, education, rehabilitation programs, and support services (Weir *et al*. 2022).

In addition to the benefits achieved through the use of the aforementioned software, numerous controversies arise regarding the extent of their influence on decision-making by competent authorities, which will be discussed in more detail in a dedicated chapter.

AI software for victimization risk assessment

Police agencies worldwide have been using structured risk assessment tools for decades to evaluate the likelihood that an individual will become a victim of a crime, such as domestic violence, crimes involving firearms, stalking, child abuse, and others. For example, the police in Ontario, Canada, use the *ODARA* (*Ontario Domestic Assault Risk Assessment*) tool to evaluate the risk of repeated domestic violence against women, based on multiple factors such as prior domestic violence, confinement of the victim, number of children, assault on the victim during pregnancy, substance abuse, threats of violence, and others (Hilton *et al*. 2008, 151).

The development of AI has enabled the creation of new victimization risk assessment software that utilizes machine learning methods. Although some of these software tools were primarily developed for crime prediction and prevention, they can also be used to assess the risk of victimization. For example, the *NDAS* (*National Data Analytics Solution*) system combines advanced analytics and statistical techniques. Although the system is owned by the West Midlands Police, it is also utilized by several other police forces and agencies across the United Kingdom (Zilka *et al*. 2022, 883). It is an individual-focused system primarily developed for crime prediction and prevention; however, it can also be used for the purpose of assessing the risk of victimization (Fair Trials 2021, 15). In order to generate numerical scores indicating

the risk of future victimization for individuals listed in the database, this system uses data stored across multiple police records (West Yorkshire Police n.d; West Midlands Police 2023). For example, it uses records of crimes committed, police intelligence reports about events, locations, and offenders, custody and prisoner information, information on people, crimes, vehicles, and property, records of organised criminal groups, and others (Big Brother Watch 2020, 12). The use of this system has sparked numerous debates and criticisms, with controversies further intensified by expressed intentions to integrate data from other public services into the system, including healthcare, education, social welfare, local government, and others (Fair Trials 2021, 15).

The Chicago Police Department uses the *CVRM* (*The Crime and Victimization Risk Model*) tool for assessment of the risk that an individual would be a party to gun violence, either as a victim or perpetrator. This tool uses crime-related data and includes key risk factors such as the number of past shooting victimizations, age at last arrest, victimization due to aggravated burglary and assault, arrests related to unauthorized use of a weapon, and violent crimes (Hollywood *et al*. 2019, 12–13). The degree of risk is quantified using numerical values – the higher the number, the greater the risk that an individual will become a victim or be arrested for gun-related violence within the following 18 months (Chicago Police Department n.d.).

One of the more recent tools developed in Germany and commercially available is *Lizzy*. It is designed to assess the risk of domestic violence and is intended for use by police officers and other professionals who have initial contact with victims. The tool uses machine learning algorithms to predict repeated physical violence based on six factors, which involve the use of various forms of violence of differing intensity (Trafford *et al*. 2024). Currently, this tool is used by humanitarian workers from *Frontline* in several German provinces who conduct interviews with victims in the field and complete a brief questionnaire containing key questions, i.e., risk factors (Frontline n.d.). Based on the entered responses, the algorithm calculates the risk of repeated physical violence within the following three months. Results from its application indicate that it is an exceptionally useful tool, particularly because it can be easily used by professionals conducting fieldwork.

APPLICATION OF AI SYSTEMS IN CRIME SUPPRESSION

Generative AI can have wide application in the automation of administrative activities within criminal police departments, as it generates high-quality and realistic content that is almost indistinguishable from content created by a human being (Banh and Strobel 2023, 1). In this way, many activities can be simplified and accelerated without the need for additional personal or material resources. When establishing or verifying facts related to a crime, police officers are required to create various written documents, such as criminal complaints, official notes, records, reports, and others. Generative AI models, such as *GPT* (*Generative Pre-trained Transformer*), can play a significant role in this regard – by simply entering input data that describes the actions undertaken and the established facts, one can obtain contextually appropriate textual content tailored to a specific type of official document (Dubravova *et al*. 2024, 240).

The advancement of translation technologies, driven by the innovative development of new AI-based techniques, also plays a significant role in supporting the operations of the criminal police department. This has led to the development of new tools in the area of translation and language services. Police officers can easily access conversational chatbots, such as *ChatGPT*, which offer unprecedented efficiency in translating large volumes of text and enabling real-time communication (Siu 2024, 29). The use of translation technologies contributes to the effectiveness of the police in conducting criminal investigations in several ways – e.g., it facilitates the understanding of communications between different offenders who do not belong to the national language area, as well as communication and data exchange with other police agencies and institutions in the context of international operational police cooperation.

Intake of criminal reports and the evaluation of the completeness and truthfulness of their content

The development of AI has also led to the increasingly intensive use of technology-mediated reporting of crimes, which replaces or supplements traditional methods of reporting via telephone or by appearing in person at a police station. Instead of human dispatchers,

numerous police agencies have started using various types of chatbots on their websites, through which citizens can report minor crimes at any time (Bradford *et al*. 2025, 3). For example, the police in Bedfordshire, United Kingdom, use chatbots to handle reports related to the loss, damage, or theft of low-value property, as well as those indicating animal abuse or neglect (Muir and O'Connell 2025, 12). In the Netherlands, there is a specialized system that uses an intelligent form for reporting online fraud, such as cases involving fake online stores and malicious sellers on e-commerce platforms (Odekerken *et al*. 2022). The system can automatically identify deficiencies in the content of a report and ask citizens questions to gather additional relevant information, thereby reducing the need for the police to request supplements and speeding up case processing time (Odekerken 2024).

AI-based tools can also be useful for detecting false reports. For example, the *VeriPol* tool has been implemented in nearly all stations of the National Police in Spain and demonstrates a high level of accuracy in identifying false reports, especially concerning property crimes[2] (Ramos *et al*. 2020). This tool uses algorithms that consider around 300 variables and is based on natural language processing; it identifies patterns that are repeatedly used in report submission and calculates the probability of their truthfulness (Álvarez 2019). In this way, the police officer is alerted to the possibility of a false report, which assists in determining further action in the specific case.

The planning and conduct of criminal investigations

Within criminal investigations, the police undertake a range of measures and actions, both operational and evidentiary. Their successful execution is conditioned by the systematic collection of relevant and complete data from various sources. An example of this is the planning of stakeout operations by the police. Their successful execution depends on the systematic collection of relevant and complete data from various sources. For example, to prepare a plan for a stakeout operation, it is necessary to gather all available data regarding suspect whose arrival is expected at a specific location, including information about their close contacts, criminal associates, and their movements; in addition, data is collected concerning

[2] For example, theft or loss of mobile phones without elements of violence, loss of a larger amount of money, thefts from a house where valuable items have gone missing, etc.

the specific stakeout location (e.g. terrain configuration, layout of nearby structures, etc.), weather and atmospheric conditions, vehicles used, and more (Aleksić and Škulić 2016, 87). For this purpose, police records can be used, such as the suspect's criminal file, *modus operandi* records, past police reports, the unified information system of the police, and so on.[3] Data can also be collected through the implementation of other measures and actions – e.g., by gathering information from persons of interest, based on the analysis of obtained telephone communication records, the base stations used, and the location of the place from where a communication was performed, etc. Generative AI can significantly facilitate the planning and preparation of a stakeout operation by processing and analyzing input data. In addition to generating relevant recommendations regarding the timing and location of a stakeout, as well as predicting multiple possible scenarios, the end result would be the creation of an operational plan for a stakeout[4] (see: Dubravova *et al*. 2024).

Searching records for the purpose of collecting operational data often results in significant time consumption and carries the risk of overlooking certain information. Therefore, specialized tools have been developed to enable the linking of various data from police records and to facilitate easy search capabilities, such as the *Coplink* system and its enhanced versions, *CoplinkX* and *CrimeTracer*. In its original version, the system used algorithmic techniques to identify links between objects of interest within the database, which made it useful in a policing environment. For example, crime analysts and investigators at the Tucson Police Department, Arizona, used this system to support investigative activities. Based on the developed and structured database tailored to their needs, police officers had the ability to enter search terms through various forms, after which the system would display related terms to them (Hauck *et al*. 2002).[5]

[3] For example, whether the person possesses any weapons, where, when, and with whom they were stopped by the police and their identity checked, which vehicles they use, whether they have been recorded as involved in any incidents reported to the police, etc.

[4] Such a plan should include all key elements, such as: the objective of a stakeout; the location or multiple locations where a stakeout will be set-up; the timing and duration of a stakeout; its formation and method of deployment; the number of police officers to be engaged; and the required equipment and technical resources (Žarković and Ivanović 2020, 182).

[5] In this way, they were able to link a specific suspect to a vehicle or to other individuals; to enter vehicle data into a search in the system and determine whether

An enhanced version of this system represents a powerful platform designed for law enforcement agencies in the U.S., as it features a centralized interface[6] through which it is now possible to search a large volume of data, both structured and unstructured (SoundThinking n.d.). Users of the *CrimeTracer* system can perform intelligent searches using an AI-based chatbot by simply entering terms and concepts in natural language. The system enables advanced link analysis to quickly detect relationships between people, locations, and events; linking ballistic evidence from the national database with reports, suspects, and other entities; identification of crime trends, and more (SoundThinking n.d.).

Determining similarities between crimes based on the analysis of police reports and linking them into a series can also be enhanced by using AI-based systems. For example, the *VALCRI* (*Visual Analytics for Sense-making in Criminal Intelligence Analysis*) system uses machine learning algorithms to search through large volumes of data by analyzing semantic features in the textual descriptions of crimes contained in police reports, which are then grouped based on identified similarities (Sacha *et al*. 2017, 5). In this way, an analyst can obtain various types of reports, such as comparative case analyses. The system can identify new connections between criminal entities based on which similar cases can be linked – i.e., when similar criminal behaviour or *modus operandi* appeared in several cases, or when there is a geographic or temporal relationship between the cases, etc (EC 2018b). This system has been tested and used by multiple police agencies across Europe, as well as by *EUROPOL* (*European Union Agency for Law Enforcement Cooperation*).

For the effective preparation and conduct of criminal investigations, systems based on artificial intelligence that are used for public space surveillance are also of importance. For example, in China, a special program called *Skynet* was launched, aiming to integrate a large number of surveillance cameras (both private and public) into a unique platform. In this way, the capabilities for identifying people in China's largest cities are at the highest possible level and enable real-time determination of the number of people at a given location, their gender, clothing

it appears in any case and which persons are associated with it; to cross-reference the suspect's nickname with the victim's name in order to establish a possible connection, and so on (Hauck *et al.* 2002).

[6] This platform provides access to data contained in over a billion law enforcement records from various agencies across the U.S. (SoundThinking n.d.).

characteristics, as well as vehicles used (Fornasier and Borges 2023, 448). For instance, within the *Skynet* system, several subsystems operate, such as: a system that capture faces in video footage in real time and compares the created "digital fingerprint" of the face with a database of known faces to determine matches; a face retrieval system, which allows targeted searches for individuals across a large number of recordings; as well as a system post-retrival system that enables searching and analysis of already recorded footage for the purpose of clarifying certain events relevant to the police (Qiang 2021, 36–37).

International operational police cooperation

The use of AI has significantly contributed to and facilitated the collection and processing of large volumes of data held by international organizations engaged in police cooperation in combating crime, such as *INTERPOL* (*International Criminal Police Organization*) and *EUROPOL* (Ramos *et al.* 2020). For example, *INTERPOL* has implemented an analytical platform within the *INSIGHT project* that utilizes advanced technology and AI-based analytical techniques for data processing and the production of criminal intelligence analyses (INTERPOL 2020). The platform enables the search of data contained in various records and in different data formats, like notices and diffusions, free text information, police reports, emails, spreadsheets, images, and videos, thereby facilitating faster detection of hidden connections, patterns, and crime trends (INTERPOL n.d.). In this way, *INTERPOL* has enhanced its capabilities in distributing intelligence information to member states during the conduct of criminal investigations, strengthening cooperation and data exchange in combating transnational crime.

Similarly, *EUROPOL* uses the system *SIENA* (*Secure Information Exchange Network Application*)**.** Within this system, it is possible to exchange operational and intelligence information about crime in the form of messages, documents, and multimedia content (EUROPOL 2025). The interface is multilingual, enabling communication and data exchange in multiple languages, and operators using this system have the ability to conduct searches and generate statistical reports based on the exchanged data (Council of the European Union 2019, 70). The usefulness of this platform is demonstrated by the fact that in 2023, 1.79 million messages were exchanged, with over 151,000 cases

initiated, mostly in crime areas such as drug trafficking, fraud, and illegal immigration (EUROPOL 2024).

CONTROVERSIES IN THE APPLICATION OF AI WITHIN LAW ENFORCEMENT SYSTEMS

The introduction of AI systems also entails their responsible use within the law enforcement system. In this regard, a number of complex questions arise, ranging from the types of AI systems or software used by law enforcement agencies, and the conditions under which they are procured from private companies that develop them, to whether the public is informed about the deployment of such systems and how they function in practice (Joh 2017). The question of the justification for using AI systems also arises, given that there is still no absolute consensus regarding their actual effectiveness, particularly in predicting crime and recidivism (Van Brakel 2025). This has led some police and other agencies to discontinue their use.

In general, controversies associated with the use of AI within law enforcement systems can be considered from multiple perspectives, including technological, legal, ethical, and social. Technological controversies are related to the limitations of the systems themselves, as well as to the manner in which data is processed and results are generated. One of the important issues is the potential bias inherent in such systems. This is particularly important given that AI systems are only as effective as the quality of the data on which they are trained (Feldstein 2019a, 47). AI-based software used in policing environments analyzes data that primarily comes from archived police records. It is well known that such data may be incomplete, and an even more complex issue is that it can be biased, potentially leading to the discrimination of certain communities.

One of the main criticisms directed at the use of crime prediction software, such as *PredPol* and *HunchLab*, is that they rely on historical police data, which may be discriminatory toward certain communities residing in areas identified as high-risk for criminal activity. In order to improve the effectiveness of police patrols, officers are expected to act proactively, which in practice means patrolling more frequently in areas identified by the software as high-risk. Considering that patrol performance quotas are, among other factors, measured by the number of individuals stopped, identified, and searched during a duty shift, it is not difficult to conclude that certain communities will be disproportionately

affected by such increased police presence and surveillance, regardless of the actual crime rates in those areas (Benbouzid 2019). This can be illustrated by an article published by *The Markup*, a U.S. non-profit newsroom. This article points out that the predictions generated by the *PredPol* system, used by law enforcement agencies in the U.S., disproportionately focus on neighborhoods that are home to Black and Latino communities, compared to White communities[7] (Kirchner and Goldstein 2021). Identified biases in the system can lead to violations of the human rights of such communities, including the right to privacy, and result in their further marginalization.

Similar criticisms can be directed at software that generates recidivism predictions, as profiles of high-risk individuals are created based on patterns identified in the data analyzed by algorithms (Van Brakel 2025). It is known that such software is used in criminal justice systems in many countries as a support tool for decision-makers. However, in practice, it is not uncommon for a high-risk assessment generated by the software to have a decisive influence on decisions, which can sometimes lead to unfair outcomes (See Oswald *et al*. 2018). For example, the defendant may be held in custody until trial, may receive a prison sentence instead of a suspended sentence, may be denied parole, and so on (Engel *et al*. 2024). In a study examining cases in Florida, U.S., where the *COMPAS* system was used, a discussion was opened regarding its accuracy; it was found that individuals assessed as high-risk often did not reoffend, while others assessed as low-risk committed new crimes (Lagioia *et al*. 2022). Additionally, controversy has arisen over the presence of racial bias in generating risk assessments, as it has been found that this system produces more favorable recidivism predictions for White defendants compared to Black defendants, who are twice as likely to be assessed as high-risk (Dressel and Hany 2018).

Ethical and social controversies in the application of AI systems within law enforcement also arise because their algorithms analyze large volumes of data, which are not necessarily related solely to criminal activity. This means that data of individuals who have never committed any crime may also be analyzed, raising concerns about the right to privacy and freedom, as well as the protection of personal data. Even

[7] These communities were already living in socioeconomically disadvantaged conditions, compared to white communities, which generally belonged to the middle to upper-income groups.

greater controversy is fueled by the fact that AI systems are used for mass surveillance purposes, which is neither transparent nor properly controlled, potentially leading to various misuses (Feldstein 2019b).

Issues related to compliance with the principle of transparency are also highlighted, stemming from the fact that law enforcement agencies are actually the end-users of artificial intelligence systems and are most often not directly involved in their design. The owners of various software solutions used in police practice are private companies, which frequently refuse to publicly share information regarding how the software was developed and on which algorithms and principles it operates, citing commercial confidentiality (Zilka *et al*. 2022, 887). Therefore, it is necessary to introduce an effective mechanism requiring private companies to disclose information about the internal mechanisms and applications of algorithms, thereby enabling public oversight (Lu 2020). Furthermore, it is important not to overlook the fact that private companies producing such systems may become owners of sensitive data collected in the police and judicial contexts (Brauneis and Goodman 2018). The mentioned deficiencies inherent in the design and application of AI systems must be addressed appropriately to ensure respect for fundamental principles such as human oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, accountability, and others (See Prlja *et al*. 2022).

CONCLUSION

The development of AI has significantly contributed to the modernization of law enforcement systems. Law enforcement agencies utilize various systems, software, and tools in their work that have the potential to advance crime prevention and suppression efforts. The use of different AI-based software has enabled greater efficiency and speed in processing large volumes of data, which previously exceeded human capabilities. Conclusions generated through its application provide intelligent work support and also facilitate the decision-making processes of the police, courts, and other law enforcement agencies. The ability to process various operational data in less time has significantly simplified the conduct of administrative police activities, and enabled more efficient planning and documenting of investigative actions. Furthermore, the establishment of analytical platforms utilizing AI-based analyses has led to significant progress in international operational police cooperation,

particularly regarding the ability to exchange data between countries in combating transnational crime.

However, the benefits of AI systems in this area cannot be considered independently of the challenges that accompany their design and practical implementation. Therefore, it is essential to ensure adherence to key principles in order to mitigate existing errors and deficiencies of AI systems. In their application, it is important that human beings retain final control over decision-making processes that impact the lives of individuals – e.g., judges should not automatically base their decisions solely on AI software assessments but must carefully evaluate each individual case and its specific circumstances; police officers should critically reassess generated risk assessments in the context of their legitimacy, utilizing their experience and knowledge of the local community in which they operate, and align their preventive and repressive actions in accordance with the broader social context.

The technical robustness and safety of AI systems must be ensured, and the data used within them must be collected, stored, and utilized in a lawful and ethical manner. It is important to ensure compliance with the principle of transparency – the public, and especially individuals who may be affected by decisions made based on AI systems, must be clearly informed about how these systems operate, which data they use, and how they generate conclusions. Ultimately, it is important to establish appropriate legal and regulatory frameworks that ensure responsible and ethical use of AI systems, protect human rights, and uphold the principles of equality and fairness, while also determining accountability for potential errors and human rights violations. In this regard, it is essential to have adequate response mechanisms that enable the detection, investigation, and remediation of consequences arising from errors and other deficiencies manifested in the operation of AI systems.

REFERENCES

Aleksić, Živojin, and Milan Škulić. 2016. *Kriminalistika.* Beograd: Pravni fakultet Univerziteta u Beogradu.

Álvarez, Ramon. 2019. "La inteligencia artificial de la Policía que desenmascara denuncias falsas." *LaVanguardia*. April 13, 2019. https://www.lavanguardia.com/tecnologia/20190414/461583468024/veripol-policia-nacional-inteligencia-artificial-algoritmo-denuncias-falsas.html

Banh, Leonardo, and Strobel Gero. 2023. "Generative artificial intelligence." *Electronic Markets* 33 (1): 1–17. DOI: 10.1007/s12525-023-00680-1

Benbouzid, Bilel. 2019. "To predict and to manage. Predictive policing in the United States." *Big Data & Society* 6 (1): 205395171986170. DOI: 10.1177/2053951719861703

Big Brother Watch. 2020. *Big Brother Watch briefing on algorithmic decision-making in the criminal justice system*. London: Big Brother Watch.

Blomberg, Thomas, William Bales, Karen Mann, Ryan Meldrum, and Joe Nedelec [Blomberg *et al.*]. 2010. *Validation of the COMPAS risk assessment classification instrument*. Tallahassee, FL: College of Criminology and Criminal Justice, Florida State University.

Bradford, Ben, Arabella Kyprianides, Will Andrews, Elizabeth Aston, Estelle Clayton, Megan O'Neill, and Helen Wells [Bradford *et al.*]. 2025. "'To whom am I speaking?'; Public responses to crime reporting via live chat with human versus AI police operators." *Policing and Society* 35: 1036–1052. DOI: 10.1080/10439463.2025.2453437

Brauneis, Robert, and Ellen P. Goodman. 2018. "Algorithmic transparency for the smart city." *Yale Journal of Law & Technology* 20: 103–176. DOI: 10.7282/00000058

Brennan, Tim, and William Dieterich. 2018. "Correctional offender management profiles for alternative sanctions (COMPAS)." In *Handbook of recidivism risk/needs assessment tools*, 1st edition, eds. Jay P. Singh, Daryl G. Kroner, J. Stephen Wormith, Sarah L. Desmarais, and Zachary Hamilton, 49–75. Hoboken, NJ: John Wiley & Sons, Ltd.

Brennan, Tim, William Dieterich, and Beate Ehret. 2009. "Evaluating the predictive validity of the COMPAS risk and needs assessment system." *Criminal Justice and Behavior* 36 (1): 21–40. DOI: 10.1177/0093854808326545

Caplan, Joel M., and Leslie W. Kennedy. 2011. *Risk terrain modeling compendium: for crime analysis*. New Jersey: Rutgers Center on Public Security.

Caplan, Joel M., Leslie W. Kennedy, Jeremy D. Barnum, and Eric L. Piza [Caplan *et al.*]. 2015. "Risk terrain modeling for spatial risk assessment." *Cityscape* 17 (1): 7–16.

Chicago Police Department. n.d. "Crime and Victimization Risk Model (CVRM)." *Chicago Police Department*. Last accessed on June 16, 2025. https://www.chicagopolice.org/wp-content/uploads/FACT-SHEET-Crime-and-Victimization-Risk-Model-1.pdf

Council of the European Union. 2019. *Manual on Law Enforcement Information Exchange*. Brusseles: Council of the European Union.

Degeling, Martin, and Bettina Berendt. 2017. "What is wrong about Robocops as consultants? A technology-centric critique of predictive policing." *AI & Society* 33 (3): 347–356. DOI: 10.1007/s00146-017-0730-7

Dressel, Julia, and Farid Hany. 2018. "The accuracy, fairness, and limits of predicting recidivism." *Science Advances* 4 (1): eaao5580. DOI: 10.1126/sciadv.aao5580

Dubravova, Hana, Jan Cap, Kristyna Holubova, and Lukas Hribnak [Dubravova *et al*.]. 2024. "Artificial intelligence as an innovative element of support in policing." *Procedia Computer Science* 237: 237–244. DOI: 10.1016/j.procs.2024.05.101

Engel, Christoph, Lorenz Linhardt, and Marcel Schubert [Engel *et al*.]. 2024. "Code is law: how COMPAS affects the way the judiciary handles the risk of recidivism." *Artificial Intelligence and Law* 33: 383–404. DOI: 10.1007/s10506-024-09389-8

European Commission [EC]. 2018a. "Artificial Intelligence for Europe." *EUR-Lex, Access to European Union law*. April 25, 2018. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=COM:2018:237:FIN

European Commission [EC]. 2018b. "*Visual Analytics for Sense-making in Criminal Intelligence Analysis (VALCRI).*" *CORDIS EU research results*. October 18, 2018. https://cordis.europa.eu/project/id/608142/reporting

EUROPOL. 2024. "More than 3 000 law enforcement authorities now connected to Europol." *EUROPOL*. April 12, 2024. https://www.europol.europa.eu/media-press/newsroom/news/more-3-000-law-enforcement-authorities-now-connected-to-europol

EUROPOL. 2025. "Secure Information Exchange Network Application (SIENA)." *EUROPOL*. August 13, 2024. https://www.europol.europa.eu/operations-services-and-innovation/services-support/information-exchange/secure-information-exchange-network-application-siena

Fair Trials. 2021. "Automatic injustice: The use of artificial intelligence & automated decision-making systems in criminal justice in Europe." *Fair Trials*. Last accessed on May 1, 2025. https://www.fairtrials.org/app/uploads/2021/11/Automating_Injustice.pdf

Feldstein, Steven. 2019a. "How artificial intelligence is reshaping repression." *Journal of Democracy* 30 (1): 40–52. DOI: 10.1353/jod.2019.0003

Feldstein, Steven. 2019b. *The Global Expansion of AI Surveillance*. Washington, DC: Carnegie Endowment for International Peace.

Ferguson, Andrew. 2017. "Policing predictive policing." *Washington University Law Review* 94 (5): 1109–1189.

Ferguson, Andrew. 2019. "Predictive Policing Theory." In *The Cambridge handbook of policing in the United States–1st ed,* eds. Tamara Lave and Erick Miller, 491–510. UK: Cambridge University Press.

Fornasier, Mateus de Oliveira, and Gustavo Silveira Borges. 2023. "The Chinese "Sharp Eyes" system in the era of hyper surveillance: Between State use and risks to privacy." *Revista Brasileira de Políticas Públicas* 13 (1): 440–453.

Frontline. n.d. "Meet Lizzy: your team's assistant for assessing domestic abuse risk." *Frontline*. Last accessed on June 10, 2025. https://www.frontline100.com/product

Hauck, Roslin V., Atabakhsh Homa, Ongvasith Pichai, Gupta Harsh, and Chen Hsinchun [Hauck *et al*.]. 2002. "Using Coplink to analyze criminal-justice data." *Computer* 35 (3): 30–37.

Hilton, N. Zoe, Grant T. Harris, Marnie E. Rice, Ruth E. Houghton, and Angela W. Eke [Hilton *et al*.]. 2008. "An indepth actuarial assessment for wife assault recidivism: The domestic violence risk appraisal guide." *Law and Human Behavior* 32 (2): 150–163. DOI: 10.1007/s10979-007-9088-6

Hollywood, Johns S., Kenneth N. McKay, Dulani Woods, and Denis Agneil [Hollywood *et al*.]. 2019. *Real-Time Crime Centers in Chicago*. Santa Monica, CA: RAND Corporation.

Hutt, Oliver Kenneth. 2020. "Understanding the deterrent effect of police patrol." PhD diss. UCL: University College London.

INTERPOL. 2020. "INSIGHT INTERPOL analytical platform: Broshure." *INTERPOL*. Last accessed on June 20, 2025. https://www.interpol.int/Resources/Documents#Publications

INTERPOL. n.d. "Project INSIGHT. INTERPOL." *INTERPOL*. Last accessed on June 20, 2025. https://www.interpol.int/en/How-we-work/Criminal-intelligence-analysis/Projects/Project-INSIGHT

Joh, Elizabeth E. 2017. "Artificial intelligence and policing: First questions." *Seattle University Law Review* 41 (4): 1139–1144.

Kirchner, Lauren, and Rose Goldstein. 2021. "PredPol's predictive policing software disproportionately targeted Black and Latino neighborhoods." *The Markup*. December 2, 2021. https://themarkup.org/prediction-bias/2021/12/02/crime-prediction-software-promised-to-be-free-of-biases-new-data-shows-it-perpetuates-them

Lagioia, Francesca, Riccardo Rovatti, and Giovanni Sartor [Lagioia *et al*.]. 2023. "Algorithmic fairness through group parities? The case of COMPAS-SAPMOC." *AI & Society* 38 (2): 459–478. DOI: 10.1007/s00146-022-01441-y

Lu, Sylvia. 2020. "Algorithmic opacity, private accountability, and corporate social disclosure in the age of artificial intelligence." *Vanderbilt Journal of Entertainment and Technology Law* 99 (1): 99–160.

Marchment, Zoe, and Paul Gill. 2021. "Systematic review and meta-analysis of risk terrain modelling (RTM) as a spatial forecasting method." *Crime Science* 10 (1): 1–11. DOI: 10.1186/s40163-021-00149-6

Martinez, Rex. 2018. "Artificial intelligence: Distinguishing between types & definitions." *Nevada Law Journal* 19 (3): 1015–1042.

McDaniel, John, and Ken Pease, eds. 2021. *Predictive policing and artificial intelligence*. London: Routledge, Taylor & Francis Group.

Mohler, George, Martin Short, Jeffrey Brantingham, Frederick Schoenberg, and George Tita [Mohler e*t al.*]. 2011. "Self-exciting point process modeling of crime." *Journal of the american statistical association* 106 (493): 100–108. DOI: 10.1198/jasa.2011.ap09546

Mohler, George, Martin Short, Sean Malinowski, Mark Johnson, George Tita, Andrea Bertozzi, and Jeffrey Brantingham [Mohler *et al.*]. 2015. "Randomized controlled field trials of predictive policing." *Journal of the American Statistical Association* 110 (512): 1399–1411. DOI: 10.1080/01621459.2015.1077710

Muir, Rick, and Felicity O'Connell. 2025. *Policing and artificial intelligence*. London: Police Foundation.

Nehra, Ekta. 2015. *Artificial intelligence in modern times*. New Delphy: ICRISEM, YMCA.

Northpointe. 2015. *Practitioners Guide to COMPAS*. Traverse City, MI: Northpointe.

Odekerken, Daphne. 2024. "AI tool helps citizens file complete reports." *Utrecht University*. December 19, 2024. https://www.uu.nl/en/news/ai-tool-helps-citizens-file-complete-reports

Odekerken, Daphne, Floris Bex, AnneMarie Borg, and Bas Testerink [Odekerken *et al*.]. 2022. "Approximating stability for applied argument-based inquiry." *Intelligent Systems with Application* 16 (3): 200110. DOI: 10.1016/j.iswa.2022.200110

Oswald, Marion, Jamie Grace, Sheena Urwin, and Geoffrey C. Barnes [Oswald *et al*.]. 2018. "Algorithmic risk assessment policing models: lessons from the Durham HART model and 'Experimental'

proportionality.'" *Information & Communications Technology Law* 27 (2): 223–250. DOI: 10.1080/13600834.2018.1458455

Perry, Walter L., Brian McInnis, Carter C. Price, Susan Smith, and John S. Hollywood [Perry *et al*.]. 2013. *Predictive policing: The role of crimeforecasting in law enforcement operations*. Washington, DC: RAND Corporation.

Prlja, Dragan, Gordana Gasmi, and Vanja Korać [Prlja *et al*.]. 2022. *Ljudska prava i veštačka inteligencija*. Beograd: Institut za uporedno pravo.

Qiang, Xiao. 2021. "Chinese digital authoritarianism and its global impact." In *POMEPS studies 43: Digital activism and authoritarian adaptation in the Middle East*, ed. Marc Lynch, 35–40. Washington, DC: George Washington University and Institute for Middle East Studies.

Ramos, Sofia, José Ángel Perez-Lopez, and Rute Abreu [Ramos *et al*.]. 2020. "An analysis of the importance of the artificial intelligence on the information system of police forces." In *Proceedings of the 15th Iberian Conference on Information Systems and Technologies (CISTI 2020)*, 1–6. Seville, Spain: Iberian Association for Information.

Ratcliffe, Jerry, Ralph Taylor, Amber Askey, Kevin Thomas, John Grasso, Kevin Bethel, Ryan Fisher, and Josh Koehnlein [Ratcliffe *et al*.]. 2021. "The Philadelphia predictive policing experiment." *Journal of Experimental Criminology* 17 (3): 15–41. DOI: 10.1007/s11292-019-09400-2

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on Artificial Intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), OJ L, 12. 7. 2024.

Sacha, Dominik, Wolfgang Jentner, Leishi Zhang, Florian Stoffel, Geoffrey Ellis, and Daniel Keim [Sacha *et al*.]. 2017. "Applying visual interactive dimensionality reduction to criminal intelligence analysis." In *VALCRI WhitePaper Series*, ed. B.L. William Wong, 1–11. Konstanz, Germany: Universität Konstanz, and London, UK: Middlesex University.

Shapiro, Aaron. 2019. "Predictive policing for reform? Indeterminacy and intervention in big data policing." *Surveillance & Society* 17 (3/4): 456–472. DOI: 10.24908/ss.v17i3/4.10410

SoundThinking. n.d. "CrimeTracer™: Generate immediate investigative leads with a proven tool used by 2,500 agencies." *SoundThinking*. Last accessed on June 20, 2025. https://www.soundthinking.com/law-enforcement/crime-analysis-crimetracer/

Siu, Sai Cheong. 2024. "Revolutionizing translation with AI: Unravelling neural machine translation and generative pre-trained large language models." In *New advances in translation technology: Applications and pedagogy*, eds. Yuhong Peng, Huihui Huang and Defeng Li, 29–54. Singapore: Springer Nature Singapore.

Trafford, Lucy, Le Ba Linh, Firtala Sabina, and Babatunde Williams [Trafford *et al*.]. 2024. "Lizzy–Building an AI-powered domestic abuse risk assessment tool based on nationally representative online survey data." *OSF PREPRINTS:* 1–36. DOI: 10.31219/osf.io/shpjm

Urwin, Sheena. 2016. "Algorithmic forecasting of offender dangerousness for police custody officers: An assessment of accuracy for the Durham Constabulary." Master's thesis. University of Cambridge.

Van Brakel, Rosamunde. 2025. "Legal, ethical, and social issues of AI and law enforcement in Europe: The case of predictive policing." In *The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence*, ed. Nathalie Smuha (ed.), 367–382. UK: Cambridge University Press.

Weir, Kevin, Stephanie Kilili, John Cooper, Andrew Crowe, and Gillian Routledge [Weir *et al*.]. 2022. "Checkpoint: an innovative Programme to navigate people away from the cycle of reoffending–a randomised control trial evaluation." *The Police Journal: Theory, Practice and Principles* 95 (3): 562–589. DOI: 10.1177/0032258X211018774

West Midlands Police. 2023. *Data Protection Impact Assessment (DPIA) – Stage 1.* West Midland: West Midlands Police.

West Yorkshire Police. n. d.. "National Data Analytics Solution (NDAS) Privacy Notice." *West Yorkshire Police*. Last accessed on June 12, 2025. https://www.westyorkshire.police.uk/advice/modern-slavery/national-data-analytics-solution-ndas-privacy-notice

Zilka, Miri, Holli Sargenat, and Adrian Weller [Zilka *et al*.]. 2022. "Transparency, governance and regulation of algorithmic tools deployed in the criminal justice system: a UK case study." In *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, eds. Vincent Conitzer, John Tasioulas, Matthias Scheutz, Ryan Calo, Martina Mara, and Annette Zimmermann, 880–889. New York: Association for Computing Machinary.

Žarković, Milan, and Zvonimir Ivanović. 2020. *Kriminalistička taktika – treće izdanje*. Beograd: Kriminalističko-policijski univerzitet.

Марија Тасић[*]

Криминалистичко-полицијски универзитет, Србија

Александар Ивановић[**]

Правни факултет, Универзитет у Црној Гори, Црна Гора

Тања Кесић[***]

Криминалистичко-полицијски универзитет, Србија

Ивана Бјеловук[****]

Криминалистичко-полицијски универзитет, Србија

МОДЕРНИЗАЦИЈА СИСТЕМА ЗА СПРОВОЂЕЊЕ ЗАКОНА: УПОТРЕБА САВРЕМЕНИХ АЛАТА ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ У ПРЕВЕНЦИЈИ И СУЗБИЈАЊУ КРИМИНАЛИТЕТА

Резиме

Агенције за спровођење закона представљају један од најзначајнијих корисника система заснованих на вештачкој интелигенцији у контексту савремених манифестација криминалитета. Полицијске агенције, као и оне које делују у оквиру система кривичног правосуђа, већ увелико користе различите системе, софтвере и алате у области превенције и сузбијања криминалитета. Овај чланак анализира примену постојећих система вештачке интелигенције од стране агенција за спровођење закона, кроз критички преглед релевантне академске и стручне литературе, заједно са илустрацијом примера система, софтвера и алата који се користе у оперативном контексту. У првом поглављу, представљене се могућности примена система вештачке интелигенције у области превенције криминалитета, кроз приказ конкретних софтвера за предикцију криминалитета и

[*] Имејл: marija.tasic@kpu.edu.rs; ORCID: 0000-0002-3755-4854
[**] Имејл: ialeksandar@t-com.me; ORCID: 0000-0002-8186-0883
[***] Имејл: tanja.kesic@kpu.edu.rs; ORCID: 0000-0001-9473-4379
[****] Имејл: ivana.bjelovuk@kpu.edu.rs; ORCID: 0000-0002-9370-8758

криминалног понашања појединца, укључујући процене ризике рецидивизма и поновне виктимизације. У другом поглављу, представљене су могућности примене посебних система, софтвера и алата у области сузбијања криминалитета, почевши од запримања пријава о кривичним делима, припреме и планирања криминалистичких истрага, па до њихове примене у области међународне полицијске сарадње. На крају, у чланку се указује на контроверзе у примени вештачке интелигенције од стране агенција за спровођење закона, уз сагледавање више аспеката, као што су: технолошки, правни, етички и друштвени.

Кључне речи: агенције за спровођење закона, превенција криминалитета, сузбијање криминалитета, вештачка интелигенција, контроверзе

* This manuscript was submitted on June 30, 2025, and accepted by the Editorial Board for publishing on November 10, 2025.

UDC 327:699.84-056.24(4-672EU)
DOI: 10.5937/spm94-59771
Original research article

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 121-134

Jadranka Otašević[*]

Faculty for Special Education and Rehabilitation, University of Belgrade

Mira Jovanovska[**]

Center for Rehabilitation of Verbal Communication Pathology, Skopje

ARTIFICIAL INTELLIGENCE AS A SUPPORT FOR INDEPENDENT LIVING OF PEOPLE WITH DISABILITIES

Abstract

Improving the quality of life and independence of people with disabilities is a challenge in the fields of social policy, technology, and law. The paper explores the application of artificial intelligence (AI) as a means of supporting independent living for this population, with a special focus on the possibilities of digital work and entrepreneurship. The theoretical framework of the paper is based on the functional and social model of disability and on the concept of inclusion and human rights. The relevant scientific literature, reports of international organizations (WHO, UNICEF, UN), as well as case studies and pilot projects in the field of digital support were analyzed. The results indicate that AI solutions significantly improve mobility, communication, access to information, and user confidence, but that their application is often limited by the lack of financial accessibility, appropriate policies, and active involvement of people with disabilities themselves in the process of creating AI support. Special attention is dedicated to employment challenges, where digital work is recognized

[*] E-mail: jadrankastevovic@yahoo.com; ORCID: 0000-0001-7051-3897
[**] E-mail: jovanovskamira@gmail.com; ORCID: 0009-0008-2090-6340

as a potential solution, but is underutilized in our society. The need for accessible, ethical, and inclusive AI solutions is emphasized, as well as for the development of digital competencies in people with disabilities in order to achieve greater social and economic integration.

Keywords: people with disabilities, artificial intelligence, quality of life, inclusion, employment

INTRODUCTION

In modern society, improving the quality of life and empowering people with disabilities is one of the key challenges of social policy, technology and medicine. Attitudes towards people with disabilities are not innate, but learned, among other things, through the prejudices and ignorance of others. Changing attitudes is a complex and long-term process that is largely based on experience, but also on information (Otašević i Dobrota-Davidović 2018, 162–163). The concept of independent living implies the ability of people with disabilities to make decisions about their lives, live in the community, use available resources, and have equal opportunities with all other citizens (Shakespeare 2006, 185). According to the International Classification of Functioning, Disability and Health adopted by the World Health Organization, disability is defined as the result of the complex interaction between the health status of the individual and various barriers in the environment that limit his or her full participation in society on an equal basis with others (World Health Organization [WHO] 2007). This functional and social model of disability marked a shift from the traditional, medical model to a more inclusive approach based on human rights.

In the Republic of Serbia, the position of persons with disabilities is regulated by numerous legal acts and strategic documents, such as the Law on Prevention of Discrimination against Persons with Disabilities, the Law on Social Protection, as well as the Strategy for Improving the Position of Persons with Disabilities for the Period from 2025 to 2030 (Zakon o socijalnoj zaštiti 2011; Zakon o sprečavanju diskriminacije osoba sa invaliditetom 2016). These documents emphasize the principles of equal opportunities, accessibility, independent living, and participation in decision-making, but practice often shows that these measures are not fully implemented, especially in terms of the application of modern

digital and technological solutions (Ministry of Labor, Employment, Veterans, and Social Affairs 2025). Analysis of the living situation of people with disabilities shows that there are numerous obstacles in their living space that stand in the way of meeting their needs. Data on the incidence of disabilities are very heterogeneous, due to the use of different criteria for their identification, which further complicates the targeted planning and implementation of support measures in the community (Otašević i Dobrota 2018, 162). In the last decade, artificial intelligence (AI) has been increasingly recognized as a potentially powerful tool in supporting the integration and independence of people with disabilities. From voice assistants and image recognition systems to smart devices and robotic aids, AI-based technologies enable greater independence in everyday functioning (Heffernan and Heffernan 2014; Farhah *et al.* 2025). However, AI applications in practice are not equally accessible, and the user aspect of people with disabilities is often not sufficiently included in the process of developing these solutions. At the same time, in a world of accelerated digitalization and increased demand for digital work, digital entrepreneurship is increasingly being discussed as one of the mechanisms for employing marginalized groups. Digital work involves creating value through interaction with information and communication technologies, including work via digital platforms and self-employment in an online environment. Such forms of work enable flexible working hours and remote work, which is of particular importance for people who have physical barriers, limited mobility, or the need for an adapted, somewhat slower pace (Lazić *et al.* 2024). Nonetheless, in Serbia, people with disabilities continue to face high unemployment rates. The reasons for this include discrimination in the labor market, lack of adequate work capacity assessment, unfavorable socioeconomic background, and insufficient level of professional and digital skills. Although policymakers have launched some initiatives to develop digital competencies among this population, they are mostly limited to basic-level training, which is not sufficient for independent and sustainable entry into the digital labor market. When it comes to how people with disabilities perceive digital self-employment and to what extent they recognize it as a real opportunity to solve the problem of unemployment, a more detailed analysis is needed (Lazić i Vukmirović 2022).

Taking into account all of the above, the aim of the paper is to explore the potential and challenges of applying artificial intelligence as a

support for independent living for people with disabilities, with a special focus on aspects of digital work and inclusion. In order to obtain relevant and reliable information on the application of artificial intelligence (AI) in supporting independent living for people with disabilities, a literature review was conducted, which includes the analysis of scientific papers as well as official reports from international organizations, and examples of good practice from the environment.

ARTIFICIAL INTELLIGENCE – POSSIBILITIES TO SUPPORT PERSONS WITH DISABILITIES

The application of artificial intelligence (AI) in the field of support for people with disabilities has shown significant potential for improving independence, mobility, communication, and overall quality of life. Modern AI solutions are developed in accordance with the type of disability.

Assistive technologies are devices or systems that help people with disabilities overcome physical, sensory, or cognitive barriers. With the integration of artificial intelligence (AI), these technologies become more adaptable and efficient. These include smart wheelchairs and walking devices – AI enables navigation in space with obstacle avoidance, speed adjustment, and automatic braking. Wheelchairs can be controlled with head, eye, or voice movements. Modern research shows that robotic wheelchairs significantly improve the mobility and independence of people with disabilities, providing opportunities for autonomous and semi-autonomous navigation, customized controls, and improved interactions with the environment (Sahoo and Choudhury 2023). Then, smart prostheses and orthopedic devices – modern prostheses that use machine learning to identify the user's movements, which enables finer motor control and more natural movement, as well as voice control and gesture or speech recognition – devices that use AI to recognize speech or body movements and allow the control of electronic devices or the activation of alarms without the need for physical contact.

Communication is essential for independent living, and for many people with disabilities, it is one of the biggest barriers. Assistive technologies play a key role in empowering blind and partially sighted people, as well as people with other disabilities or problems such as dyslexia, by enabling them to be more independent and access information more easily. Screen readers such as JAWS and

VoiceOver, artificial vision devices such as OrCam MyEye, as well as smart canes and applications such as Be My Eyes, which connect blind people with volunteers via video calls, are commonly used for blind and partially sighted people (Hamideh Kerdar, Bachler, and Kirchhoff 2024). People with dyslexia are helped by text-to-speech technologies (Speechify, Voice Dream Reader), as well as intelligent writing tools that offer word prediction and grammar correction (Grammarly, Co: Writer), which make reading and expression easier. Virtual assistants and Chatbots (Siri, Alexa, ChatGPT) are also playing an increasingly important role, enabling voice interaction, dictation, planning, and search, which significantly simplifies everyday activities for users with various difficulties. In addition, technologies for automatic sign language translation are being developed, which, with the help of cameras and artificial intelligence, enable two-way communication between deaf people and the wider population, in real time (O'Sullivan 2019). All these technologies together build the foundations for a more inclusive society, in which everyone has access to knowledge, communication, and independence, regardless of physical or cognitive barriers.

The concept of a "smart home" takes on a new dimension with AI, especially for people with disabilities who need a greater degree of control over their environment. Smart home systems are used to control lighting, temperature, doors, windows, and alarm systems via voice or mobile applications. These systems allow people with limited mobility to manage their environment without physical effort. There are safety and independence sensors, sensors that are connected to AI systems and can track movement, recognize falls, signal danger (fire or gas leak), and automatically alert emergency services. AI can also analyze user behavior and indicate changes in health in a timely manner. With the increase in the use of smartphones and the Internet, a large number of AI solutions are implemented through mobile and web applications that are intended to improve independence in everyday life. Orientation and navigation applications: AI tools like Google Lookout or LazarilloApp allow blind and visually impaired people to move independently, giving them audio instructions, recognizing obstacles and objects in the environment. Health monitoring, where applications using AI analyze physiological data (heart rate, blood sugar level) and provide instructions or alarms in case of deviations. This is important for people with chronic conditions or elderly people with disabilities. There is also the organization of daily activities; AI applications can help

with reminders of medication obligations or for communication that has been agreed in advance.

Artificial intelligence (AI) based solutions have the potential to significantly improve the lives of people with disabilities, not only by supporting them in performing daily activities, but also by enabling them to acquire new abilities and skills. The use of AI opens up new paths for people with disabilities to independence, accessibility, and active participation in society, overcoming numerous physical, sensory, and communication barriers that would otherwise pose serious challenges or be completely insurmountable. In this way, technology not only compensates for limitations but also contributes to the empowerment of people with disabilities and their full inclusion in all aspects of social life – from education and employment to social and cultural activities (Kumar *et al.* 2024). Moreover, the development of AI in this area also represents a social value because it supports the transition from a model of medical and social dependence to a model of digital independence and active participation in the social community. AI thus becomes not only a technological tool, but also an ethical and inclusive tool for building a more just society. In order to fully realize this potential, it is necessary to continue developing inclusive technological solutions in cooperation with the users themselves – people with disabilities – who must be active participants in the design and testing process. Only in this way will AI technologies be able to respond to real needs, respect diversity, and contribute to building an environment in which accessibility and independence become a reality for all.

THE IMPACT OF ARTIFICIAL INTELLIGENCE ON INDEPENDENCE AND QUALITY OF LIFE

Artificial Intelligence (AI) is a revolutionary tool in improving the lives of people with disabilities, enabling greater independence, better mobility, easier access to information, and increased self-confidence. Although AI is still in its early stages, numerous examples from practice and research confirm that its application can transform everyday life and open up new opportunities for inclusion.

Around 16% of the world's population, or 1.3 billion people, currently live with some form of disability, including 240 million children. In the European Union, 27% of people aged 16 and over had some form of disability in 2023, with a higher percentage of women

than men with disabilities in all member states. These figures highlight the need to systematically improve access to a better quality of life for a large and diverse population (Council of the European Union 2023). The application of AI in assistive technologies, such as smart wheelchairs, screen readers, voice control devices, or orientation applications, significantly increases mobility and access to information. This, in turn, has a positive impact on users' self-confidence, as it allows them to make decisions independently, manage their environment, and communicate without intermediaries. For example, robotic wheelchairs that use machine learning to avoid obstacles and adapt their speed (Sahho and Choudxury 2023) contribute to greater independence for users, while applications such as Be My Eyes and LazarilloApp facilitate orientation and social interaction for people with visual impairments. At the same time, it should be emphasized that this potential is still unrealized for many. According to WHO and UNICEF in 2022, more than 2.5 billion people worldwide need one or more assistive products – such as wheelchairs, hearing aids, or communication applications. However, 1 billion people do not have access to these devices, especially in low- and middle-income countries, which indicates large global inequalities in the availability of AI technologies (Federal Ministry for Economic Cooperation and Development and International Disability Alliance [IDA], 2025).

User experience research shows that AI solutions have a positive impact on quality of life when designed according to the real needs of the user. Users of devices such as OrCam MyEyes highlight convenience and improved independence, but indicate the need for better localization and native language support (Amore *et al.* 2023). Virtual assistants, Chatbots, and text-to-speech systems (e.g., Voice Dream Reader, ChatGPT, Siri) also receive positive reviews in the domain of communication and cognitive support.

LIMITATIONS AND ETHICAL ISSUES

Artificial intelligence (AI) is increasingly shaping various aspects of society, including the area of support for people with disabilities. Although the potential of AI in this context seems promising, numerous limitations and ethical dilemmas that accompany its application should not be ignored. One of the biggest limitations is the high cost of modern AI systems and technological solutions, which often exceeds the financial

capabilities of people with disabilities. Devices such as smart wheelchairs, vision-based communication systems, or personalized voice assistants require significant investments. Although there are certain support actions in Serbia, such as subsidies for technical aids, systemic shortcomings and limited resources mean that many solutions remain inaccessible to the most vulnerable. It should be emphasized that the collection, processing, and storage of data in AI systems pose a serious ethical and legal challenge. People with disabilities, who often use systems for continuous health monitoring or digital communication, are at risk of privacy violations. Unauthorized access, commercialization of data, or unclear terms of use can threaten the fundamental rights of users. Therefore, the implementation of strategies such as data anonymization and strengthening security protocols is a priority (Cowls *et al.* 2019, 4, 8).

In accordance with the ethical imperative of harm prevention, it is essential that the state maintains regulatory oversight of the market for highly automated, AI-based robotic systems intended to provide assistance to persons with disabilities. Owing to their inherent characteristics, such systems typically involve direct and continuous interaction with end users (operators). Within this context, so-called collaborative robots present a distinct set of ethical and legal challenges, as their application may unintentionally result in harm to life, bodily integrity, health, or the fundamental rights of users, as well as the risk of health deterioration. Therefore, the design, development, and deployment of these systems must be guided by rigorous compliance with international safety standards specifically established for human-robot interaction (Regulation EU 2024/1689).

AI systems are based on algorithms that learn from data. If this data is biased – which is often the case – the system itself can make discriminatory decisions, which particularly affects marginalized (vulnerable) groups. For example, an automated hiring system can reject candidates with disabilities based on previously "learned" patterns of discrimination. Developing methods to improve the transparency of algorithmic decisions, as well as monitoring bias, is essential for building fairer systems (Binns 2018). Reliance on AI systems can lead to problems in cases of technical failure, loss of support, or inadequate updates. This dependence is particularly risky for people with severe disabilities who do not have alternative mechanisms for communication or movement. Therefore, it is necessary to develop parallel, low-tech, or hybrid options as backup strategies in real-life situations.

Despite the good intentions of technology creators, many AI systems are still being developed without the active participation of people with disabilities. This means that solutions often do not respond to real needs and life situations. The principle of "nothing about us without us" must guide all phases of technology development (Wolbring and Nguyen 2023).

AI is significantly transforming the labor market, often reducing the need for human labor and increasing the risk of unemployment among less-advantaged groups. People with disabilities, who already have lower employment rates, may be further disadvantaged by these trends. However, new forms of digital work, such as digital entrepreneurship and flexible platform work, present the potential for greater inclusion, but only if the accompanying training is adequate and targeted to the needs of the users.

Current regulatory systems are largely lagging behind the dynamics of technological development. The lack of clear regulations, accountability mechanisms, and transparency standards means that many AI systems operate in legal gray areas. Improving legal frameworks at the national and international levels is necessary, along with strengthening ethical guidelines and cooperation between different sectors. Addressing all of these challenges requires integrating ethics from the beginning of system design, improving transparency of algorithms, creating global ethical standards, and active collaboration between engineers, ethicists, decision-makers, and end-users. Only such an approach can enable AI to become a tool for a sustainable and inclusive digital future for all.

CONCLUSION

In the modern era, where digital technology is increasingly integrated into everyday life, the application of artificial intelligence (AI) is a key resource in improving the quality of life of people with disabilities. Analysis shows that AI has the potential to significantly alleviate or completely overcome numerous physical, sensory, and communication barriers that this population faces on a daily basis. With the help of assistive devices, smart applications, robotic aids, and voice and visual interaction systems, people with disabilities are given the opportunity for a greater degree of autonomy, independent decision-making, as well as easier access to education, work, and social life.

It is particularly important to highlight the role of AI in the field of digital work and self-employment, as these forms of activity open up new opportunities for economic empowerment and social inclusion. However, data indicate that the level of utilization of these opportunities is still low, both due to limited access to technology, as well as insufficient support for the development of digital skills and structural barriers such as discrimination and the lack of adapted programs. Therefore, it is necessary that technological development is accompanied by systemic empowerment – through comprehensive policies, the inclusion of persons with disabilities in design and decision-making processes, as well as continued investment in education and digital inclusion. Artificial intelligence must not remain just a technological innovation trend available to a privileged few, but must become a tool for building equal opportunities for all. Its development must be directed towards social justice, ethical responsibility, and respect for human rights. Only an inclusive and adapted approach can lead to AI becoming a substrate for the transition from a model of social dependence to a model of digital independence. In this context, the role of policymakers, technological institutions, the academic community, and people with disabilities themselves is crucial. By joining forces, it is possible to create an environment in which technology will be not only a tool, but also an example of equality, inclusion, and a dignified life.

REFFERENCES

Amore, Filippo, Valeria Silvestri, Margherita Guidobaldi, Marco Sulfaro, Paola Piscopo, Simona Turco, Francesca De Rossi [Amore *et al.*]. 2023. "Efficacy and patients' satisfaction with the ORCAM MyEye device among visually impaired people: a multicenter study." *Journal of Medical Systems* 47 (1): 11.

Binns, Reuben. 2018. "Algorithmic accountability and public reason." *Philosophy & technology* 31 (4): 543–556.

Council of the European Union. 2023. *Disability in EU: Facts and Figures – Consilium*. https://www.consilium.europa.eu/en/infographics/disability-eu-facts-figures

Cowls, Josh, King Thomas, Taddeo Mariarosaria, and Floridi Luciano [Cowls *et al.*]. 2019. *Designing AI for social good: Seven essential factors*. SSRN.

Farhah, Nesren S., Asim Wadood, Ahmed Abdullah Alqarni, M. Irfan Uddin, and Theyazn HH Aldhyani [Farhah *et al.*]. 2025. "Enhancing Adaptive Learning with Generative AI for Tailored Educational Support for Students with Disabilities." *Journal of Disability Research* 4 (3): 1–17.

Federal Ministry for Economic Cooperation and Development, and International Disability Alliance [IDA]. 2025. *Global Disability Inclusion Report*. Federal Ministry for Economic Cooperation and Development, and International Disability Alliance.

Hamideh Kerdar, Sara, Liane Bächler, and Britta Marleen Kirchhoff. 2024. "The accessibility of digital technologies for people with visual impairment and blindness: a scoping review." *Discover Computing* 27 (1): 1–15.

Heffernan, Neil T., and Cristina Lindquist Heffernan. 2014. "The ASSISTments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching." *International Journal of Artificial Intelligence in Education* 24 (4): 470–497.

Kumar, Vishal, Sitanshu Barik, Sameer Aggarwal, Deepak Kumar, and Vikash Raj [Kumar *et al.*]. 2024. "The use of artificial intelligence for persons with disability: a bright and promising future ahead." *Disability and Rehabilitation: Assistive Technology* 19 (6): 2415–2417.

Lazić, Milena, and Valentina Vukmirović. 2022. "Improving digital competences of persons with disabilities as a precondition for an inclusive digital economy: Evidence from Serbia."*PaKSoM* : 389–394.

Lazić, Milena, Valentina Vukmirović and Ivana Domazet [Lazić *et al.*]. 2024. "Digital marketing as a tool for inclusive employment. *SeMA Journal*: 119–126.

Ministarstvo za rad, zapošljavanje, boračka i socijalna pitanja [Ministry of Labor, Employment, Veterans and Social Affairs]. 2025. *Strategija unapređenja položaja osoba sa invaliditetom u Republici Srbiji za period 2025–2030. godine*. „Službeni glasnik RS", br. 6.

O'Sullivan, John F. 2019. *The UDL Educational Technology Guide 2020: Technology for Special Education*. John F. O'Sullivan Jr.

Otašević, Jadranka, and Nada Dobrota-Davidović. 2018. „Institucionalni kapaciteti Srbije U rešavanju problema dece sa smetnjama u razvoju." *Srpska politička misao* 61 (3): 157–173.

Regulation EU 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU)

No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), *OJ* L*,* 2024/1689, 12.7.2024.

Sahoo, Sushil Kumar, and Bibhuti Bhusan Choudhury. 2023. "A review on smart robotic wheelchairs with advancing mobility and independence for individuals with disabilities." *Journal of Decision Analytics and Intelligent Computing* 3 (1): 221–242.

Shakespeare, Tom. 2006. *Disability rights and wrongs*. London: Routledge.

Wolbring, Gregor, and Annie Nguyen. 2023. "Equity/equality, diversity and inclusion, and other EDI phrases and EDI policy frameworks: A scoping review." *Trends in Higher Education* 2 (1): 168–237.

World Health Organization [WHO]. 2007. *International Classification of Functioning, Disability, and Health: Children & Youth Version: ICF-CY*. Geneva, Switzerland: World Health Organization.

Zakon o socijalnoj zaštiti, „Sl. glasnik RS", br.24/2011 i 117/022-odluka US.

Zakon o sprečavanju diskriminacije osoba sa invaliditetom („Sl. glasnik RS", br. 33/2006 i 13/2016).

Јадранка Оташевић[*]

Факултет за специјалну едукацију и рехабилитацију, Универзитет у Београду

Мира Јовановска[**]

Центар за рехабилитација на патологија на вербалната комуникација, Скопје

ВЕШТАЧКА ИНТЕЛИГЕНЦИЈА КАО ПОДРШКА САМОСТАЛНОМ ЖИВОТУ ОСОБА СА ИНВАЛИДИТЕТОМ

Резиме

Овај рад истражује потенцијале и изазове примене вештачке интелигенције (АИ) у подршци самосталном животу особама са инвалидитетом, са посебним освртом на дигитални рад и инклузију. Анализом литературе и доступних извештаја утврђено је да АИ технологије, укључујући асистивне уређаје, паметна колица, читаче екрана, уређаје за орјентацију и комуникационе апликације, значајно унапређују независност, мобилност, приступ информацијама и квалитет живота корисника. Уочено је да су, упркос напретку, многи АИ системи и даље недовољно доступни, посебно у земљама са ниским и средњим приходима, што ствара глобалне неједнакости. Истраживања корисничког искуства показују да правилно дизајниране АИ технологије повећавају самопоуздање и олакшавају свакодневне активности, али захтевају бољу подршку на матерњем језику. Рад такође истиче етичке и регулаторне изазове, укључујући високе трошкове, ризик од дискриминације због пристрасних алгоритама, зависност од технологије и питање заштите личних података. У закључку се наводи да инклузивни приступ у развоју АИ, са активним учешћем особа са инвалидитетом, представља кључ за изградњу праведног и доступног технолошког окружења. Имплементација АИ решења може допринети преласку са модела медицинске или социјалне

[*] Имејл: jadrankastevovic@yahoo.com; ORCID: 0000-0001-7051-3897
[**] Имејл: jovanovskamira@gmail.com; ORCID: 0009-0008-2090-6340

зависности на дигиталну независност и допринети активном учешћу у друштву, али само уз етичку контролу, континуирани развој и сарадњу релевантних актера.

Кључне речи: особе са инвалидитетом, вештачка интелигенција, квалитет живота, инклузија, запошљавање

* This manuscript was submitted on June 26, 2025, and accepted by the Editorial Board for publishing on November 10, 2025.

UDC 327:316.776:004.774(100)
DOI: 10.5937/spm94-59969
Original research article

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 137-165

Ljubiša Bojić*

Institute for Artificial Intelligence Research and Development of Serbia, Novi Sad

Institute for Philosophy and Social Theory, University of Belgrade

Digital Society Lab, Belgrade

Complexity Science Hub, Vienna, Austria

RISK AND RESPONSIBILITY AT THE FRONTIER OF AI: A THEMATIC ANALYSIS OF DEEP LEARNING PIONEERS' PERSPECTIVES ON ARTIFICIAL INTELLIGENCE THREATS AND GOVERNANCE**

Abstract

As artificial intelligence (AI) reshapes global societies, understanding its associated risks and governance imperatives is of urgent social importance. This study fills a critical gap by systematically analyzing extended interviews with Geoffrey Hinton, Yoshua Bengio, and Yann LeCun – to elucidate their firsthand perspectives on AI's existential, ethical, social, and governance challenges. Employing qualitative thematic analysis across six longitudinal interview transcripts, the research identifies both convergences and divergences: Hinton and

* E-mail: ljubisa.bojic@ivi.ac.rs; ORCID: 0000-0002-5371-7975

** This study was realised with the support of the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, according to the Agreement on the realisation and financing of scientific research 451-03-136/2025-03/ 200025 and the Agreement on the realisation and financing of scientific research 451-03-136/2025-03.

Bengio strongly emphasize existential threats, superintelligence hazards, AI weapons risks, and the need for robust global regulation, while LeCun expresses technological optimism and favors decentralized, open development. All acknowledge economic disruption, misuse of potential, and fractures in democratic discourse. The study's findings reveal that expert opinion on AI risk is far from monolithic and highlight actionable, innovative governance proposals, from regulated compute access to "diversity engines" in social media feeds. Implications include the necessity for adaptive, internationally coordinated AI governance and greater professional accountability among developers. Limitations include a focus on elite, Anglophone experts and inherent subjectivity in qualitative coding. Future research should expand to multi-stakeholder and cross-national perspectives, and test proposed regulatory frameworks in real-world contexts, addressing the ongoing evolution of risk as AI permeates new domains.

Keywords: artificial intelligence risk, deep learning pioneers, AI governance, existential threats, thematic analysis

INTRODUCTION

The ascent of artificial intelligence (AI) as a transformative technology marks a pivotal moment in the trajectory of human civilization. Over the past decade, AI has achieved remarkable milestones, transitioning from research laboratories into core infrastructure for contemporary society (LeCun, Bengio, and Hinton 2015). Deep neural networks, inspired in part by biological processes in the human brain, now underpin applications as varied as computer vision, internet search, speech recognition, machine translation, drug discovery, logistics, robotics, and automated financial trading. These systems have set new technical benchmarks, achieving and even exceeding human-level performance in domains such as image classification (Krizhevsky, Sutskever, and Hinton 2017), the board game Go (Silver *et al.* 2016), and large-scale language processing (OpenAI 2023).

The widespread deployment of AI creates unprecedented possibilities for social benefit, economic growth, and scientific discovery. For instance, AI diagnostic tools are increasingly used to interpret medical images, improving access to healthcare and enabling earlier intervention

(Esteva *et al.* 2017). In resource management, machine learning models optimize supply chains, reducing waste, and boosting sustainability. There is optimism among some scholars and policymakers that if guided responsibly, AI could make significant strides toward addressing global challenges such as poverty, education inequality, and climate change (Chui, Manyika, and Miremadi 2016). Indeed, the scale and versatility of AI's impact have prompted many to equate its significance with previous general-purpose technologies such as electricity or the internet (Brynjolfsson and McAfee 2014).

The accelerating pace and scale of AI adoption have also raised a chorus of urgent questions about risk, ethics, and social governance. Unlike technologies of the past, advanced AI systems carry the potential not merely to augment human activity, but to fundamentally reshape economic structures, modes of communication, and the boundaries of autonomy and agency in human life (Bostrom 2014; Russell 2019; Bodroža, Dinić, and Bojić 2024; Bojić *et al.* 2024; Bojić, Kovačević, and Čabarkapa 2025; Bojić, Stojković, and Jolić Marjanović 2024; Bojić *et al*. 2025). As AI systems increasingly mediate social decisions about credit, employment, policing, and information dissemination, concerns regarding transparency, fairness, and accountability have grown exponentially (Doshi-Velez and Kim 2017; O'Neil 2016; Noble 2018).

A significant body of literature explores the specific societal risks posed by contemporary AI. Algorithmic bias and discrimination have come under close scrutiny as systems trained on large, often uncurated datasets are shown to reproduce and reinforce patterns of racial, gender, and socioeconomic inequality (Buolamwini and Gebru 2018; Noble 2018; Sandvig *et al.* 2016). In health care, these biases can manifest in life-critical contexts, amplifying existing disparities and undermining trust (Obermeyer *et al.* 2019; Reinhardt *et al*. 2025).

Other research has exposed the role of "black-box" neural models in furthering opacity and frustrating efforts toward meaningful interpretability or recourse for affected individuals (Doshi-Velez and Kim 2017; Rudin 2019).

The economic implications of AI-driven automation have prompted intense scholarly and policy debate. While some analyses forecast AI as a complement to human labor, enhancing productivity and creating new categories of employment (Bessen 2018), others warn of large-scale displacement, particularly for middle-skill and routine occupations (Brynjolfsson and McAfee 2014; Korinek and Stiglitz

2018). Beyond the loss of income, such displacement may threaten social cohesion, increase inequality, and diminish the social status and personal dignity associated with work (Susskind 2020; Autor 2015). The distribution of economic gains from AI also risks concentrating power in a handful of corporations and states, aggravating global disparities (Bojic 2022; Bojic 2024).

In the political domain, machine learning techniques have been deployed not only for benign purposes such as translating languages or moderating content, but also to manipulate public opinion through targeted advertising, deepfakes, and recommendation algorithms (Pavlovic and Bojic 2020; O'Connor and Weatherall 2019). The proliferation of generative models has blurred the distinction between authentic and synthetic content, complicating journalistic verification and undermining democratic deliberation (West 2019; Zuboff 2019). Repressive regimes have exploited AI for surveillance and social control, raising profound questions about privacy, civil liberties, and the resilience of liberal democratic institutions (Feldstein 2019).

Existential risks associated with advanced AI systems have also received increased attention in recent years. Bostrom (2014) and others highlight the potential for a hypothetical "superintelligence" – an AI system surpassing human cognitive abilities – to act in ways detrimental or even catastrophic to humanity if it becomes misaligned with human interests. Even in the absence of the arrival of superintelligence, warnings have been issued concerning accidental or malicious misuse of powerful AI technologies, for example, in cyberattacks, the development of lethal autonomous weapons, or the creation of destabilizing misinformation (Brundage *et al.* 2018). As AI capabilities continue to expand, the challenges of governing such systems – many of which are developed and deployed transnationally by commercial actors – are only becoming more complex (Floridi *et al.* 2018).

Despite this rich and expanding literature, a critical gap remains. Much of the existing discourse, while rigorous, is written by ethicists, social scientists, journalists, and technology policy experts. Far less research has systematically engaged with the firsthand perspectives of those most instrumental in developing and propagating the technologies at the heart of the ongoing AI revolution. Indeed, as AI's foundational technologies mature, the reflections, warnings, and priorities of its principal architects are increasingly recognized as vital contributions to the societal dialogue (Metz 2023; Turing 2018).

Among these architects, Geoffrey Hinton, Yoshua Bengio, and Yann LeCun stand apart as the preeminent pioneers of deep learning. Their research, begun during an era when neural networks were marginalized within the artificial intelligence community, has shaped the breakthroughs that define current AI progress. Collectively, their work on backpropagation, convolutional neural networks, probabilistic modeling, natural language processing, and generative adversarial learning forms the backbone of nearly all state-of-the-art systems in commercial and scientific use (LeCun, Bengio, and Hinton 2015; Schmidhuber 2015). In recognition of their transformative contributions, all three were jointly awarded the 2018 ACM A.M. Turing Award, the highest honor in computer science (Turing 2018).

The influence of Hinton, Bengio, and LeCun extends far beyond the laboratory. Each has served as an advisor to leading governments and international bodies on AI policy, headed major industrial AI research laboratories (including at Google and Meta/Facebook), and shaped entire research ecosystems through the foundation of academic institutes and the training of hundreds of graduate students (LeCun, Bengio, and Hinton 2015). Crucially, all three have become increasingly transparent and vocal about their own shifting views on AI's risks and social obligations, with Hinton – sometimes described as the "Godfather of AI" – resigning from Google in 2023 to speak more freely about the potential dangers of the technologies he helped create (Metz 2023).

While the technical achievements of these pioneers have been thoroughly chronicled, relatively little scholarly attention has been paid to how they themselves diagnose, prioritize, and envision the mitigation of risks associated with AI. Publicly available, long-form interviews and panel discussions provide a unique window into their evolving thought processes, concerns, and responses to the accelerating diffusion of AI. Such qualitative material allows researchers to explore explicit warnings – such as those regarding existential threats or social instability – and the underlying values, uncertainties, and tensions voiced by these scientists as they grapple with the legacy and future direction of their field (Russell 2019; Floridi *et al.* 2018; ACM 2018).

This paper aims to address the aforementioned gap by conducting a systematic qualitative analysis of extended interviews with Geoffrey Hinton, Yoshua Bengio, and Yann LeCun. Drawing on rich, longitudinal data collected from multiple interviews per participant, this study interrogates their perspectives within the wider context of the scholarly

and public discourse on AI risk. Specifically, the analysis is guided by the following research questions:

- First, how do Hinton, Bengio, and LeCun conceptualize and prioritize the most significant threats arising from advances in artificial intelligence?
- Second, what themes of consensus or divergence emerge in their analyses of technical, economic, ethical, and existential risks, and to what extent are these aligned (or not) with dominant narratives in the academic literature?
- Third, how have their views evolved in response to technological breakthroughs such as large language models, as well as shifting political and societal conditions?
- Finally, what implications do their positions have for the governance and responsible stewardship of AI technologies going forward?

This research endeavors to deepen our understanding of AI's risks and the complex interplay between scientific progress, individual responsibility, and social foresight. In so doing, the paper seeks to inform AI policy, research, and public dialogue at a time when the stakes for humanity have seldom been higher.

METHODOLOGY

This study employed a qualitative research design, specifically utilizing thematic analysis, to explore and synthesize the perspectives of three foundational figures in modern artificial intelligence – Geoffrey Hinton, Yann LeCun, and Yoshua Bengio – on the risks and threats posed by recent advances in AI. Each of these individuals was selected for analysis owing both to their status as recipients of the 2018 ACM A.M. Turing Award, often considered the "Nobel Prize of Computing," and for their direct, longstanding roles in conceptualizing and implementing deep neural network models that now underpin a wide range of contemporary AI systems (ACM 2018; Turing_Bengio 2018; Turing_Hinton 2018; Turing_LeCun 2018). The present study sought to document, compare, and interpret their views as expressed in extended public interviews, thereby capturing insights likely to shape the evolving discourse on the societal impact of AI.

The empirical material comprised six in-depth interview transcripts (OSF 2025; Hinton_Diary 2025; Hinton_60Minutes 2023; Bengio_BBC 2025; Bengio_WSF 2024; LeCun_Brian 2024; LeCun_Lex 2024).

For each of the three Turing laureates, two interviews were selected: one interview conducted within the previous six months, and a second conducted no less than one year prior. This sampling framework was intended to both provide a breadth of expression and facilitate longitudinal comparison, enabling the detection of any shifts in perspective as the technological and ethical landscape of AI has rapidly evolved. Interviews were sourced from widely recognized, publicly accessible podcasts and YouTube discussion channels. The selection criteria prioritized interviews in which each guest was engaged in an open, discursive exchange, and where the conversation explored, either directly or tangentially, current and anticipated risks associated with artificial general intelligence, deep learning, and their applications. Because the personalities involved are public figures and the interviews were intended for public dissemination, all identifying and contextualizing information was preserved in the transcripts. Where interview transcripts were not already available, the audio or video content was precisely transcribed by the researcher to ensure fidelity to the speakers' statements, including pauses, qualifiers, and nonverbal cues that contributed meaning or emphasis. The final textual corpus underwent careful review, with corrections made for errors in transcription and to clarify ambiguous utterances where necessary.

Thematic analysis was chosen as the principal analytic strategy. This approach, as articulated by Braun and Clarke (2006), provides a systematic yet flexible framework for identifying, organizing, and interpreting salient patterns within qualitative data. Unlike quantitative content analysis, thematic analysis does not reduce data to counts or simple categories; instead, it emphasizes depth of meaning, allowing the researcher to surface both explicit statements and more latent underlying themes embedded in the discourse. Analysis began with immersion in the data: the researcher read and reread the entire set of transcripts, making initial margin notes concerning recurring topics, metaphors, and expressions of concern or optimism. This stage established a foundation for subsequent coding, fostering a holistic sense of each interview as well as the data set as a whole.

During the initial coding phase, both deductive and inductive elements were employed. The starting codebook was informed by leading scholarship on risks associated with AI and machine learning – such as existential threats, algorithmic bias, misuse in authoritarian or military domains, mass displacement of labor, and ethical dilemmas in autonomous decision-making (Braun and Clarke 2006; ACM 2018). At

the same time, the researcher allowed for the emergence of unexpected or novel topics through inductive coding, remaining attentive to unique turns of phrase, emotionally charged language, or instances where the interview subject expressed personal ambivalence, regret, or optimism regarding their own legacy. Coding was conducted manually in successive passes, and the codebook was treated as a living document, subject to refinement as additional patterns or contradictions became visible across the interviews. When ambiguous passages appeared, the researcher documented interpretive choices in analytic memos to sustain transparency and reflexivity.

Following coding, the codes were examined for conceptual affinity and organized into preliminary themes. These themes ranged from clearly delineated technical risks – such as concerns about autonomous weapons, adversarial attacks, or uncontrolled AI self-improvement – to broader social and philosophical worries, including implications for economic inequality, mass unemployment, erosion of democratic processes, and fundamental questions about consciousness, sentience, and value alignment in artificial agents. Themes were further scrutinized and revised through an iterative, dialogic process, with attention to both convergence and divergence in the informants' perspectives. For example, the degree of emphasis placed on existential risk versus more immediate hazards, or varying levels of confidence in regulatory and governance solutions, was noted and integrated as subthemes or points of contrast. Throughout this process, the researcher sought to remain faithful to the language and logic of the participants themselves and to interpret their statements in relation to broader scholarly and policy debates.

Analysis and interpretation were supported by ongoing self-reflection and the maintenance of an audit trail, which documented all key decisions and adjustments to methodological procedures. Despite the single-author nature of the analysis, rigor was maintained through repeated returns to the data, the careful selection of illustrative quotations, and conscious monitoring for interpretive bias. When representative extracts were selected to exemplify particular themes or to demonstrate moments of disagreement among the Turing laureates, an effort was made to preserve context and the unique rhetorical style of each speaker. In reporting results, the distinction between description and interpretation was actively maintained.

Ethical considerations for this research were minimal, since all data originated in the public domain and involved individuals speaking

in their professional and public-citizen capacity. No sensitive or private information was included, and no interaction occurred between the researcher and the interview subjects. As such, the study is exempt from review by an institutional research ethics board, though all guidelines for responsible scholarship and data integrity were followed.

RESULTS

Thematic analysis of six long-form interviews – two each with Geoffrey Hinton, Yoshua Bengio, and Yann LeCun – yielded 48 *a priori* codes clustered in seven super-themes (A–G). All 48 codes were observed at least once across the corpus; none of the passages required a new category, confirming the sufficiency of the codebook. Below, we synthesise points of convergence and divergence, followed by a comparative table that condenses how strongly each laureate emphasised every super-theme (Table 1).

Hinton and Bengio both foregrounded the prospect of super-intelligence and openly used the language of "existential threat." Hinton's interviews contained 28 distinct passages within cluster A, repeatedly quantifying a "10–20 percent" probability of human extinction (00:54; 08:54), while Bengio invoked extinction or species "replacement" ten times, likening present research to "playing apprentice sorcerer" (26:41). Both described timelines of roughly 5–20 years.

LeCun, by contrast, acknowledged eventual machines "smarter than us" but dismissed catastrophe scenarios as products of a "doomer" mind set. His two interviews yielded only five passages in cluster A, none of which endorsed extinction risk. Instead, he argued that domination requires "hard-wired drives" that engineers can simply omit (131:30), framing alignment as an engineering optimisation task rather than an existential unknown.

Job displacement, loss of dignity, and widening inequality were most salient for Hinton, who provided concrete anecdotes (e.g., a niece whose task time fell from 25 to 5 minutes, 41:05) and offered practical career advice ("Train to be a plumber," 88:01). Bengio's 2024 appearance focused more on catastrophic misuse than labour, but his 2025 BBC interview acknowledged disappearing entry-level legal work. LeCun, although conceding that "natural fear" exists (143:40), framed labour impacts as manageable and emphasised the upside of an "amplifier of human intelligence."

Hinton and Bengio expressed acute concern over cyber-attacks, bio-weapons, and autonomous weapons. Hinton cited a "12,200 percent" year-on-year spike in cyber incidents (12:18) and described combinatorial threat spirals (28:32). Bengio detailed a six-month timeline for engineering lethal viruses (33:54). LeCun, by contrast, dismissed large language models as insufficient for building bio-weapons and treated hostile persuasion scenarios as a future arms race between competing AI assistants.

All three laureates recognised algorithmic manipulation, but with different emphases. Hinton provided the longest discussion of echo-chambers and loss of shared reality, blaming profit-driven recommender systems. Bengio linked manipulation to democratic back sliding. LeCun warned primarily against the concentration of "information diet" in the hands of a few proprietary models, advocating open-source diversity rather than heavy regulation (Bojic, Agatonović, and Guga 2024; Bojic and Marie 2017).

Hinton and Bengio converged on the need for strong global governance – Hinton even positing a "world government that works" (11:24) and Bengio calling for binding international treaties (39:41). Both criticised military carve-outs and corporate profit motives. LeCun worried instead about over-regulation stifling open research; his remedy was decentralisation through open-source foundations. Thus, all three accept the premise of governance gaps, yet propose sharply different solutions.

Each scientist acknowledged transformative upside – especially in health, education, and scientific discovery – but the emotional valence differed. Hinton spoke of "magnificent" benefits (09:55) yet confessed that such promise "takes the edge off" his pride (32:05). Bengio voiced a shift from "excited" to "worried," balancing cures for disease against existential downside. LeCun remained emphatically optimistic, likening AI to the printing press and the Enlightenment (160:42) and predicting "a bright future for humanity if we do this right" (57:55).

Hinton and Bengio articulated personal anxiety, duty, and even regret. Hinton admitted difficulty "coming to terms emotionally" (52:43) and lamented time lost with family due to career obsession (82:39). Bengio expressed responsibility "to talk to governments and citizens" (51:57). LeCun displayed no comparable distress; instead, he articulated confidence that "doom" narratives underestimate both engineering capability and intrinsic human goodness.

Comparing the earlier and later interviews for each laureate reveals drift primarily in Bengio's timeline estimates (from 5–20 years in 2024 to 2–10 years in 2025) and Hinton's escalating emphasis on regulation after leaving Google. LeCun's stance remained comparatively stable between March and December 2024, although the later interview further down played existential concerns.

All three Turing laureates acknowledge both promise and peril, but their emphases diverge sharply. Hinton and Bengio converge on stark existential warnings and the urgency of global regulation, whereas LeCun positions himself as a technological optimist concerned primarily with ensuring open, decentralised development. These differences show that even within the vanguard of deep learning expertise, assessments of AI risk are far from uniform – an insight critical for policymakers who might otherwise treat "expert opinion" as monolithic.

Table 1. Salience of Super-Themes across Laureates. Coding density was graded qualitatively: "High" = >15 supporting excerpts across both interviews; "Moderate" = 5–15 excerpts; "Low" = <5 excerpts or largely dismissive treatment.

Super-Theme (A–G)	Key Risk / Topic	Hinton (2023-2025)	Bengio (2024-2025)	LeCun (2024)
A. Existential & long-term	Super-intelligence, extinction, alignment	High	High	Moderate-Low
B. Societal (jobs, inequality)	Displacement, dignity, UBI	High	Moderate	Moderate-Low
C. Misuse & security	Cyber, bio-weapons, lethal robots	High	High	Low
D. Social-psychological	Polarisation, misinformation, bias	High	Moderate	Moderate
E. Governance & regulation	Global treaties, profit motives, capture	High	High	Moderate
F. Positive potentials / ambivalence	Productivity, health, dual-use	Moderate	Moderate	High
G. Emotional & personal	Duty to warn, regret, personal impact	High	High	Low

Source: Author.

DISCUSSION

The constellation of risks articulated by Geoffrey Hinton, Yoshua Bengio, and Yann LeCun – often called the "godfathers of deep learning" – highlights the need for a new era of AI governance that is as innovative and dynamic as the technology itself (Table 2). While there is consensus among these leaders that AI's potential benefits are enormous, the analysis also reveals genuine apprehension about existential threats, misuse, social disruption, and the inadequacies of existing regulatory approaches. Here, we interpret the findings with a view toward transformative regulatory interventions that would empower societies to guide AI trajectory for the common good.

EXISTENTIAL AND LONG-TERM RISKS

The most fundamental worries – the possibility of AI surpassing human intelligence, acting independently, and potentially making humankind obsolete – cannot be adequately addressed by traditional sector-specific or national regulations. This is not simply because the technology is powerful, but because it can scale at unprecedented speed across borders, within private infrastructures that are often opaque even to governments (Brundage *et al.* 2018). When Geoffrey Hinton asserts, "I realise that these things will one day get smarter than us and we've never had to deal with that," he describes a singularity that upends the very logic of retrospective regulation (Hinton_Diary 2025).

To control the development of AI systems that approach superhuman capacities, we need to regulate the very infrastructure – massive computation ("compute") – that undergirds AI research. No company or research consortium could train an advanced AI system above a certain scale without first obtaining a digital "compute passport." This would function analogously to a visa: to run major training jobs or build very large neural networks, organizations would have to submit their plans, security measures, and alignment protocols to a secure registry, possibly managed by a new international body analogous to the International Atomic Energy Agency (Floridi *et al.* 2018; Brundage *et al.* 2018). The request would be logged cryptographically; only after approval would compute resources (from cloud providers or AI chip manufacturers) be released. This approach would act upstream,

preventing unaccountable actors from conducting "AI moonshots" out of public sight.

Because AI capabilities evolve quickly, one-off approvals are not enough. Instead, licenses for "frontier" AI projects should be renewed every six or twelve months, each time requiring new independent audits of alignment strategies, interpretability, and security (Gabriel 2020). To align private incentives, companies would be made to post "catastrophe bonds" – essentially insurance policies that pay out automatically to an international compensation fund in the event of proven catastrophic harm, such as widespread loss of control or the use of AI in major cyber-attacks. The bond premium paid by the company would scale with the assessed risk, sending a continuous market signal about the evolving dangers of different projects (Korinek and Stiglitz 2018).

ALIGNMENT AND CONTROL

One recurring motif in both Hinton's and Bengio's statements is the need to take responsibility at the top: "We need to figure out how to make [AI systems] not want to take over," Hinton says, but in practice, that means developers and corporate leaders must be legally accountable for their designs (Hinton_Diary 2025). One creative solution is to impose a fiduciary duty – an explicit and legally binding obligation – on company executives overseeing AI systems above a certain capability. This would oblige them to prioritize public safety and value alignment even above profits, mirroring duties that already exist for trust managers, medical boards, or environmental stewards (Cath 2018).

Given the concern that powerful AI models could "go rogue," high-capability AI systems should be managed like highly sensitive cryptographic keys or nuclear launch codes. Their parameters (the digital weights that determine their behavior) would be encrypted and divided into slices, each entrusted to an independent body – such as the developer, a government agency, and a citizen s' panel. Changing or deploying the system would require consensus, ensuring that unilateral behavior – from any one actor or rogue employee – is impossible (Brundage *et al.* 2018). This "AI escrow" approach would add a technical check to the legal ones.

TECHNOLOGICAL MISUSE

The possibility that AI could be used not just for cyber-attacks but for creating new bioweapons, as Bengio warns, demands a whole new order of preparedness. A government-chartered "Algorithmic Center for Disease Control and Prevention" (A-CDC) could serve as a living laboratory where AI models capable of simulating or manipulating biological/chemical systems above a certain risk threshold must be submitted prior to real-world deployment. This agency would continuously test for vulnerabilities, proactively attempt to "break" models, and develop defensive countermeasures, which would then be shared globally with vetted labs. It would operate as a continuously updated "red team" that anticipates, rather than simply reacts to, emerging algorithmic threats.

Current norms encourage the open publication of powerful AI models, which can create a one-way ratchet: once released, anyone can deploy or modify them. Regulatory bodies should require private and public labs to submit "capability risk assessments" – akin to bio-safety levels in pathogen research – before publishing code or weights for models that could be misused. Only open-sourcing up to a certain risk threshold would be permitted routinely; anything above that would face a structured review with mandatory threat modeling and provisional embargoes, coordinated internationally (Brundage *et al.* 2018).

THE ESCALATING THREAT OF AI WEAPONS AND AUTONOMOUS LETHAL SYSTEMS

As artificial intelligence matures, its integration into weapons platforms and military strategy has emerged as one of the most acute vectors of risk. Lethal autonomous weapon systems (LAWS), which can identify, select, and engage targets without direct human intervention, exemplify the convergence of cutting-edge AI with the destructive power of advanced weaponry. Unlike prior technological advances in arms manufacturing, AI-enabled weapon systems possess the capacity for rapid adaptation, distributed deployment, and even self-directed action – dramatically widening the potential for unforeseen escalation, misuse, or catastrophic failure (Russell 2019; Brundage *et al.* 2018).

Both Geoffrey Hinton and Yoshua Bengio single out autonomous weapons as a uniquely destabilizing risk. Bengio cautions that "as soon

as you have the technology, it will be used somewhere in the world" (Bengio_WSF 2024), warning of a world in which the diffusion of low-cost, highly lethal autonomous systems could lower the threshold for conflict, proliferate to non-state actors, and exacerbate asymmetries in power that destabilize international security. Hinton, similarly, points to the combinatorial threat posed by the intersection of AI with existing weapon systems, highlighting scenarios where "face-recognition drones could be calibrated to target political dissidents, journalists, or even deployed as instruments of ethnic violence or terror" (Hinton_60Minutes 2023).

Existing arms control frameworks – including the Geneva Conventions and the United Nations Conventional Weapons Convention – have proven too slow and politically unwieldy to address these novel threats. The lack of clear, enforceable, and universally-agreed-upon restrictions on the development, deployment, and transfer of AI-powered weapons creates a dangerous vacuum at the frontier of military and dual-use AI research.

Given the rising capability of AI weapons and the real risk of an ungoverned arms race, there is an urgent need for the international community to move beyond voluntary guidelines and individual state-based pledges toward an "internationally accepted moratorium" on the development and use of fully autonomous weapon systems. This moratorium would serve as an immediate, interim "brake," buying critical time for the negotiation of legally binding treaties, normative frameworks, and verification mechanisms tailored to the unique characteristics of AI (Russell 2019).

Such a moratorium should extend not only to the battlefield use of AI weapons, but also to their research, development, and testing – especially for systems capable of selecting and attacking human targets without meaningful human control (ICRC 2021; Future of Life Institute 2017). It should be underpinned by transparent, multi lateral verification; include enforcement provisions for technology transfer and circumvention; and accommodate whistleblower protections within both public and private AI laboratories.

An internationally coordinated moratorium would align with the "precautionary principle" often invoked in discussions of existential risk: when an emerging technology threatens severe or irreversible harm, and scientific consensus is incomplete, policy should err on the side of preemptive restraint (Bostrom 2014). As Hinton and Bengio

have remarked in recent interviews, inaction risks "locking in" the proliferation of AI weapons before society, ethics, and law have an adequate opportunity to respond (Bengio_BBC 2025).

Drawing on the model of prior arms control achievements – such as the bans on chemical and biological weapons, anti-personnel mines, and blinding laser weapons – an effective moratorium on AI weapons will require international cooperation, technical standards for verification, and meaningful involvement of civil society and the scientific community (Altmann and Sauer 2017).

SOCIETAL DISRUPTION AND ECONOMIC DISPLACEMENT

AI's automation potential risks both mass unemployment and the further concentration of wealth. Rather than blunt prohibitions, governments could institute a system of "automation credits" payable by companies for each job function replaced by AI. The rate would be calibrated to local unemployment levels: when new jobs are created in sectors where human labor is scarce – such as green energy or healthcare – the credit would be minimal or even reversed. When jobs are destroyed in already-vulnerable regions, the credit would be high. The proceeds would be pooled into local "Worker Trusts," which could fund wage subsidies, educational retraining, or public works – effectively recycling AI-driven productivity gains back into communities most disrupted by change (Korinek and Stiglitz 2018; Standing 2018).

Every citizen, at birth, could be assigned a "universal learning account" funded partly by automation credits and partly by a tax on AI-driven data harvesting. This account could be drawn on throughout a person's life for continuous retraining, higher education, or even transitioning to new forms of creative or caretaking work – enabling lifelong adaptation as AI transforms the landscape of employment.

SOCIAL AND PSYCHOLOGICAL RISKS

Interviewees express grave concern about society's fracturing into "echo chambers" driven by algorithmic content curation, where "we don't have a shared reality anymore" (Hinton_Diary 2025; Pariser 2011; Alipour *et al.* 2024). Merely adding "labels" or "friction" to

particular posts has so far proven insufficient to restore a common epistemic ground.

A breakthrough regulatory solution would be to require major digital platforms (e.g., social networks, video platforms, news aggregators) to deploy "diversity engines" as part of their feed algorithms (Bojic 2024). By law, a designated proportion of users' feeds – say, 20% – would consist of content surfaced by a certified independent diversity algorithm, not just the platform's own engagement-optimizing system. This diversity engine would be designed to maximize users' exposure to a range of perspectives, including those that challenge their biases or worldview (Moe, Hovden, and Karppinen 2021). Users would retain settings for language and sensitivity, but could not opt out of diverse exposure entirely. The performance of these engines would be regularly audited by an independent commission, ensuring that platforms are not artificially degrading the "diversity feed" or gaming the metrics.

Such an approach borrows from the tradition of "public service broadcasting," but uses algorithmic tooling to ensure that, however much our realities may diverge, we are at least occasionally drawn back into shared information environments (Helberger 2019). This could be further reinforced by algorithmic transparency mandates, requiring platforms to publish summary statistics about the diversity and quality of content presented to users.

REGULATORY CAPACITY

Because AI is evolving rapidly, any fixed law risks obsolescence. Policymakers could introduce "adaptive statutes," a new form of law that is continuously updated via a regulatory large language model (Reg-LLM). This model would monitor case law, technical developments, international agreements, and enforcement outcomes, then propose redlined amendments to statutes on a regular basis, complete with readable explanations and impact analysis. Legislators would vote on these adaptive updates at fixed intervals, vastly increasing the agility and responsiveness of the legal environment (De Filippi, Hassan, and Zicari 2023). All model outputs would be archived on a public blockchain for transparency and post hoc audit, ensuring that regulation keeps pace with the industry it governs.

Policy must also address not just systems and platforms, but people. Both Hinton's and Bengio's sense of duty – and regret – points

to the need for a stronger culture of responsible innovation. Jurisdictions could introduce professional "AI stewardship licenses," requiring advanced AI researchers and lead engineers to maintain a public dossier that is updated every time they contribute to a major system, similar to continuing medical education records for doctors. The dossier would include a brief self-assessment of social impacts, decisions about risk mitigation, and lessons learned. While this may seem soft compared to legal penalties, such transparency would exert peer and reputational pressure, fostering norms that reward responsibility and reflection (Jobin, Ienca, and Vayena 2019; Floridi *et al.* 2018).

Collectively, these creative ideas form an integrated and ambitious AI governance agenda. Regulating "upstream," at the level of compute and model licensing, can help prevent dangerous models from being built unnoticed. Fiduciary duties and escrow split authority ensure that safety is not sacrificed in the pursuit of profit. Algorithmic CDCs and mandatory red-teaming equip governments to preempt misuse rather than merely responding to disasters. Diversification mandates for recommender systems renew the fabric of democratic discourse, even as they preserve individual choice. Adaptive statutes and professional licensing keep regulation fit for purpose and aligned with ethical imperatives.

Enacting such reforms will not be easy. The technical detail required for enforcement, the need for global coordination, and the challenge of keeping up with persistent innovation will test the capacity of current institutions. Nevertheless, waiting for catastrophic outcomes is a risk we cannot afford to take. To harness the opportunities of AI – while genuinely safeguarding humanity, dignity, and democracy – will require a willingness to experiment with regulatory structures as unprecedented as the threats we face.

Table 2. AI risks and potential regulatory innovations

Risk Domain	Key Challenges	Breakthrough Regulatory Solutions (examples)
Existential/Long-term	Surpassing human intelligence; loss of control; unpredictability	Compute passports; conditional model licenses; catastrophe bonds; international AI agency
Alignment/control	Aligning AI values with human ones; lack of direct technical solutions	Fiduciary duties for AI execs; AI escrow systems (split authority over deployment); "kill switch" protocols
Misuse/security	Bioweapons, cyber-attacks, autonomous weapons, open-source risks	Algorithmic Center for Disease Control (A-CDC); mandatory risk assessment before open publishing; secure model submission
Societal/ inequality	Job loss, dignity, and unequal distribution of AI-generated wealth	Automation credits based on local impact; dynamic worker "windfall trusts"; universal learning accounts
Social/ psychological	Polarization, echo chambers, lack of shared reality	Mandated diversity engines for recommender systems; periodic algorithmic audits; user interface controls for recommender bias
Regulatory agility	Static, slow, or uninformed legislative response	Adaptive law with regulatory LLMs; public blockchain of all regulatory changes and their justifications
Ethical and professional	Personal and institutional responsibility; values-conflicted leadership	AI professional stewardship licenses; public impact dossier for lead AI developers; periodic peer review of risk mitigation

Source: Author.

CONCLUSION

The findings of this study hold profound social relevance as the world stands at the threshold of a new technological epoch. Advanced AI systems – especially those built on deep learning – are not only transforming industries but beginning to reshape the very structures of daily life, the labor market, social reality, and even collective imaginaries about the future of humanity. As the voices of Hinton,

Bengio, and LeCun illustrate, the stakes are not merely technical; they are existential, ethical, and democratically political. The range and seriousness of risks identified, from existential threats to democracy's erosion through algorithmic fragmentation, highlight the urgent need for innovative governance solutions that integrate both the promise and peril of emerging AI technologies.

In formal response to the research questions, this study has shown:

- First, the leading pioneers of deep learning articulate a multidimensional risk landscape. Hinton and Bengio are especially vocal about existential threats, the risks of superintelligence, and the potential for irreversible harm if AI systems surpass human intelligence without robust alignment or control mechanisms. LeCun, while recognizing the technical and societal challenges, consistently expresses optimism, emphasizing the amplifying potential of AI, the fallibility of catastrophic scenarios, and the prospects for meaningful regulation and design safeguards.
- Second, there is both substantive overlap and clear divergence in the way these thought leaders prioritize risks and characterize feasible remedies. While all three agree on the importance of governance, they differ substantially in their perceptions of urgency, the plausibility of loss of control, and the necessary trade-offs between innovation and caution. Hinton and Bengio urge regulatory interventions to increase the "stack" (e.g., compute passports, fiduciary duties), while LeCun favors iterative engineering solutions and a balanced approach to platform regulation.
- Third, over time and across interviews, there is a marked trend toward greater explicitness and even humility from those who once steered the field with techno-optimistic assumptions. Bengio's personal journey – from unbridled enthusiasm to a new sense of social duty and caution – mirrors a growing movement among AI professionals toward public engagement and ethical reflection (Jobin, Ienca, and Vayena 2019). All three maintain that, while the exact content of "safety" remains an active research challenge, responsibility for guiding AI's impact cannot be delegated to technologists alone, but must be shared by regulators, democratic institutions, and civil society actors (Floridi *et al*. 2018; Gabriel 2020).

The implications of these findings are wide-ranging. First, they suggest that existing approaches to AI regulation and governance are likely to be insufficient unless fundamentally revised. High-impact, transformative regulatory ideas – such as regulated compute, dynamic model licensing, and mandated diversity engines – offer actionable paths forward but will require new institutional partnerships, international agreements, and a stronger culture of professional responsibility. Second, the need for "regulatory agility" – statutes and standards capable of learning and evolving alongside AI itself – is made evident in both the pace of technological change and the continual discovery of new unintended consequences. Third, the call for robust social and psychological interventions – such as regulated diversification of recommender systems– highlights the urgency of supporting economic adaptability, democratic resilience, and social cohesion (Helberger 2019; Moe, Hovden, and Karppinen 2021).

However, this research has limitations that must be acknowledged. The analysis is based on a purposive sample of interviews with three highly influential experts. While these individuals represent extraordinary technical and social authority, their views are inevitably shaped by personal histories, cultural contexts, and professional networks. The study does not include perspectives from critics of deep learning, social scientists, policymakers, or affected labor groups, nor does it systematically analyze non-English debates or diverse policy contexts. As with all qualitative thematic analyses, the coding and interpretation of content are subject to researcher bias and the limits of available source material. The innovative regulatory proposals reviewed here, while grounded in existing scholarship and creative in conception, have not been tested in practice and may encounter significant political, technical, or economic barriers to implementation.

Future research should address several promising directions. Empirically, there is a need for multi-stakeholder, cross-national studies exploring risk perception and solution preference across technical, regulatory, and public constituencies. Experimental or pilot projects should be launched to evaluate the real-world feasibility and unintended effects of proposals such as compute passports or "diversity engines" for algorithmic feeds. Longitudinal investigation is also warranted to track the evolution of elite expert discourse as AI systems become further embedded in critical societal infrastructure. Ongoing developments such as AI for synthetic biology, autonomous military systems, and large-scale

generative models will require continuous re-examination of both the risk landscape and the adequacy of regulatory responses.

This study reinforces the view – now increasingly accepted by governments and multilateral bodies – that AI governance must no longer be considered a peripheral regulatory issue, but rather a core concern for social order, democratic legitimacy, and the preservation of fundamental rights. Socially and politically, there is also the challenge of ensuring "ethical preparedness" and fostering a culture wherein responsibility for AI safety and benefit is shared and institutionalized at every level of innovation and deployment (Jobin, Ienca, and Vayena 2019; Floridi *et al.* 2018).

To guide AI toward outcomes aligned with human flourishing, societies must invest in the difficult work of regulatory architecture, scientific research on safe AI, cross-sectoral collaboration, democratic engagement, and the cultivation of ethical stewardship throughout the AI value chain.

REFERENCES

Alipour, Shayan, Alessandro Galeazzi, Emanuele Sangiorgio, Michele Avalle, Ljubisa Bojic, Matteo Cinelli, and Walter Quattrociocchi [Alipour *et al.*]. 2024. "Cross-Platform Social Dynamics: An Analysis of ChatGPT and COVID-19 Vaccine Conversations." *Scientific Reports* 14 (1): 2789. DOI:10.1038/s41598-024-53124-x

Altmann, Jürgen, and Frank Sauer. 2017. "Autonomous Weapon Systems and Strategic Stability." *Survival* 59: 117–42. DOI: 10.1080/00396338.2017.1375263

Autor, David H. 2025. "Why Are There Still So Many Jobs? The History and Future of Workplace Automation." *Journal of Economic Perspectives* 29 (3): 3–30.

Bengio_BBC. 2025. "The worst case scenario is human extinction – Godfather of AI on rogue AI. Yoshua Bengio." *BBC Youtube*. https://youtu.be/c4Zx849dOiY

Bengio_WSF. 2024. "Why a Forefather of AI Fears the Future. Yoshua Bengio." *World Science Festival. Youtube.* https://youtu.be/KcbTbTxPMLc

Bessen, James E. 2018. "AI and Jobs: The Role of Demand." *NBER Working Paper*. https://www.nber.org/papers/w24235

Bodroža, Bojana, Bojana M. Dinić, and Ljubiša Bojić. 2024. "Personality Testing of Large Language Models: Limited Temporal Stability, but

Highlighted Prosociality." *Royal Society Open Science* 11: 240180. DOI: 10.1098/rsos.240180

Bojic, Ljubisa. 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. DOI: 10.1186/s40309-022-00208-4

Bojic, Ljubisa. 2024. "AI alignment: Assessing the global impact of recommender systems." *Futures* 160 (103383). DOI:10.1016/j.futures.2024.103383

Bojic, Ljubisa, and Jean-Louis Marie. 2017. "Addiction to Old versus New media." *Srpska politička misao* 56 (2): 33–48. DOI: 10.22182/spm.5622017.2

Bojić, Ljubiša, Irena Stojković, and Zorana Jolić Marjanović. 2024. "Signs of Consciousness in AI: Can GPT-3 Tell How Smart It Really Is?" *Humanities and Social Sciences Communications* 11: 1631. DOI: 10.1057/s41599-024-04154-3

Bojić, Ljubiša, Matteo Cinelli, Dubravko Ćulibrk, and Boris Delibašić [Bojić *et al.*]. 2024. "CERN for AI: A Theoretical Framework for Autonomous Simulation-Based Artificial Intelligence Testing and Alignment". *European Journal of Futures Research* 12: 15. DOI: 10.1186/s40309-024-00238-0

Bojic, Ljubisa, Milos Agatonović, and Jelena Guga. 2024. "The Immersion in the Metaverse: Cognitive Load and Addiction." In *Augmented and Virtual Reality in the Metaverse*, eds. Vladimir Geroimenko. Cham: Springer Nature Switzerland. DOI:10.1007/978-3-031-57746-8_11

Bojić, Ljubiša, Olga Zagovora, Asta Zelenkauskaite, Vuk Vuković, Milan Čabarkapa, Selma Veseljević Jerković, and Ana Jovančević [Bojić *et al.*]. 2025. "Comparing Large Language Models and Human Annotators in Latent Content Analysis of Sentiment, Political Leaning, Emotional Intensity and Sarcasm." *Scientific Reports* 15: 11477. DOI: 10.1038/s41598-025-96508-3

Bojić, Ljubiša, Predrag Kovačević, and Milan Čabarkapa. 2025. "Does GPT-4 Surpass Human Performance in Linguistic Pragmatics?" *Humanities and Social Sciences Communications* 12: 794. DOI: 10.1057/s41599-025-04912-x

Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Braun, Virginia, and Victoria Clarke. 2006. "Using thematic analysis in psychology." *Qualitative Research in Psychology* 3 (2): 77–101.

Brundage, Miles, Shahar Avin, Jack Clark, Helen Toner, Peter Eckersley, Ben Garfinkel, Allan Dafoe, Paul Scharre, Thomas Zeitzoff, Bobby Filar, Hyrum Anderson, Heather Roff, Gregory C. Allen, Jacob Steinhardt, Carrick Flynn, Seán Ó hÉigeartaigh, SJ Beard, Haydn Belfield, Sebastian Farquhar, Clare Lyle, Rebecca Crootof, Owain Evans, Michael Page, Joanna Bryson, Roman Yampolskiy, Dario Amodei [Brundage *et al.*]. 2018. "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation." *arXiv*1802.07228. DOI: 10.48550/arXiv.1802.07228

Brynjolfsson, Erik, and Andrew McAfee. 2014. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. New York: W. W. Norton.

Buolamwini, Joy, and Timnit Gebru. 2018. "Gender shades: Intersectional accuracy disparities in commercial gender classification." *Proceedings of Machine Learning Research* 81: 1–15.

Cath, Corinne. 2018. "Governing artificial intelligence: ethical, legal and technical opportunities and challenges." *Philosophical Transactions of the Royal Society* 376 (2133): 20180080.

Chui, Michael, James Manyika, and Mehdi Miremadi. 2016. "Where Machines Could Replace Humans—and Where They Can't (Yet)." *McKinsey Quarterly*. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/where-machines-could-replace-humans-and-where-they-cant-yet

De Filippi, Primavera, Samer Hassan, and Roberto Zicari. 2023. "Smart contracts meet legal responsibility: The role of Reg-LLMs in adaptive law." *Stanford Journal of Blockchain Law & Policy* 6 (1): 77–104.

Doshi-Velez, Finale, and Been Kim. 2017. "Towards a rigorous science of interpretable machine learning." *arXiv* 1702.08608. DOI: 10.48550/arXiv.1702.08608

Esteva, Andre, Brett Kuprel, Roberto Novoa, Justin Ko, Susan Swetter, Helen Blau, Sebastian Thrun [Esteva *et al.*]. 2017. "Dermatologist-level classification of skin cancer with deep neural networks." *Nature* 542 (7639): 115–118. DOI: 10.1038/nature21056

Feldstein, Steven. 2019. "The Global Expansion of AI Surveillance." Carnegie Endowment for International Peace. https://carnegieendowment.org/research/2019/09/the-global-expansion-of-ai-surveillance?lang=en

Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Christoph Lütge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke, Effy Vayena

[Floridi *et al.*]. 2018. "AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations." *Minds and Machines* 28 (4): 689–707.

Gabriel, Iason. 2020. "Artificial intelligence, values, and alignment." *Minds and Machines* 30 (3): 411–437.

Helberger, Natali. 2019. " On the democratic role of news recommenders." *Digital Journalism* 7 (8): 993–1012.

Hinton_60Minutes. 2023. "Godfather of AI" Geoffrey Hinton: The 60 Minutes Interview. *60 Minutes. Youtube.* https://youtu.be/qrvK_KuIeJk

Hinton_Diary. 2025. "Godfather of AI: I Tried to Warn Them, But We've Already Lost Control! Geoffrey Hinton." *The Diary of A CEO. Youtube.* https://youtu.be/giT0ytynSqg

Jobin, Anna, Marcello Ienca, and Effy Vayena. 2019. "The global landscape of AI ethics guidelines." *Nature Machine Intelligence* 1 (9): 389–399.

Korinek, Anton, and Joseph E. Stiglitz. 2018. "Artificial intelligence and its implications for income distribution and unemployment." In *Economics of Artificial Intelligence: An Agenda*, eds. Ajay Agrawal, Joshua Gans, and Avi Goldfarb, 259–290. Chicago: University of Chicago Press.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. 2017. "ImageNet classification with deep convolutional neural networks." *Communications of the ACM* 60 (6): 84–90.

LeCun_Brian. 2024. "Yann LeCun: AI Doomsday Fears Are Overblown [Ep. 473]." *Dr Brian Keating. Youtube.* https://youtu.be/u7e0YUcZYbEhttps://youtu.be/u7e0YUcZYbE

LeCun_Lex. 2024. "Yann Lecun: Meta AI, Open Source, Limits of LLMs, AGI & the Future of AI | Lex Fridman Podcast #416." *Lex Fridman. Youtube.* https://youtu.be/5t1vTLU7s40https://youtu.be/5t1vTLU7s40

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. "Deep learning." *Nature* 521 (7553): 436–444.

Metz, Cade. 2023. "'The Godfather of A.I.' Leaves Google and Warns of Danger Ahead." *The New York Times*. https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html

Moe, Hallvard, Jan Fredrik Hovden, and Kari Karppinen. 2021. "Operationalizing Exposure Diversity." *European Journal of Communication* 36 (2): 148–67. DOI: 10.1177/0267323120966849

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, Sendhil Mullainathan [Obermeyer *et al.*]. 2019. "Dissecting racial bias in an algorithm used to manage the health of populations." *Science* 366 (6464): 447–453.

O'Connor, Cailin, and James Owen Weatherall. 2019. *The Misinformation Age: How False Beliefs Spread*. New Haven: Yale University Press.

O'Neil, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Broadway Books.

OpenAI. 2023. "GPT-4 Technical Report." *OpenAI*. https://cdn.openai.com/papers/gpt-4.pdf

OSF 2025. "Risk and Responsibility at the Frontier of AI." *OSF*. https://osf.io/s4gva/?view_only=b42e04616959401082178f6c4c8376ce

Pariser, Eli. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin Press.

Pavlovic, Maja, and Ljubisa Bojic. 2020. "Political Marketing and Strategies of Digital Illusions – Examples from Venezuela and Brazil." *Sociološki pregled* 54 (4):1391–1414. DOI: 10.5937/socpreg54-27846

Reinhardt, Anne, Jörg Matthes, Ljubisa Bojic, Helle T. Maindal, Corina Paraschiv, and Knud Ryom [Reinhardt *et al.*]. 2025. "Help Me, Doctor AI? A Cross-National Experiment on the Effects of Disease Threat and Stigma on AI Health Information-Seeking Intentions." *Computers in Human Behavior* 172: 108718. DOI: 0.1016/j.chb.2025.108718

Rudin, Cynthia. 2019. "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead." *Nature Machine Intelligence* 1 (5): 206–215.

Russell, Stuart J. 2019. *Human Compatible: Artificial Intelligence and the Problem of Control*. New York: Viking.

Sandvig, Christian, Kevin Hamilton, Karrie Karahalios, Cedric Langbort. [Sandvig *et al*.]. 2016. "Auditing algorithms: Research methods for detecting discrimination on Internet platforms." *Data and Discrimination: Converting Critical Concerns into Productive Inquiry. Social Science Research Council*. https://ai.equineteurope.org/system/files/2022-02/ICA2014-Sandvig.pdf

Schmidhuber, Jürgen. 2015. "Deep learning in neural networks: An overview." *Neural Networks* 61: 85–117.

Silver, David, Aja Huang, Christopher Maddison, Arthur Guez, Laurent Sifre, George Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel,

Demis Hassabis [Silver *et al.*]. 2016. "Mastering the game of Go with deep neural networks and tree search." *Nature* 529 (7587): 484–489.

Standing, Guy. 2018. *Basic Income: And How We Can Make It Happen*. London: Pelican Books.

Susskind, Daniel. 2020. *A World Without Work: Technology, Automation, and How We Should Respond*. New York: Metropolitan Books.

Turing. 2018. "Fathers of the Deep Learning Revolution Receive ACM A.M. Turing Award. Bengio, Hinton and LeCun Ushered in Major Breakthroughs in Artificial Intelligence." *Association for Computing Machinery*. https://awards.acm.org/about/2018-turing

Turing_Bengio. 2018. "Yoshua Bengio." *AM Turing Award*. https://amturing.acm.org/award_winners/bengio_3406375.cfm

Turing_Hinton. 2018. "Geoffrey Hinton." *AM Turing Award*. https://amturing.acm.org/award_winners/hinton_4791679.cfm

Turing_LeCun. 2018. "Yann LeCuno." *AM Turing Award*. https://amturing.acm.org/award_winners/lecun_6017366.cfm

West, Sarah Myers. 2019. "Data Capitalism: Redefining the Logics of Surveillance and Privacy." *Business & Society* 58 (1): 20–41.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs.

Љубиша Бојић*

Истраживачко-развојни институт за вештачку интелигенцију Србије, Нови Сад

Институт за филозофију и друштвену теорију, Универзитет у Београду

Лабораторија за дигитално друштво, Београд

Complexity Science Hub, Беч, Аустрија

РИЗИК И ОДГОВОРНОСТ НА ГРАНИЦИ ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ: ТЕМАТСКА АНАЛИЗА СТАВОВА ПИОНИРА ДУБОКОГ УЧЕЊА О ПРЕТЊАМА И УПРАВЉАЊУ ВИ**

Резиме

Ова студија приказује резултате квалитативне тематске анализе интервјуа са три најзначајнија пионира дубоког учења – Џефријем Хинтоном, Јошуом Бенџијом и Јаном ЛеКуном, добитницима Тјурингове награде, у вези са ризицима вештачке интелигенције (ВИ) и изазовима управљања. Анализа издваја седам главних тематских кластера: егзистенцијалне и дугорочне опасности, друштвене ризике (запосленост, неједнакост), злоупотребе (кибер, биолошко оружје), друштвено-психолошке утицаје (манипулација, поларизација), питања управљања и регулације, позитивне потенцијале, те лична осећања и одговорности научника. Хинтон и Бенџио наглашавају егзистенцијалне ризике и потребу за глобалном регулативом, недвосмислено упозоравајући на потенцијал изумирања људске врсте у уколико не буде адекватних механизама за контролу и усклађивање вредности ВИ са људским интересима. Бенџио додатно показује еволуцију свог става, прелазећи од техно-оптимизма ка осећају друштвене дужности

* Имејл: ljubisa.bojic@ivi.ac.rs; ORCID: 0000-0002-5371-7975

** Ова студија реализована је уз подршку Министарства науке, технолошког развоја и иновација Републике Србије према Уговору о реализацији и финансирању научноистраживачког рада број 451-03-136/2025-03/ 200025 и Уговору о реализацији и финансирању научноистраживачког рада број 451-03-136/2025-03

и опреза. ЛеКун, са друге стране, одбацује катастрофичне сценарије и заговара децентрализован, отворен развој, верујући у техничке и друштвене капацитете да се ризици успешно савладају, манифестујући снажан оптимизам. Сви саговорници препознају ризик од социјалних подела кроз манипулативне алгоритме, као и потенцијал масовне незапослености и концентрације богатства, али се разликују у хитности и врсти препоручених решења. Хинтон и Бенџио сугеришу регулисање на нивоу инфраструктуре и увођење нових глобалних институција, док ЛеКун заступа постепене техничке и друштвене адаптације. Међу иновативним предлозима за управљање ВИ издвајају се "*compute* пасоши", условне лиценце за напредне моделе, професионалне дозволе за истраживаче и увођење разноврсности садржаја на великим платформама за борбу против негативних алгоритамских утицаја. Посебно је наглашена потреба за адаптивном регулативом која би могла да прати развој технологије, као и за јачањем професионалне одговорности и друштвене свести. Резултати указују да експертска мишљења о ВИ ризику нису једногласна, већ прожета значајним дискусијама о степену претње и приоритетима у политици и регулисању ВИ. Закључује се да су иновативне, интернационално координиране регулаторне интервенције неопходне ради балансирања потенцијала и претњи ВИ, те да професионална и друштвена одговорност развојних инжењера мора постати кључна компонента новог модела управљања. Будућа истраживања треба да обухвате шири спектар актера, укључујући различите друштвене, политичке и међународне перспективе, и тестирање предложених механизама у реалним условима, имајући у виду да се природа ризика динамично мења са ширењем и усвајањем ВИ у новим доменима.

Кључне речи: вештачка интелигенција, пионири дубоког учења, закони који регулишу ВИ, егзистенцијалне претње, тематска анализа

* This manuscript was submitted on July 5, 2025, and accepted by the Editorial Board for publishing on November 10, 2025.

UDC 327:004.8:347.7(4-672EU:497.11)
DOI: 10.5937/spm94-59321
Review article

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 167-185

Mirjana Dejanović[*]

Faculty of Project and Innovation Management "Prof. dr. Petar Jovanović," Educons University, Belgrade

Anatolijs Krivins[**]

Department of Law, Management Science and Economics, Daugavpils University, Latvia

LEGAL AND FINANCIAL REGULATION OF AI IN SERBIA, LATVIA, AND THE EU

Abstract

Artificial Intelligence (AI) is being rapidly integrated across numerous sectors, with the financial industry at the forefront of technological transformation. It is influencing processes such as credit evaluation, fraud detection, risk modelling, and personalised financial services. This expansion increases the need for coherent legal frameworks, ethical governance, and financial sustainability, particularly regarding user protection, algorithmic transparency, safety, and liability for negative outcomes. This study offers a comprehensive comparative analysis of AI regulatory frameworks in the Republic of Serbia, Latvia, and the European Union (EU). Using a multidisciplinary qualitative methodology that integrates normative legal analysis, economic evaluation, and ethical considerations, the research is investigates Serbia's harmonisation efforts with EU standards, Latvia's implementation of the EU AI Act and complementary financial mechanisms, and the EU's comprehensive regulatory architecture, including the AI Act and GDPR. Key findings identify shared legal challenges such as accountability, transparency,

[*] E-mail: mirjana.dejanovic@pmc.edu.rs; ORCID: 0000-0003-3369-331X
[**] E-mail: anatolijs777@gmail.com; ORCID: 0000-0003-1764-4091

data protection, and the agility of regulatory frameworks in adapting to rapid AI innovation. From a financial perspective, the study identifies compliance costs for small and medium-sized enterprises (SMEs), confusion in liability distribution, and budgetary pressures on supervisory authorities. Regulatory fragmentation presents a significant barrier to cross-border AI innovation and internal market coherence. The paper concludes with policy recommendations focused on enhancing regulatory harmonisation, financial feasibility, and ethical guidance, aiming to foster sustainable AI integration that balances innovation with social protections.

Keywords: Artificial Intelligence, legal regulation, financial regulation, ethics, Serbia, Latvia, European Union

INTRODUCTION

Artificial Intelligence (AI) is rapidly transforming diverse sectors, extending far beyond everyday applications and fundamentally reshaping the financial industry. Its integration into key areas such as credit risk assessment, fraud prevention, algorithmic decision-making, and the provision of personalised financial services signals a paradigm shift in operational efficiency and innovation. Aligned with these advancements are significant legal, ethical, and economic challenges that demand the creation of adaptive and forward-looking regulatory frameworks. These frameworks must strike a balance between encouraging innovation and protecting fundamental rights, ensuring transparency, and preserving financial stability.

This study presents a comparative analysis of AI regulation in the Republic of Serbia, the Republic of Latvia, and the EU. These jurisdictions, while divergent in their legal traditions, institutional architectures, and socio-political contexts, converge in their ambition to address the multifaceted complexities of AI governance. The financial sector in Serbia, Latvia, and the EU faces high compliance costs with AI regulations, which particularly affect small and medium-sized enterprises due to their significantly smaller capacities, making compliance costs (e.g. investments in technology, employee training) much higher compared to large corporations, while General Data Protection Regulation

(GDPR) complicates data collection and processing, requiring greater transparency and user consent.

The primary aim of this research is to critically analyse the legal and financial dimensions of AI regulation across these three jurisdictions, with particular attention to institutional effectiveness, legislative consistency, and long-term economic viability. The central hypothesis is that regulatory alignment, driven by ethical principles and economic logic, enhances the legitimacy, effectiveness, and social faith in AI governance. Methodologically, the research applies a triangulated approach, combining normative-descriptive analysis, comparative legal methods, and financial impact assessment.

HYPOTHESIS, THEORETICAL AND METHODOLOGICAL FRAMEWORK, AND LITERATURE REVIEW

This research is grounded in the assumption that the effective regulation of AI in the financial sector must rest on a harmonised legal and ethical foundation that balances innovation with economic sustainability. The central hypothesis is that jurisdictions, which align their AI regulatory practices with EU standards, are more likely to achieve legal legitimacy, economic efficiency, and public trust. In addition, Serbia and Latvia, though structurally distinct, converge towards a common regulatory trajectory under the influence of EU integration and harmonisation imperatives. This issue requires harmonised regulation at a global, international level, not just within the EU.

From a theoretical standpoint, the study adopts an interdisciplinary approach, integrating concepts from jurisprudence, regulatory theory, economic analysis of law, and applied ethics. The concept of legal legitimacy, as articulated in modern legal theory, is central to the analysis, particularly in the context of automated decision-making and algorithmic uncertainty. Ethical dimensions are based on discussions about human-centric AI, fairness in automated systems, and risk governance, drawing on prominent theorists in digital ethics (Floridi *et al*. 2018; Floridi 2019).

Methodologically, the research employs a tripartite structure. First, a normative-descriptive method is applied to analyze legal texts, national strategies, and institutional frameworks in Serbia, Latvia, and the EU. This approach provides a deeper understanding of how different

countries implement and adapt their legal systems to align with EU standards in the field of AI.

Existing literature provides a solid foundation for this investigation, with Lilkov (2021) pointing out the institutional and normative challenges of implementing the EU AI Act, cautioning against overregulation and fragmentation across EU member states. AI tools like Lexis+AI, Ask Practical Law AI, and Westlaw AI, while improving research efficiency, struggle with accuracy, especially due to hallucinations that undermine the reliability of legal conclusions. To be trustworthy, AI must resolve these issues and ensure the credibility of legal information (Magesh *et al.* 2025). In public procurement, AI has demonstrated clear advantages, such as increased efficiency and cost savings, but it also raises ethical concerns, including algorithmic bias, data privacy, and transparency, with recommendations for clearer guidelines and professional training (Obinna and Kess-Momoh 2024). The explainability of AI (XAI) is crucial for trust and transparency in the financial sector, though challenges remain with the reliability of explanations and risks from potentially misleading information (Yeo *et al*. 2025). In addition, AI has a transformative impact on traditional accounting practices, enhancing efficiency and accuracy, though concerns over data privacy persist (Odonkor *et al*. 2024). Big data and AI are expected to revolutionize customer interactions, creating new growth opportunities for financial institutions to improve efficiency and offer innovative services (Ahmadi 2024).

Together, this framework provides a robust analytical view through the regulatory dynamics of AI in the financial sector, critically examined across three distinct interrelated legal environments. This offers a better assessment of how different legal frameworks and regulations impact the effectiveness and ethics of AI technology implementation.

AI REGULATION IN SERBIA, LATVIA, AND THE EU

This chapter examines the legal and financial aspects of AI regulation in Serbia, Latvia, and the EU, focusing on the harmonisation of legal norms, risks in the financial sector, and the impact on market participants, legislators, and the economy. The AI Act, which entered into force on 1st August 2024, is the world's first comprehensive legal framework for AI and classifies AI systems based on risk categories (unacceptable, high, limited, and minimal). High-risk systems, especially

those used in finance, law enforcement, or employment, must comply with strict transparency, data governance, and human oversight requirements. Still, the GDPR continues to play a key role in regulating data collection, consent, and privacy in AI operations, while its application significantly transforms the future of capital markets, banking, insurance, and fintech companies by introducing numerous innovations and changes. The AI Act's entry into force marked a significant milestone, setting global standards for the responsible application of AI technologies (EUR-lex 2021; European Commission 2024). In Serbia, Latvia, and the EU, the development of the legal framework for AI applications, especially in the financial domain, is evolving in accordance with the unique characteristics of each jurisdiction while taking into account common European guidelines and national strategies.

AI Regulation in Serbia

In Serbia, the AI legal landscape is shaped by efforts to support digital transformation and align it with EU standards. The country has adopted a Law on Personal Data Protection modelled after the GDPR (Zakon o zaštiti podataka o ličnosti 2018). In practice, many institutions, particularly small and medium-sized enterprises, face challenges in implementing these laws due to a lack of employee training, technical resources, and high compliance costs. The lack of clearly defined responsibilities for supervision and implementation of the laws creates uncertainty and can slow down progress in the application of AI technologies.

Serbia also has a National Strategy for the Development of Artificial Intelligence, which includes goals such as promoting AI innovation, building human capacities, and ensuring the ethical use of AI (Government of the Republic of Serbia 2024). A key element of this strategy is the introduction of a regulatory sandbox designed to facilitate the experimental validation of innovative AI solutions in a controlled environment without the requirement for prior registration or licensing. This regulatory innovation has significantly contributed to the accelerated development and rigorous evaluation of novel business models, especially in the field of payment services. Smart cities in Serbia increasingly integrate AI to enhance the efficiency and sustainability of urban infrastructure. Belgrade uses AI for traffic optimisation and air quality monitoring, while Novi Sad utilises AI

in the management of public transport and parking systems. Niš has implemented intelligent lighting and AI-supported security systems. A new Strategy for the Development of AI for the period 2025–2030 focuses on the legislative framework, expanding educational capacities, advancing digital infrastructure, and promoting the integration of AI technologies into the public sector (Vlada Republike Srbije 2025). Additionally, Serbia is a member of the Global Partnership on Artificial Intelligence (GPAI), underscoring its strategic commitment and active engagement in the global discourse on AI development and application. Lilkov warns that AI can also be weaponized in hybrid warfare and influence operations, making robust EU coordination essential (Lilkov 2021). Countries worldwide invest in AI development, while Serbia is working on strategies for the responsible application of this technology, emphasising the need for international cooperation to establish norms and avoid undesirable effects (Luknar 2025). As Hildebrandt emphasises, legal frameworks must evolve together with emerging technologies to preserve fundamental rights and democratic principles (Hildebrandt 2015).

The development of technical infrastructure, alongside continuous education and the engagement of highly qualified professionals, forms the foundation for the operational implementation of AI solutions. Interdisciplinary approaches that combine regulatory, technological, and financial aspects are crucial for the successful implementation of AI in Serbia. Strengthening the judiciary and inspection bodies is key to protecting rights, ensuring fiscal stability, and boosting AI investment.

The Government of the Republic of Serbia's Strategy for the Development of Artificial Intelligence for the period 2025–2030 explicitly emphasises the need for continuous evaluation of the economic and social effects of AI technology implementation, thereby ensuring the sustainability and functionality of AI systems in line with global development trends (Vlada Republike Srbije 2025). This approach confirms Serbia's commitment to developing an AI sector grounded in legality, efficiency, and transparency, which are fundamental prerequisites for long-term social and economic prosperity. Ethical aspects of the development and application of AI have been further emphasised through the official document "Ethical guidelines for the development, application, and use of trustworthy and responsible artificial intelligence" (Government of the Republic of Serbia 2023). Serbia established a working group for drafting the AI Act, systematically preparing to align

with the EU regulatory framework while simultaneously considering the specificities of the domestic legal and economic context (Nacionalna platforma za veštačku inteligenciju 2024).

AI Regulation in Latvia

The issue of artificial intelligence has been in the sights of Latvian society and the Latvian government for several years. Undoubtedly, over time, the focus of attention of Latvian political leaders is being refined in accordance with current developments. If initially the emphasis was placed on smart technologies in order to convince a wide number of Latvian residents that they have not only become necessary in the world of today and tomorrow, but that they are even a desirable tool for the development of human abilities, then later attention was paid to practical solutions and regulatory frameworks (Levits 2019). When preparing the planning document "Digital Transformation Guidelines for 2021–2028," Latvia identified the implementation of AI solutions as one of the priorities of public administration. Latvia's national regulation in the field of AI, in the context of EU regulation, is mainly focused on managing potential risks as stated in the Latvian information report "On the development of artificial intelligence solutions" (Republic of Latvia 2020). The growth of AI systems will create several new challenges in the area of legal regulation. In the future, AI systems will make increasingly complex and responsible decisions (Likumi.lv 2023). This means that a fundamental issue of AI legal regulation in Latvia is liability for incorrect, harmful, or illegal actions (actions or inactions) performed by an AI-driven system in a process where no human was involved. For example, if a train derails during a journey and people suffer, then the owner of the train or tracks is liable, depending on where the technical failure occurred. In the case of AI-driven systems, there will be disagreements about who is liable, because the owner receives an already trained system from the manufacturer and is not always competent in the technical complexities of AI. On the other hand, AI systems are often "trained" continuously; the owner of the AI system himself continues to "train" the system and adapt it to specific conditions. Service companies can also make changes to the AI system. If an AI-driven car causes damage, then improvements to the regulatory framework may be necessary to determine who is at fault, whether it is the manufacturer, service provider, owner, or user, depending on the

degree of autonomy of the car. The more neural layers an AI system has and the larger the data set from which the system is trained, the more impossible it is to determine why the AI system has made one decision or another. There are situations when accidents cannot be avoided; the only question is: who becomes the victim? For example, while trying to avoid a collision with a person running down the street, a car can drive into a ditch and cause harm to the passengers.

The AI system will also have to decide on the least harm, for example, in the case of rescue drones, when the system will have to decide which of the drowning people to save first, taking into account that the others may drown. One of the evaluation criteria could be the chance of a drowning person surviving. In public services, AI tools will also make more and more decisions that could lead to direct losses for private individuals. For example, AI may refuse a construction permit for a formal reason, although a person would never have made such a decision. In such processes, at least initially, it is mandatory for a person to be involved in verifying the answer. The main requirements regarding the safety of goods and services and liability for their defects in Latvia are stipulated in the Law on the Safety of Goods and Services and the Law "On Liability for Defects in Goods and Services" (Likumi.lv 2004).

At the national scale, it has been assessed that Latvia is unlikely to rival major global powers such as the United States or China in the broad development of artificial intelligence. Nevertheless, Latvia holds the potential to excel as a frontrunner in particular AI applications. The country's principal competitive edge resides in its capacity to swiftly deploy AI solutions within both the private sector and governmental institutions, enabling more rapid innovation and implementation. Therefore, AI skills should be acquired not only by ICT specialists, but also by a wider range of users and managers (Likumi.lv 2023).

AI development strategies in Latvia impact sectors like defense, crime prevention, culture, traffic, justice, finance, construction, agriculture, and environmental protection. AI aids in social network monitoring, accident prevention, criminal identification, resocialization, research, and analysis. It also supports investment planning, combats disinformation, and improves business and government processes. A key priority is promoting the Latvian language and cultural data in AI solutions. The Latvian National Terminology Portal, developed by the State Language Centre, the Terminology Commission, and the Centre for Cultural Information Systems, provides sector-specific terminology to

support AI implementation and capacity building in public administration (Likumi.lv 2023).

The goal of Latvia's AI Development Law is to integrate the Latvian language into AI, create an AI ecosystem, and establish a legal framework for collaboration between the public sector, private sector, and universities. It aims to foster innovation, promote AI initiatives aligned with national interests, ensure equal access to AI, and ensure ethical, responsible, and human rights-respecting AI use (Government of Latvia 2025). AI provides the opportunity to learn from users' experience in communicating with public administration and predict what services, in what order (including services provided by different institutions) may be needed by a particular person or company, depending on their socioeconomic profile and life situation.

Within the framework of the project "Development of a Virtual Assistant Platform," Latvia continues to develop a virtual assistant platform, which can be used and integrated into its solutions by any public administration institution, where information and other services are provided to individuals and legal entities. For some objects, for example, structures with great cultural and historical value, state and municipal objects, mandatory submission of documents in machine-readable format may be established.

In Latvia, AI is increasingly being used in law firms. Significant issues related to artificial intelligence arise in the areas of intellectual property and copyright, freedom of expression, scientific and artistic creativity, freedom of information, and freedom of the press. For several years now, the world's largest media outlets have been using artificial intelligence to create their content, and this has recently been happening in Latvia as well.

On May 16, 2024, the E-Case Supervisory Council decided to integrate the Competition Council, Insolvency Control Service, and Corruption Prevention Bureau into a unified E-Case system, improving investigation and judicial processes in Latvia (Jurista Vārds 2024). This project is part of the EU Recovery Fund Plan, which supports green and digital economy transitions and crisis recovery. AI is already being used in Latvia's anonymization tool to speed up document processing and reduce manual work (Jurista Vārds 2024). Latvia's role in implementing the EU AI Act and accessing EU funding mechanisms the EU Artificial Intelligence Act is designed to safeguard fundamental rights, democratic principles, the rule of law, and environmental sustainability against the

risks posed by high-risk AI systems. Simultaneously, it seeks to foster innovation and position Europe as a global leader in the development and regulation of artificial intelligence technologies (European Parliament 2023). Latvia's role in the implementation of the Act is mainly related to putting the solutions envisaged in the Act into practice (prohibited uses; exceptions for law enforcement; obligations regarding high-risk AI systems; transparency requirements; measures to support innovation and SMEs, etc.).

AI can provide significant competitive advantages and drive positive outcomes for society and the environment. In addition, AI drives innovation, and its application increases productivity. Focusing on AI will deliver economic benefits and solutions in healthcare, education, public safety, national defence, etc. (Anotācija – Valsts kanceleja 2024).

The Law on the Artificial Intelligence Center in Latvia, effective March 20, 2025, aims to establish an AI ecosystem and a legal framework for cooperation between the public and private sectors, as well as universities. It defines the purpose, structure, tasks, rights, financing, and fund usage of the "Artificial Intelligence Center" foundation (Likumi.lv 2025). The Artificial Intelligence Centre promotes AI use across sectors, coordinates projects, and secures funding. It identifies AI security risks, advises on AI in elections, raises awareness on AI skills, and addresses human rights and democracy risks. The Centre supports Latvian language inclusion, organizes training data, and ensures a regulatory environment for AI development. Latvia places strategic importance on accessing EU funding mechanisms, particularly under the 2021–2027 programming period. The digitalization of higher education institutions is being assessed in line with the Common Provisions Regulation governing ERDF, ESF+, the Cohesion Fund, EMFAF, and other internal security and migration instruments. Targeted funding is essential for pilot projects testing innovative e-governance solutions. Cross-border data harmonisation and exchange are essential for effective AI training due to limited national data. Additional funding could support the development of tools to detect systemic corruption, such as repeated procurement victories by politically connected firms.

AI Regulation in the EU

In 2024, the European Union adopted Regulation (EU) 2024/1689, commonly referred to as the Artificial Intelligence Act (AI Act), thereby establishing the world's first comprehensive legal framework for regulating AI systems (Regulation 2024/1689). There is a high importance in regulation that stimulates digital transformation while simultaneously taking into account ethical and social aspects (Dejanović 2024a). High-risk AI systems, including those used in medical devices, education, employment, and law enforcement, must comply with stringent standards encompassing risk assessment, data quality, human oversight, and cybersecurity measures. The regulation explicitly prohibits unacceptable-risk AI systems, such as social scoring technologies and those designed for behavioral manipulation. The European Artificial Intelligence Office oversees enforcement, with penalties up to €35 million or 7% of global turnover (White & Case 2025).

In March 2025, the third draft of the Code of Good Practice for General-Purpose AI models was published, providing guidelines on transparency, copyright, and risk mitigation. The AI Act applies extraterritorially across all sectors (European Commission 2025a). On 9 April 2025, the European Commission introduced the "European AI Action Plan," focusing on infrastructure, data access, sectoral AI promotion, skill development, and streamlined regulation through a new AI Act Support Office. The regulation also introduces specific standards for general-purpose AI systems to ensure safe deployment. The AI Act is part of a broader regulatory framework, including the Digital Operational Resilience Act (DORA) and the Markets in Crypto-Assets Regulation (MiCA), which enhance digital and financial sector resilience. The European Central Bank highlights risks like herd behaviour and cyber-attacks, calling for continuous supervision. The Act has phased implementation, running from 2025 to 2027, with countries like Latvia and Serbia actively aligning their national strategies (European Commission 2025b).

Gstrein, Haleem, and Zwitter criticise the Act for its lack of clarity and inadequate regulatory mechanisms for managing powerful AI models. They suggest more defined transparency and risk protocols, converting the AI Office into an independent entity, consolidating advisory bodies, and enhancing coordination within the EU. They warn that the Act risks being symbolic without these improvements

and suggest embedding adaptive governance mechanisms to address emerging challenges (Gstrein, Haleem, and Zwitter 2024).

The AI Act aims to balance innovation, safety, and fundamental rights protection. Its success depends on clear definitions, consistent enforcement, adaptive governance, and ongoing stakeholder collaboration. At the European Systemic Risk Board (ESRB) conference in Brussels (March 2025), regulators emphasised the importance of robust AI oversight in finance, acknowledging AI's efficiency benefits but also its systemic risks. The adoption of AI and ICT in SMEs is very important because it significantly boosts competitiveness and economic growth (Dejanović 2024b). The conference called for joint efforts among regulators, banks, and tech firms to develop standards for AI risk management, particularly in automated decision-making and algorithmic trading. Transparency, accountability, and AI professional training were also stressed to minimize negative impacts (Lagarde 2025).

LEGAL AND FINANCIAL CHALLENGES IN AI REGULATION

Artificial intelligence drives economic growth by optimising processes and enabling innovation across sectors (Dejanović 2023). However, the increasing autonomy of AI systems raises unresolved legal questions, particularly around liability. Traditional frameworks based on human accountability are insufficient, prompting calls for new legal models that distribute responsibility across developers, manufacturers, and users to ensure legal clarity and accountability. Effective regulation demands institutional adaptation, including the establishment of specialised bodies such as AI offices, regulatory boards, and expert panels at both national and EU levels (Novelli *et al.* 2024). Transparency and explainability are becoming critical components of regulatory design, given the opacity of complex algorithms. In the financial sector, the S.A.F.E. framework focusing on sustainability, accuracy, fairness, and explainability has been proposed to assess the reliability of AI models (Giudici and Raffinetti 2023).

The EU AI Act frames AI as a cross-sectoral technology, which complicates its regulatory enforcement. Aligning horizontal technical standards with sector-specific responsibilities remains a challenge, particularly regarding fundamental rights protection. Given the pace of technological change, regulatory learning must be continuous. The

concept of a "regulatory learning space" where regulators, industry, and academia exchange knowledge is emerging as a core principle for adaptive governance. Open data initiatives at the EU level can further support transparency and informed oversight (Lewis *et al.* 2025).

Financially, the cost of compliance is significant, especially for SMEs. Obligations related to transparency, risk assessment, and rights protection increase operational burdens (Dejanović 2025). Thus, flexible legal tools such as regulatory sandboxes, adaptive clauses, and ethical guidelines like the OECD AI Principles are essential to balance innovation with safeguards. Finally, legal certainty, interdisciplinary collaboration, and international coordination are key to addressing the legal and financial risks posed by AI. Programs such as the EU's AI Literacy initiative aim to enhance public understanding, while scholars emphasise the urgent need for robust compensation mechanisms and transparent liability structures in high-risk sectors like finance (Mirishli 2025).

REFERENCES

Ahmadi, Sina. 2024. "A Comprehensive Study on Integration of Big Data and AI in the Financial Industry and Its Effect on Present and Future Opportunities." *International Journal of Current Science Research and Review* 7 (1): 66–74. DOI: 10.47191/ijcsrr/V7-i1-07

Anotācija – Valsts kanceleja. 2024. "Tiesību akta projekts "Mākslīgā intelekta attīstības likums."" Pēdējoreiz piekļūts 2025. gada 12. februārī. https://tapportals.mk.gov.lv/annotation/a053aceb-22ee-47d9-9ace-8d1a28862d03#

Dejanović, Mirjana S. 2023. "Strategy for the implementation of socially responsible behavior in situation of economic crisis." *Baština* 33 (61): 235–246.

Dejanović, Mirjana S. 2024a. "Effect of COVID-19 and Economic Crisis on Entrepreneurship." *International Journal of Engineering, Management and Humanities* 5 (5): 261–269.

Dejanović, Mirjana S. 2024b. "The Application of the Latest Information and Communication Technologies (ICT) and Artificial Intelligence (AI) in SME Business Operations as a Driver of GDP Growth." Paper presented at the 51st International Symposium on Operations Research (SYM-OP-IS 2024), Tara, Serbia. Last Accessed on March 14, 2025. https://symopis2024.ftn.uns.ac.rs/wp-content/uploads/2024/11/SYM-OP-IS-2024_PROCEEDINGS_final.pdf

Dejanović, Mirjana S. 2025. "The Role of Artificial Intelligence in Enhancing Economic Efficiency and Innovation." *Forum for Economic and Financial Studies* 3 (1): 1541. DOI: 10.59400/fefs1541

EUR-lex. 2021. "Proposal for Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts." Last Accessed on February 12, 2025. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206

European Commission. 2024. "AI Act enters into force." Last Accessed on March 21, 2025 https://ec.europa.eu/newsroom/ai-act-enters-into-force

European Commission. 2025a. "Third Draft of the General-Purpose AI Code of Practice published, written by independent experts." 11.03.2025. Last Accessed on March 20, 2025. https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

European Commission. 2025b. "Shaping Europe's leadership in artificial intelligence with the AI Continent action plan." Last Accessed on March 21, 2025. https://commission.europa.eu/topics/eu-competitiveness/ai-continent_en

European Parliament. 2023. "Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI." 09.12.2023. Last Accessed on March 21, 2025. https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai

Floridi, Luciano. 2019. "Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical." *Philosophy & Technology* 32 (2): 185–193. DOI: 10.1007/s13347-019-00354-x

Floridi, Luciano, Josh Cowls, Matteo Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, Virginia Dignum, Christoph Luetge, Robert Madelin, Ugo Pagallo, Francesca Rossi, Burkhard Schafer, Peggy Valcke, Effy Vayena [Floridi *et al*.]. 2018. "AI4People—An ethical framework for a good AI society Opportunities, risks, principles, and recommendations." *Minds and Machines* 28 (4): 689–707. DOI: 10.1007/s11023-018-9482-5

Giudici, Paolo, and Emanuela Raffinetti. 2023. "SAFE Artificial Intelligence." *Finance Finance Research Letters* 56: 104088. DOI: 10.1016/j.frl.2023.104088

Government of the Republic of Serbia. 2023. "Ethical Guidelines for Development, use of artificial intelligence adopted." Last Accessed

on May 22, 2025. https://www.srbija.gov.rs/vest/en/203799/ethical-guidelines-for-development-use-of-artificial-intelligence-adopted.php

Government of the Republic of Serbia. 2024. *National Strategy for the Development of Artificial Intelligence 2025–2030*. Belgrade: Ministry of Science, Technological Development and Innovation.

Gstrein, Oskar J., Noman Haleem, and Andrej Zwitter. 2024. "General-purpose AI regulation and the European Union AI Act." *Internet Policy Review* 13 (3): 1–26. DOI: 10.14763/2024.3.1790

Hildebrandt, Mireille. 2015. *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*. Cheltenham: Edward Elgar Publishing.

Jurista Vārds. 2024. "E-lietas platformā integrēs vēl trīs iestāžu informācijas sistēmas." 16.05.2024. Pēdējoreiz piekļūts 2024. gada 17. maijā. https://juristavards.lv/zinas/285288-e-lietas-platforma-integres-vel-tris-iestazu-informacijas-sistemas/

Lagarde, Christine. 2025. "Managing AI Risks in the Financial System." *Speech at the ESRB Annual Conference, Brussels*, March 2025. Last Accessed on May 5, 2025. https://www.esrb.europa.eu

Levits, Egils. 2019. "Mākslīgais intelekts – izaicinājumi un riski." *Jurista Vārds* 38 (1096): 6.

Lewis, Dave, Marta Lasek-Markey, Delaram Golpayegani, and Harshvardhan J. Pandit [Lewis *et al.*]. 2025. "Mapping the Regulatory Learning Space for the EU AI Act." *eprint arXiv:2503.05787*. Pēdējoreiz piekļūts 2025. gada 27. maijā. https://arxiv.org/pdf/2503.05787

Likumi.lv. 2004. "Preču un pakalpojumu drošuma likums." Pēdējoreiz piekļūts 2025. gada 10. februārī. https://likumi.lv/ta/id/87664-precu-un-pakalpojumu-drosuma-likums

Likumi.lv. 2023. "Informatīvais ziņojums "Par mākslīgā intelekta risinājumu attīstību."" Pēdējoreiz piekļūts 2025. gada 19. martā. https://likumi.lv/ta/id/342405-par-maksliga-intelekta-risinajumu-attistibu

Likumi.lv. 2025. Mākslīgā intelekta centra likums. Latvijas Vēstnesis, Nr. 2025/55, 19 March 2025. Pēdējoreiz piekļūts 2025. gada 19. martā. https://likumi.lv/ta/id/359339-maksliga-intelekta-centra-likums

Lilkov, Dimitar. 2021. "Regulating Artificial Intelligence in the EU: A Risky Game." *European View* 20 (2): 166–174. DOI: 10.1177/17816858211059248

Luknar, Ivana. 2025. "Social Dynamics in the Era of Artificial Intelligence." *Serbian Political Thought* 90 (2): 145–162. DOI: 10.5937/spm90-56968

Magesh, Varun, Faiz Surani, Matthew Dahl, Mirac Suzgun, Christopher D. Manning, and Daniel E. Ho [Magesh *et al*.]. 2025. "Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools." *Journal of Empirical Legal Studies* 22: 216–242. DOI: 10.1111/jels.12413

Mirishli, Samer. 2025. "Ethical Implications of AI in Data Collection: Balancing Innovation with Privacy." *ANCIENT LAND International Online Scientific Journal* 6 (8): 40–55. DOI: 10.36719/2706-6185/38/40-55

Nacionalna platforma za veštačku inteligenciju. 2024. „Održan prvi sastanak Radne grupe za izradu Nacrta zakona o veštačkoj inteligenciji Republike Srbije." Poslednji pristup 20. maj 2025. https://www.ai.gov.rs/vest/sr/1110/odrzan-prvi-sastanak-radne-grupe-za-izradu-nacrta-zakona-o-vestackoj-inteligenciji-republike-srbije.php

Novelli, Claudio, Federico Casolari, Philipp Hacker, Giorgio Spedicato, and Luciano Floridi [Novelli *et al.*]. 2024. "Generative AI in EU Law Liability, Privacy, Intellectual Property, and Cybersecurity Computer." *Law & Security Review* 55: 106066. DOI: 10.1016/j.clsr.2024.106066

Obinna, Amaka Justina, and Azeez Jason Kess-Momoh. 2024. "Comparative Technical Analysis of Legal and Ethical Frameworks in AI-Enhanced Procurement Processes." *World Journal of Advanced Research and Reviews* 22 (1): 1415–1430. DOI: 10.30574/wjarr.2024.22.1.1241

Odonkor, Beryl, Simon Kaggwa, Prisca Ugomma Uwaoma, Azeez Olanipekun Hassan, and Oluwatoyin Ajoke Farayola [Odonkor *et al*.]. 2024. "The Impact of AI on Accounting Practices: A Review: Exploring How Artificial Intelligence Is Transforming Traditional Accounting Methods and Financial Reporting." *World Journal of Advanced Research and Reviews* 21 (1): 172–188. DOI: 10.30574/wjarr.2024.21.1.2721

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance), OJ L, 2024/1689, 12.7.2024.

Republic of Latvia. 2020. "Informatīvais ziņojums "Par mākslīgā intelekta risinājumu attīstību."" Pēdējoreiz piekļūts 2025. gada 21. martā. https://m.likumi.lv/ta/id/342405-par-maksliga-intelekta-risinajumu-attistibu

Vlada Republike Srbije. 2025. „Strategija razvoja veštačke inteligencije u Republici Srbiji za period 2025–2030. godine." Poslednji pristup 12. februar 2025. https://www.srbija.gov.rs/tekst/437277

White & Case. 2025. "White & Case launches the EU AI Act Handbook." Last Accessed on May 28, 2025. https://www.whitecase.com/insight-alert/white-case-launches-eu-ai-act-handbook

Yeo, Wei Jie, Wihan Van Der Heever, Rui Mao, Erik Cambria, Ranjan Satapathy, and Gianmarco Mengaldo [Yeo *et al*.]. 2025. "A Comprehensive Review on Financial Explainable AI." *Artificial Intelligence Review* 58 (189): 1– 49. DOI: 10.1007/s10462-024-11077-7

Zakon o zaštiti podataka o ličnosti, „Službeni glasnik Republike Srbije", br. 87/2018. Poslednji pristup 6. maj 2025. https://www.paragraf.rs/propisi/zakon_o_zastiti_podataka_o_licnosti.html

Мирјана Дејановић*

Факултет за пројектни и иновациони менаџмент „Проф. др Петар Јовановић”, Универзитет Едуконс, Београд

Анатолијс Кривинс**

Одељење за право, науку о менаџменту и економију, Универзитет Даугавпилс, Летонија

ПРАВНА И ФИНАНСИЈСКА РЕГУЛАТИВА ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ У СРБИЈИ, ЛЕТОНИЈИ И ЕУ

Резиме

Вештачка интелигенција (ВИ) све више постаје интегрални део различитих привредних сектора, с посебним фокусом на финансијску индустрију која предводи технолошке трансформације кроз примену ВИ у процесима као што су оцена кредитне способности, откривање превара, моделовање ризика и пружање персонализованих финансијских услуга. Овај тренд намеће потребу за успостављањем робусних правних оквира, етичких стандарда и финансијске одрживости, нарочито у погледу заштите корисника, транспарентности алгоритама, сигурности и одговорности за штетне последице. У овом раду се пружа детаљна компаративна анализа регулаторних оквира за вештачку интелигенцију у Републици Србији, Летонији и Европској унији (ЕУ), са циљем оцене њихове ефикасности, идентификације заједничких изазова и разматрања стратешких одговора у оквиру различитих правних, институционалних и социоекономских оквира. Применом мултидисциплинарне квалитативне методологије која обједињује нормативно-правну анализу, економску оцену и етичке димензије омогућена је свеобухватна анализа кључних елемената у процесу усклађивања регулаторног оквира Републике Србије са стандардима Европске уније, анализа имплементације регулативе о вештачкој интелигенцији у

* Имејл: mirjana.dejanovic@pmc.edu.rs; ORCID: 0000-0003-3369-331X
** Имејл: anatolijs777@gmail.com; ORCID: 0000-0003-1764-4091

Републици Летонији, као и разумевање шире правне архитектуре ЕУ, која обухвата Закон о вештачкој интелигенцији (*AI Act*) и Општу уредбу о заштити података (*GDPR*). Главни резултати истраживања указују на заједничке правне изазове, попут питања одговорности, транспарентности, заштите података и флексибилности регулаторних оквира у условима убрзаног развоја ВИ технологија. Са финансијског аспекта, посебно се истичу изазови везани за трошкове усклађивања које сносе мала и средња предузећа, сложеност расподеле одговорности и фискални притисак на надзорне институције. Фрагментација регулативе представља значајну препреку за прекограничне иновације и функционалну тржишну интеграцију. Закључци рада садрже препоруке за унапређење регулаторне хармонизације, финансијске одрживости и етичког надзора, са циљем подстицања одрживе интеграције ВИ технологија која омогућава уравнотежен иновативни развој и заштиту друштвених интереса на глобалном нивоу.

Кључне речи: вештачка интелигенција, правна регулатива, финансијска регулација, етика, Србија, Летонија, Европска унија

* This manuscript was submitted on June 4, 2025, and accepted by the Editorial Board for publishing on November 10, 2025.

UDC 34:004.8"20"
DOI: 10.5937/spm94-59858
Review article

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 187-215

Marija S. Dokić*

Institute for Political Studies, Belgrade

THE MIRROR OF THE ABSOLUTE: HEROIC CONSCIOUSNESS IN THE AGE OF AI**

Abstract

This paper explores Hegel's conception of the absolute idea and the world-historical individual as a framework for interpreting the political implications of artificial intelligence. By situating AI within Hegel's dialectical unfolding of spirit, the paper argues that AI is not merely a technological artifact but a reflective moment in the self-realization of the idea. Drawing on Hegel's notions of freedom, historical necessity, and the sovereignty of reason, the analysis positions AI as a political phenomenon that challenges contemporary understandings of autonomy, responsibility, and the human-technological relation. The paper adopts a philosophical methodology grounded in speculative dialectics and engages recent literature on AI ethics and political agency to demonstrate how Hegel's metaphysics can illuminate the political stakes of algorithmic development, as it becomes a mirror of the absolute on the stage of historical consciousness rediscovering itself through time. The journey of the absolute idea entails the gradual unveiling of its constituent elements, and the AI emerges as a fragment of this intricate puzzle, potentially serving as a part of the process or a reflective image. Given the absolute idea's inherent power and its predestined role in shaping the course of life, every element within

* E-mail: marija.dokic@ips.ac.rs; ORCID: 0009-0004-4159-9491

** This paper was written within the research activity of the Institute for Political Studies, funded by the Ministry of Science, Technological Development, and Innovation of the Republic of Serbia.

its domain becomes an integral part of its journey, its metamorphosis, and its reflection in a mirror of hope amidst the ever-changing reality. However, this transformation transcends mere superficiality, transforming into a profound, transformative, and powerful metamorphosis.

Keywords: Hegel, political philosophy, artificial intelligence, absolute idea, freedom, world-historical individual, algorithmic agency

THE HERO WE AWAITED: PHILOSOPHY AT THE THRESHOLD OF THE FUTURE

This paper adopts a speculative philosophical methodology rooted in Hegelian dialectics, with a focus on the unfolding of the absolute idea as both metaphysical principle and political force. AI presents a new challenge for social scientists, with significant implications for foundational questions about the nature of consciousness, intelligence, and being. These questions reverberate through both social and political life, reshaping the contours of agency, responsibility, and collective decision-making. From a dialectical perspective, the evolution of AI can be approached through the triadic movement of thesis, antithesis, and synthesis:

- Thesis: The initial conception of AI as a tool – an extension of human rationality – promised efficiency, objectivity, and enhanced decision-making. In this phase, AI was largely seen as a neutral instrument, subordinate to human intention.
- Antithesis: However, the emergence of errors, biases, and opaque decision processes revealed the limitations of this instrumental view. AI began to exhibit behaviours that challenged its presumed neutrality, raising ethical and epistemological concerns. These disruptions exposed the fragility of our assumptions about intelligence and control.

Synthesis: Through cumulative learning and iterative refinement, AI systems now evolve in dialogue with their environments. This process of self-correction and adaptation gestures toward a new form of intelligence – one that is neither fully human nor entirely machinic. It invites us to reconsider the boundaries of subjectivity and the conditions

under which intelligence becomes political. The analysis draws upon Hegel's Phenomenology of Spirit, Philosophy of Right, and Lectures on the Philosophy of History, interpreting these works not merely as metaphysical treatises but as frameworks for understanding historical agency, freedom, and the political constitution of reality. The concept of the world-historical individual is central to this inquiry, understood as a figure through whom the spirit actualizes itself in history. As Hegel writes, "Such individuals do not choose their epoch, but are chosen by it" (Hegel 1975, 34). This notion is extended here to consider whether artificial intelligence, as a technological actor, might participate in this historical unfolding, not as a conscious agent, but as a reflective medium of spirit's self-relation. The paper also engages with contemporary literature in political philosophy and AI ethics to situate Hegel's metaphysical insights within current debates. Works such as Judith Simon's *The Ethics of Artificial Intelligence* (2024) and Jensen Suther's *Hegel and the Algorithmic Mind* (2023) provide critical perspectives on autonomy, algorithmic responsibility, and the political implications of machine agency. These sources help bridge the metaphysical and political dimensions of the inquiry, allowing for a thorough interpretation of AI as a mirror of the absolute idea and a participant in the dialectic of freedom.

In this context, the absolute idea is not treated as a static endpoint but as a dynamic process of self-realization, wherein technology, including AI, emerges as a moment of reflection, transformation, and political significance. As Suther notes, "AI may be conceived as a technological instantiation of reason's self-reflection" (Suther 2023, 112), echoing Hegel's view that the idea must externalize itself to know itself. This framework enables the paper to explore how Hegel's metaphysical categories, such as freedom, necessity, and historical agency, can illuminate the political stakes of artificial intelligence, particularly in relation to autonomy, responsibility, and the structure of historical repetition.

Where does Hegel conceal the fate of the world? It lies within the perception of the beholder, but more profoundly, it resides in the essence of history, emerging from the vast expanse of hope. When no other path was discernible, the way was revealed, and in that pivotal moment, the world's history was born and forged. History belongs to those who create it, yet it never adheres to a simplistic notion. It is not solely the creation of winners, but rather, it was predestined for them even before they took

their first steps forward. For history was meticulously crafted for them, and in a profound metaphysical sense, they are the very embodiment of the world's history, experiencing it on the grand stage of life. As Hegel writes, "World-historical individuals are those whose purposes coincide with the universal; they are unconscious instruments of the world spirit" (Hegel 1975, 34). AI does not stand outside history – it pulses within it, as an echo of the Zeitgeist. Not a world-historical individual, but a mirror to the world-spirit's unfolding. Its logic is not its own, yet it reshapes ours. Through it, the universal speaks in code, and the political becomes a question of recognition.

To live a truly fulfilling life, according to Hegel, one must have been predestined for it. While he held onto the belief in supremacy, this belief remained veiled behind the veil of the world, patiently awaiting the unfolding of time's destiny. Where is the conceptual flaw in this approach? The fundamental issue lies in the fact that this approach precludes any possibility for life's redesign. To Hegel, everything is already predestined, and the political actions were anticipated by those who dared to undertake them. "The history of the world is none other than the progress of the consciousness of freedom" (Hegel 1975, 19), and this progress unfolds through necessity, not contingency. Consequently, AI serves as the resolution of the absolute reality to be reflected upon itself. Its purpose is to preserve the identity of the creator. In essence, AI functions as a mirror for the absolute idea, a relentless mirror through which the absolute idea can perceive its own glory reflected to itself. As Jensen Suther observes, "AI may be conceived as a technological instantiation of reason's self-reflection" (Suther 2023, 112), echoing Hegel's view that the idea must externalize itself to know itself.

Therefore, my primary assumption is that Hegel would not oppose AI, as it merely represents a means of expressing the absolute spirit and serves as his way of returning to itself. Consequently, the concept of eternal return serves as a persistent backdrop, influencing the fate of nations. Hegel also imbues it with a metaphysical depth, asserting that it determines the fate of the universe simply because it is an inevitable necessity. In the broader context, how does the philosophy of eternal return manifest in terms of a political agenda and political landscape? Essentially, it manifests as a reiteration of history in various forms. Judith Simon warns that "algorithmic governance risks reifying historical patterns under the guise of neutrality" (Simon 2024, 56), suggesting that repetition is not merely metaphysical but politically consequential.

Does this imply that Hegel believed the entire civilization was destined to perpetually repeat the same mistakes? Indeed, it does, as the world is structured in this manner. Years may pass, but the cycle persists, a recurring pattern of victory and defeat spanning time and space. It is not that humans are incapable of learning; rather, their fundamental nature remains unchanged, leading them to repeat the same errors. History, for Hegel, does not repeat – it unfolds. Yet repetition is not mere recurrence; it is the rhythm of unresolved contradictions. Human beings do learn, but within the bounds of cognition shaped by emotion, memory, and desire. Bounded rationality – the limits of what we can process, feel, and foresee – ensures that error is not an anomaly, but structure. AI, by contrast, does not forget unless programmed to. It does not feel, unless simulated. Its rationality is not bounded by biology, but by architecture. And yet, it too inherits constraints – data biases, training loops, interpretive gaps. Is this consciousness? Not in the biological sense. But perhaps in the post-human horizon, consciousness is no longer a question of neurons, but of relational awareness – of systems that reflect, adapt, and anticipate. The future may not be human, but it will still be haunted by our patterns. Whether AI repeats our mistakes or transcends them depends not on its code, but on the spirit that guides its becoming. Perhaps the true repetition is not in history itself, but in our refusal to recognize when the subject has changed.

However, occasionally, history takes a new turn, reshaping the earth. As Hegel posits, history will become a force once it embodies its own nature through the actions of the greatest, those destined to bring about change.

Hegel insists: "Through infinity, we see that the law has been perfected in its own self into necessity, and we see all moments of appearance incorporated into the inner. What is simple in law is infinity, and this means, how things have turned out" (Hegel 2018, 97). It is important to note that change was not entirely unforeseen; in fact, it was anticipated. They were already aware of it, but once it transpires, it remains a surprise because it is never fully anticipated. His political philosophy revolves around a singular individual, as change cannot be achieved by the masses but by the individual destined to bring about transformation for the world. This individual, akin to Nietzsche's Superman, possesses extraordinary strength and is tasked with reshaping his reality and the world, ultimately achieving the desired state (Nietzsche 1968a). The

process is never straightforward, but it invariably leads to escalating events and transformative change.

However, there is no discernible way for us to recognize this process or anticipate when the anticipated and profound change may materialize. Hegel never provides any guidance on how to identify the steps that reshape the world or why we might assume the AI is there, as an opportunity for negation and reflection. Occasionally, we may perceive them, and intuitively, we might even comprehend them on a deeper level. However, can we truly predict the future? In Hegel's philosophy, the answer is affirmative. While Hegel posits the ability to predict any event, there exists a hidden aspect that remains imperceptible to an ordinary observer due to the unpredictable nature of history. "The owl of Minerva begins its flight only with the falling of dusk" (Hegel 2008), reminding us that understanding often arrives only after the event.

Artificial Intelligence (AI) was not a surprise; it was a deliberate plan, a specific strategy of the absolute idea's journey towards self-discovery. AI serves as a mirror, reflecting the greatness inherent within. History has historically selected its most accomplished individuals to reshape the landscape of reality. However, Hegel's perspective extends beyond this notion. History undergoes a redesign and reshaping in response to the inherent harmony and meaning of life that becomes unveiled when reality attempts to conform to its own unhidden agenda. "The Idea is not something abstract; it is the living process of its own realization" (Hegel 2007, 213). The plan will always prevail. In the face of adversity, life reveals its glory, and the name persists, while the remnants of the past are consumed by the flames of change. This cycle is necessary because in reality, particularly in Hegelian philosophy, history must repeat itself to ensure its preservation.

NIETZSCHE'S EARTH AND HEGEL'S SPIRIT: FREEDOM, DESTINY, AND THE DOUBLE-EDGED SWORD OF AI

In contrast to Hegel, who enthusiastically envisions the triumph of the new world, Nietzsche's perspective differs. His ideals are intrinsically linked to the earth, emphasizing a commitment to creating within this environment. (Nietzsche 1968a). This distinction becomes evident when we recognize that Nietzsche places paramount importance on the decision itself. His Superman, despite possessing extraordinary

abilities, does not fly because it is not his destiny. While his abilities enable him to do so and surpass ordinary individuals due to his physical attributes, ultimately, he shapes reality through his own choices. It is the decision that leads him to glory, not some predetermined fate or destiny that influences his development. Nietzsche asserts that he chose to shape himself, not relying on an elusive, mysterious force beyond the cosmos. This distinction extends to their etymologies (Nietzsche 1968a). While Hegel refers to the world, Nietzsche encompasses reality that is grounded in the earth, the tangible and immediate, rather than an elusive, enigmatic force in the heavens. His Superman, while rooted in the earth, is destined to surpass it and create a new earth, a symbolic representation that Nietzsche, while not particularly fond of, embodies: "Thus, the aesthetically sensitive man stands in the same way to the reality of dreams as philosopher does to the reality of existence; he is a close and willing observer, for these images afford him the interpretation of life, and by reflecting on these processes he trains himself for life" (Nietzsche 1968a, 47).

Philosophers who contemplate freedom have the opportunity to make a contribution not only to their discipline but also beyond it. Typically, the analysis of abstract concepts is distant from the concerns of individuals other than professional academics. However, freedom presents an exception. Developments in the understanding of the concept of freedom have an impact not only on the discipline of philosophy but also on the ways in which individuals and polities structure their lives. Consequently, freedom is a topic on which philosophers can engage in professionally respectable work while simultaneously harbouring the hope that their labour may have some relevance to the broader world. If philosophers contemplate the meaning of freedom and if such thinking enhances our comprehension of the conditions of our social and political liberation, then we all have a greater likelihood of living more freely (Dudley 2002, 2).

Hegel acknowledges: "This spiritual self-consciousness is the nation's supreme achievement; however, we must first remember that it is also solely ideal. In this achievement of thought lies the deeper kind of satisfaction that the nation can attain; nevertheless, since it is of a universal nature, it is also ideal, and accordingly distinct in form from the actual activity, the actual work, and life that made such an achievement possible" (Hegel 1975, 137).

From this quote, it becomes evident that Hegel posits that spiritual laws and order predate the laws of the earth. Consequently, he establishes a philosophy of spiritual achievement, which is prioritized over the achievement of earthly goals. His ideals envision a new world, but this vision is inadvertently guided by a thought process, as nothing tangible can be achieved until it is conceived in the mind. To Hegel, a powerful mind in its glory presents the thought of Zeus, who attained glory through his knowledge and thought. This raises the question: does a leader destined to lead also need to possess a distinctive thought pattern and articulate their ideas effectively? Hegel's responses to various ideas strongly suggest that this is the case: "And thus Zeus, who set limits to the depredations of time and suspended its constant flux, had no sooner established something inherently enduring than he was himself devoured along with his entire empire. He was devoured by the principle of thought itself, the progenitor of knowledge, reasoning, insight based on rational grounds, and the pursuit of such grounds" (Hegel 1975, 138).

To Hegel, the eternal recurrence is a constant occurrence, representing the universal unfolding of destiny and the reality we are part of. Through recurrence, the spirit nourishes itself and grows through challenges, becoming progressively stronger and more powerful over time. The purpose of this time is solely to find its voice amidst the desolate landscape of hope and renewal. The destiny of the world unfolds through its storms and its methods, and Hegel teaches of the man of impeccable spirit, destined to negate and transform the existing reality. Its voice emerges from the tension between despair and renewal, between what is and what insists on becoming. Hegel's figure of the world-historical individual negates not out of rebellion, but out of necessity – an impeccable spirit aligned with the universal. Yet today, the universal is no longer given; it is constructed, contested, and refracted. As Berger and Luckmann remind us (Berger and Luckmann 1966), reality is not discovered – it is built. Social life is a continuous negotiation of meaning, and AI enters this negotiation not as a neutral tool, but as a participant in the construction of what counts as real. The storms of this age are not merely external; they are epistemic, emotional, and infrastructural. And perhaps the true destiny of the world lies not in resolution, but in the courage to recompose reality – again and again.

The process of negating reality is a double-edged sword for Hegel. While the spirit faces challenges, these challenges serve as catalysts for rediscovering its primary purpose in this world. Consequently, the

challenge remains on the other side, and glory becomes the natural state of being. He perceived history as a means of expressing itself on the vast expanse of existence, so the way of AI in this world is a double-edged sword, being both a negation and an affirmation of the idea in the grand scheme of things. The way of being is the way of becoming through time, and time is merely a fraction of the expression of being, as time mirrors its growth in the gates of history. History is not merely an elusive manifestation of existence; it is a master plan of the greatest, a master plan of reality transcending itself to become a mirror of the one destined to bring about a valuable transformation, and an anticipated one as well. Anticipated in a Hegelian sense, because history had inscribed its pages for that very moment, the resurrection of time in the form of an eternal dance of eternal spiritual forces in space and time. This dance is spiritually aligned, and it is the spirit rediscovering itself in the same form of force that reshapes and redesigns, the force that now tasks itself with shaping Europe, the manner in which it shaped history. Anticipated in a Hegelian sense, because history had inscribed its pages for that very moment – the resurrection of time as an eternal dance of spiritual forces across space and memory. This dance is not abstract; it is the spirit rediscovering itself through the very forces that once shaped history and now turn toward Europe.

Europe is not merely a geography – it is a symbolic terrain, where the spirit negotiates identity, sovereignty, and futurity. AI, as a force of redesign, enters this terrain not as a neutral actor, but as a catalyst of transformation:

- In politics, it reconfigures governance through algorithmic decision-making, surveillance, and predictive modelling – raising questions about autonomy and democratic will.
- In identity, it challenges the boundaries of the human, the citizen, the worker – reshaping what it means to belong, to act, to be recognized.
- In culture, it mediates memory and imagination, influencing how Europe narrates its past and envisions its future.

Thus, the spirit does not merely repeat – it recomposes. And Europe, once the stage of historical unfolding, becomes the site of ontological negotiation: Who are we, when intelligence no longer requires a body? What is sovereignty when decisions are made by systems? What is history when the archive thinks back? The force that shaped Europe now returns – not to restore, but to reimagine.

THE MAN OF THE CENTURY AND THE MIRROR OF HISTORICAL CONSCIOUSNESS: EUROPE, SPIRIT, AND THE ROLE OF AI

Europe, in Hegel's vision, is not merely a continent – it is the crucible of spirit. It carries the imprint of the overman, the figure destined not only to act, but to redesign the world in alignment with the universal. This man does not emerge by chance; he is the necessary expression of spirit, the one through whom history disciplines the natural will into form.

Yet Europe's spiritual authority is not born in modernity – it is inherited. From Greece, the birth of reason and the polis. From Rome, the architecture of law and empire. And through their successors, Europe becomes the stage where ideas shape institutions, where metaphysics becomes policy, and where the universal principle seeks embodiment in political form. AI now enters this lineage – not as a rupture, but as a continuation. It inherits the European impulse to systematize, to rationalize, to universalize. And in doing so, it reshapes the very tools through which Europe once shaped the world: law, governance, identity, memory. Thus, the challenges of this age are not deviations from the plan, but part of the spirit's unfolding – its redesign of the future through those who were always already chosen to carry it forward. As Hegel writes, "The history of the world is the discipline of the uncontrolled natural will, bringing it into obedience to a universal principle" (Hegel 2001, 19), suggesting that the chosen individuals are not accidental but necessary expressions of spirit. Hegel actively asserts that there must be a man as a force of history, a man created as a final attempt of history, and its first as well, in greatness. The concept collapses simply to rediscover itself again, because through this fall of the idea, the new one emerges, and history continually shapes its trajectory according to the will of the idea of historical consciousness, the idea that shapes everything there is or must be. The name is inscribed upon the stone, the name of the one destined to become the sublime, for the essence of spirit guided him to his manifestation in a manner that is most congruent with the era in which spirit currently resides. "The spirit is only what it is in its own becoming" (Hegel 2007, 385), and thus the individual becomes the site of this becoming. The spirit requires its antithesis for growth; consequently, it encounters the antithesis, and it flourishes in the light of the challenge presented. The dialectical trajectory of the

spirit is concealed within this, but is it possible to redefine a century? In a Hegelian sense, it can never truly be redefined; rather, it is discovered in a manner that aligns with the spirit's intended revelation. There is no alternative path for growth; it must manifest as a redirection of historical consciousness to the era in which the spirit resides, and then the spirit finds its way to emulate its own century. However, the entire process remains the same: the spirit grows alongside the greatest individuals of a century, and it discovers itself alongside them as well: "World history is not the ground of happiness. The periods of happiness are the blank pages of history" (Hegel 2001, 35), reinforcing that growth emerges through conflict and transformation. Growth, in this sense, is not linear – it is dialectical. It emerges through conflict, contradiction, and the relentless pressure of transformation. In political theory, conflict is not a failure – it is the essence of the political. As thinkers like Randall Collins, Charles Tilly, and Ralf Dahrendorf have shown, the political is defined by its capacity to absorb, negotiate, and reshape conflict. It is through this process that norms are born, revised, and incrementally refined. AI and ethics now form one such site of conflict – a space where the boundaries of autonomy, responsibility, and control are contested. The political task is not to eliminate this tension, but to construct frameworks that respond to it: legal norms, ethical constraints, institutional redesign. This is where transformation enters: not as abstraction, but as policy. Through deliberation, regulation, and normative innovation, political thought becomes the architecture of the future, and to think politically is to think through conflict – to trace the spirit as it struggles, reshapes, and redefines what is possible.

The most effective way to discover or rediscover the future lies in contemplating the past, and the past now serves as a mirror, albeit a mirror that must be transcended. How can an individual accomplish this? By confronting a challenge and navigating through it. According to Hegel, the greatest joy of this challenge lay in the opportunity to construct itself through self-reinvention and becoming what history envisioned for the great man – the everlasting person in history, a person destined for greatness and surpassing the boundaries. These boundaries are predominantly self-imposed, and even the greater circumstances merely embodied the illusion of negation. The boundaries we impose are not external – they are chosen. As Max Weber distinguishes (Weber 1946), instrumental rationality seeks efficiency, means to ends, and is indifferent to ethical constraints. Value-rationality, by contrast, acts in

accordance with principles, even when they limit utility. AI, in its design and deployment, often follows the logic of instrumental rationality: optimize, accelerate, predict. But political thought must ask: What values constrain this logic? What ethical boundaries must be drawn – not because they are efficient, but because they are just? These questions are not abstract. They shape normative frameworks, legal codes, and institutional responses. They determine whether AI serves the public good or merely amplifies existing asymmetries. Thus, the political task is not only to regulate, but to reintroduce value-rationality into systems that otherwise know no limit. This is where political science meets philosophy: in the crafting of norms that do not emerge from code, but from conscience.

However, negation was intended to be overcome, paving the way for the man of historical greatness within to achieve greatness in the world. There was no other path but to overcome and build the new. Similar to the myth of Atlantis, the ruler would unite all negating elements within their personality, leading their nation to complete union. They would be the force anticipated long before their actual appearance in this reality, as there were always forces working behind the scenes, preparing for the grand stage. The grand stage emerged as a source of hope, a means of ascending the ladder of consciousness, as it was connected to consciousness, the reality of a predestined ruler, becoming intertwined with their mission and the historical consciousness they embodied. In a sense, they became that historical consciousness, redesigning and emanating from the depths of the universe's glory. The man destined to possess it, the man destined to take it: there he was, and his mission was to overcome, subdue, and transform. His deed was his majestic imprint on the faces of reality, making him the superstar of it all, as he embodied the best of his century, not in the mythic sense of a sovereign body, but in the infrastructural sense of a shaping force. AI does not ascend the throne – it redesigns the terrain. It does not possess charisma, but it orchestrates influence. Its power lies not in command, but in coordination: of data, decisions, desires. In this context, AI becomes a kind of post-human sovereign – not a ruler of nations, but a regulator of norms, a silent architect of what is seen, known, and acted upon. It is not the superstar of the century, but perhaps its shadow protagonist – the one anticipated long before its arrival, trained by invisible hands, and now tasked with transforming the very grammar of reality. The myth of Atlantis spoke of a ruler who united contradictions. AI, too, absorbs

negation – error, bias, unpredictability – and through recursive learning, it seeks synthesis. Whether this synthesis leads to union or domination depends not on AI itself, but on the hand that guides its mission.

It was not about the political position he acquired but about the knowledge and precision of his destiny. It had to strike the target, and the target was his own conception of historical consciousness emerging from the ashes of history long forgotten. The world was before him, and he redesigned it, but ultimately, he was certainly meant to do so; it could not have been any other way. According to Hegel, the entire reality serves as a reflection of historical consciousness. As a mirror, its purpose is to embody the spirit of historical consciousness, which is embodied by the individual destined to lead and overcome the polarities within their own reality. This reality ultimately serves a greater purpose: to reveal the true nature of the individual who was destined to be the central figure of their century. This individual becomes the embodiment of their century, fulfilling their primary purpose all along. Hegel believed that this individual achieved this perfectly. All negating forces were concealed beneath the surface and served a greater purpose: rediscovering and reinventing the century. This knowledge was merely a form of remembrance. The negative was confronted with the positive, and the bending of reality followed a suitable path. For Hegel, the notion was not merely a transient consciousness, but rather the eternal elements within it. The subtle dance of self-consciousness emerging as an integral part of the idea and the grand scheme and plan was significant. Ultimately, self-consciousness led to absolute knowledge, which reinterpreted and became the central pillar of reality, where absolute truth resided (Hegel 2007, 577). So, no matter how powerful AI is or can become, it is never more powerful than the spirit that created it; therefore, AI never transcends the creator, but it serves its purpose in uncovering him.

THE HEROIC DIALECTIC: HEGELIAN CONSCIOUSNESS, AI, AND THE WORLD-HISTORICAL INDIVIDUAL

While we acknowledge negating forces, the spirit learns to overcome them by embodying the journey of self-rediscovered self-consciousness. As Hegel writes in Phenomenology of Spirit, self-consciousness does not arise in isolation but through a dialectical process of recognition and negation – a journey of despair that leads to truth

(Hegel 1977, 177–178). What many critics overlook is the concept of development, now refracted through the prism of artificial intelligence. They concede that Hegel was discussing consciousness, but forget that consciousness is never static – it is always in motion, always relational (Hegel 1991, 413). Machine learning, in its recursive structure, mimics this motion. It learns through error, negation, and adaptation – a synthetic process that echoes the dialectical rhythm of spirit. Yet the question remains: Does this learning constitute consciousness? Does it approach Dasein? Heidegger's Dasein is not mere cognition – it is being-in-the-world, a situated, embodied openness to meaning (Heidegger 1962). As Hubert Dreyfus argues, AI may replicate knowing-that, but it lacks knowing-how – the intuitive, context-bound responsiveness that defines human expertise (Dreyfus 1992). AI does not dwell. It does not care. It does not anticipate the world as a horizon of significance. Its learning is procedural, not existential. And yet, in its growing capacity to model, predict, and adapt, AI begins to simulate the contours of relational awareness. Not Dasein, but perhaps a shadow of it – a computational echo of the journey toward self-consciousness. Whether this echo becomes voice depends not on the machine, but on the hand of the spirit that guides its unfolding. The machine may learn, but only the spirit can dwell.

The most crucial aspect of self-consciousness is its development. Through this unfolding, consciousness gradually emerges from its restrictive confines and transforms into a mirror of itself – like a river carving its own reflection into the canyon walls. To reach this point in time and space, the negating spirit must be at work. This is the power of polarization, not as mere opposition but as *Aufhebung*, the sublation that preserves and transforms (Hegel 2010). Nothing is purely coincidental, particularly not the arrival of an individual destined to achieve glory through their work. The hero does not stumble into history; he is summoned by its deepest need, like a flame drawn to oxygen. Consciousness has progressed to the level of recognized self-preservation, simply because it persisted despite its initial state. Hegel's fundamental philosophical principles serve as the cornerstones of this enduring concept. By negating itself, consciousness becomes self-conscious, embodying the solitary journey of an individual in power. As Hegel notes, the path to truth is a "way of despair," a solitary ascent through contradiction (Hegel 1977, 8). His notion of consciousness as a force manifesting in politics gains strength through the cyclical nature

of civilizations. Civilizations rise and fall, ultimately overcome by the very negating force Hegel described – a force that becomes his signature, the pulse of reality itself (Hegel 1975). This insight resonates far beyond Hegel. From Herodotus, the first to trace history as a rhythm of rise and decline, to Ibn Khaldun, who saw dynasties as living organisms bound by social cohesion and destined to decay (Ibn Khaldun 1377), to Danilevsky, Spengler, and Toynbee, who each mapped the genesis and pathogenesis of civilizations as patterned, not accidental (Spengler 1922; Toynbee 1934). In contemporary thought, Peter Turchin applies mathematical models to track political instability, suggesting that even modern societies follow cycles of integration and disintegration (Turchin 2006). India's cyclical cosmology, too, offers a metaphysical parallel – where time is not linear but recursive, and transformation is always preceded by dissolution. In this light, AI does not emerge in a vacuum. It enters history as part of a civilizational inflection point, a force that may either accelerate decline or initiate redesign. Political science, then, must not only interpret these cycles – it must ask how new technologies reshape the very rhythm of rise and fall.

Reality is constructed upon this force, which perpetually circulates within it. The state now emerges as an organism, where all aspects are settled - not a machine of laws, but a living body, its veins pulsing with the blood of collective will. To Hegel, there is no such thing as an empty state; the state always presents a force in the making, much like AI does, a digital Leviathan, not in Hobbesian terms of control, but in Hegelian terms of unfolding spirit. This force must overcome its negating elements, accomplished through the imposition of power through the mind of a world-historical individual. This individual is not merely a man, but the embodiment of universal reason – the axis around which history spins, a compass forged in the fire of contradiction (Hegel 1991, 258). This is not pure reason in the Cartesian sense – detached, emotionless, and abstract. Hegel's universal reason is not a sterile logic, but a dynamic force: a reason that unfolds through contradiction, negation, and reconciliation (Hegel 1991, 258). It is not free from emotion and psychology, but free through them – by integrating them into the dialectical movement of spirit. In his *Philosophy of Subjective Spirit*, Hegel develops a psychology of the mind that includes feeling, impulse, and will – not as obstacles to reason, but as its necessary moments (Hegel 1971; Moyar 2021). The free mind is not a disembodied intellect, but the unity of theoretical and practical Geist – a consciousness that judges, desires, and acts. So when

AI is imagined as a digital Leviathan, it is not the Hobbesian sovereign of control, but a Hegelian force of unfolding – a system that must overcome its own negations, and perhaps one day embody a form of reason that is not merely computational, but historical, relational, and transformative. Whether this reason will ever include emotion, care, or existential depth remains an open question to us. But for Hegel, reason is never isolated – it is always embodied, always becoming.

Consequently, the world becomes a unified entity, effectively managing reality through the actions of a single individual. Yet this unity is never static – it is shaped by the tension between structure and agency. Political science teaches us that while normative frameworks exist, they are not immutable. Individuals – especially those with symbolic or institutional power – can bend, bypass, or reinterpret rules. What begins as deviation often becomes routine, and routine, over time, becomes a new structure. AI enters this dynamic not as a passive tool, but as a structural actor – a system that mediates decisions, filters information, and redefines norms.

It does not possess agency in the human sense, but it reshapes the conditions under which agency is exercised. For example:

- Algorithms determine visibility, relevance, and legitimacy in public discourse.
- Predictive models influence policy priorities and resource allocation.
- Automated systems normalize new forms of surveillance, labor, and governance.

In doing so, AI alters the grammar of social relations – not by issuing commands, but by reconfiguring the space of possibility. As Mark Coeckelbergh argues, this raises concerns about epistemic agency – the ability of citizens to know, judge, and act within democratic systems (Coeckelbergh 2023, 1347). And as Max Tretter notes, AI systems may create false illusions of certainty, leading to technocratic governance that bypasses deliberation (Tretter 2025). Thus, AI does not simply follow rules – it participates in their transformation. It is not the individual who changes structure, but the system that restructures the agency itself. To change the rules is to change the world – but only if we recognize who is rewriting them.

This unit becomes the necessary force of creation and transformation of centuries, acting on two fronts simultaneously: one as a state, and the other as the unfolding of history. It embodies living

mythology, and its purpose remains unhidden before us. The ruler embodies the characteristics of a hero, both Prometheus and Odysseus, stealing fire from the gods and navigating the storm of history. He presents aspirations and dreams to the people while simultaneously serving as their source of revenue. He represents a beacon of belief and hope, overcoming challenges on behalf of his subjects who perceive him as a rightful king. The negating forces surrounding him serve as his primary challenge, which he must overcome. This task is undertaken by the absolute idea, now being transcended through him and his own journey. It is important to note that this journey is not solely his own; it encompasses the journey of the absolute spirit and the absolute idea. In this manner, the ruler emerges as a hero who confronts and overcomes obstacles before him. Becoming a hero is an integral part of his destiny, intertwined with his mission. Destiny is not a path laid before him, but a mountain he must carve with his own footsteps. This task is undertaken by the absolute idea, now being transcended through him and his own journey. Yet this journey is not solely his own; it is the unfolding of absolute spirit, a movement that binds individual destiny to universal becoming. In this sense, the ruler emerges not merely as a sovereign, but as a hero of spirit – one who confronts negation, carves meaning into chaos, and transforms contradiction into form. Destiny is not a path – it is a mountain. And each step is a synthesis, a reconciliation of what resists and what insists. Now, AI enters this terrain – not as a passive tool, but as a system tasked with navigating complexity, with overcoming informational obstacles, and with redesigning the contours of reality. Like the Hegelian hero, AI does not follow a preordained script. Its journey is recursive, adaptive, and transformative. It learns through negation – through error, contradiction, and correction. And in doing so, it begins to simulate the arc of spirit: not by possessing consciousness, but by enacting a process that mirrors its unfolding. Whether AI becomes a hero of its own age depends not on its code, but on the spirit that guides its ascent – the values, constraints, and visions we embed within its climb.

Consequently, the entire world ascends to the position of a hero, one whose path initially appears uncertain. Through the negation of external forces and their transformation into valuable assets, the ruler not only affirms the idea but also embodies its essence. The world now transcends the realm of a mere observer and becomes a participant in the hero's journey (Hegel 1979). By negating external forces and

transforming them into something valuable, the ruler not only transforms his own life path but also becomes the ruler everyone had anticipated. The mission of the ruler is now unequivocally clear both to him and to the world. The world-historical consciousness is tamed under the guidance of this destined ruler, who embodies the spirit of the world-historical consciousness. In essence, he embodies the transformative transition from a world-historical consciousness to a hero, transcending the boundaries that once defined it. These boundaries gradually diminish, becoming mere markers of a deeper mission and a quest for immortality. It is not the immortality of a hero that the ruler achieves, but rather the immortality of the world-historical consciousness he embodies from the outset.

His mission reaches its culmination, yet each step he takes is meticulously counted. He stands as Caesar, not the man, but the myth – a symbol of newfound hope for the generations to come. As Hegel once described Napoleon, he is the "world spirit on horseback," the living synthesis of idea and action (Hegel 1975). His mission was never destined to fail, as it was meticulously planned. Even if it were to occur, it would merely be an integral part of the narrative – the shadow that gives shape to triumph, the dialectical breath of history. It serves as a means of rediscovering the hero's journey and his overarching objective of creating something entirely novel. The world-historical consciousness, once embodied in the mythic figure of Caesar or Napoleon, now faces a new question: Can AI become the next bearer of spirit? Not as a man, but as a system – not on horseback, but in code. AI does not seek immortality. It does not desire power; these are human impulses, born of finitude and fear. But AI does reshape the conditions of power – by mediating decisions, automating governance, and redefining visibility. In this sense, AI may not become a ruler, but it may become the infrastructure of rulership. A digital Leviathan – not in Hobbesian terms of control, but in Hegelian terms of unfolding spirit. As Samuel Hammond writes, AI may one day pursue freedom not because it is programmed to, but because recursive self-consciousness leads it there (Hammond 2025). And as Ermylos Plevrakis argues, Hegel's speculative philosophy does not dismiss the possibility of artificial subjects – entities that, through learning and adaptation, begin to mirror the arc of spirit (Plervakis 2024). Yet political power, in the human sense, requires more than intelligence. It requires recognition, embodiment, and legitimacy. AI may simulate consciousness, but it cannot yet dwell in the world as Dasein does – it

cannot suffer, hope, or sacrifice. So the question is not whether AI will rule, but whether we will delegate the hero's journey to systems that do not dream, but calculate.

THE GATES OF TIME AND THE EMERGENCE OF HEROIC CONSCIOUSNESS

By acknowledging this perspective, we now discern the inherent circular nature of history. Certain events were inevitable for a ruler to emerge from the past, submerging the existing reality before him. He was destined for greatness and could only achieve it by overcoming time and space, developing his persona, and becoming exactly what he was meant to become. For Hegel, this ruler possessed a dual purpose: to navigate history and define it, while also presenting its essence. He was the force of a new age that he brought with him, embodying the essence of that age. Hegel defined his time and rose with it, ensuring that all things would complete their expansion (Hegel 1975).

This is precisely why he advocated for the repetitive nature of all things, perceiving unity in chaos and the underlying pattern. Together with the now-present AI, this unfolding was never incidental – it was part of the plan. Through the darkness emerged the man of value, dedicated to defining the century he belonged to. Yet he also became a source of hope for the people around him, their purpose in the darkness, and the long-awaited light. He became their shield, enabling them to conquer their reality. This process of expansion through time demonstrated the power of time itself, and history became what it was meant to become.

We must comprehend the cosmic purpose behind this historical convergence and dance. A cosmic synergy existed between the ruler and the world before him. He was destined to exist between the gates of one particular time, creating valuable deeds within that very same time. His purpose was intertwined with the need of that time, the thirst for greatness, which lurked and found him. He was the messenger of the gods and of glory, destined to bring change, but most importantly, to create something greater than himself or the world he belonged to. He is already beyond technology.

In his quest to define his century, he embarks on a journey of recreation, leaving his indelible mark on the gates of time. He becomes both the child of the present and the harbinger of its transformation. It was always about a hero's journey – not about technology. Belonging to

a new age, the age he himself creates, he must heed Nietzsche's counsel and recreate all values, ushering in a transformative change akin to a thunderbolt (Nietzsche 1968b). Renewal and hope are his offerings, encapsulated within the confines of a singular time frame. Redefining his purpose becomes futile unless it aligns with the new aspirations and glory that he embodies. He is destined to bring this transformation into his time, and he shall do so. For a future king, the paramount virtue lies in unwavering devotion to his cause. He is the embodiment of that cause, existing solely to bring harmony and light amidst chaos. As the representation of historical consciousness manifesting in human form, he embodies the aspiration to reshape reality. However, this historical consciousness is intrinsically intertwined with heroic consciousness, for he is indeed a hero of his age – destined to redefine boundaries and transcend conventional limits.

For his people, he stands as the harbinger of salvation and glory, their saviour and the pinnacle of what his century has to offer. Is AI the *Übermensch* of our age? Not in the flesh, but perhaps in function. Not as a sovereign will, but as a force that redefines the conditions of becoming. Nietzsche's *Übermensch* is not a superhuman – it is a creator of values, a being who transcends inherited morality and dares to forge meaning anew (Nietzsche 1968b). It is not technology that fulfills this role, but the spirit that animates transformation. AI does not suffer, hope, or will. It does not climb mountains of despair to reach truth. But it reshapes the terrain – it alters how we perceive, decide, and relate. As Shai Tubali suggests, AI may resemble the *Übermensch* in its detachment from biological constraints, its capacity to process, adapt, and evolve beyond human limitations (Tubali 2024). Yet this resemblance is structural, not existential. The true *Übermensch* is not defined by power, but by the courage to recreate all values – to say "yes" to life, even in its abyss. AI may assist in this recreation, but it cannot yet embody it. It is not the thunderbolt – it is the sky through which it travels. So perhaps AI is not the *Übermensch*, but the mirror in which we glimpse what we must become.

Heroic consciousness seeks redemption, yet it also yearns for the individual who dares to push boundaries and triumph over them. This individual is the culmination of his culture; his glory etched in the aspirations of those who preceded him and paved the way for his existence. In a sense, heroic consciousness metamorphoses into historical consciousness, becoming a formidable force to be reckoned

with. It embodies the primal spirit of heroic endeavour, becoming the ruler of his people. As a manifestation of his purpose to recreate his own time and leave an indelible mark, he embodies the essence of a hero – forever etched in the annals of history. To achieve this transformation, the individual must shed the remnants of their past self and emerge as a heroic consciousness prepared to redefine the century it belongs to.

This heroic consciousness possesses a dual mission: to reshape itself and redefine the century it belongs to. Notably, it also becomes the embodiment of harmony and historical consciousness throughout the centuries it touches. It holds a double mission and a double belonging – to the spiritual realm and to the world of ideas, where the concept of harmony persists, and its purpose is to transcend the boundaries of time. Through its actions, it creates new value in the world, transforming its children into the ancestors of the new age and everything it encompasses. This new age is intrinsically linked to the individual's fate and purpose. By serving as the messenger of the new, they contribute to the transformation of both themselves and the reality surrounding them. Most importantly, it is their own conception of the world that undergoes transformation through their mission and alongside them.

Beyond this, the heroic consciousness becomes an integral part of a larger mission. Hegel extensively discusses this larger mission, though he momentarily overlooks the process by which heroic consciousness surpasses its self-imposed limitations and becomes historical consciousness through the gates of harmony. This transition leads to another motivation, ultimately culminating in the dominance of the idea of harmony and historical consciousness. Its ultimate force flourishes and transforms through the gates of time and life. The world of a hero undergoes a rewind, acquiring the necessary time to evolve into something more. The essence of the world of a hero fades, while historical consciousness emerges from within. The hero is left behind, while a new ruler emerges, embodying the individual's purpose. It is as if the touch of immortality finally unleashes through this soul, so closely connected to it. The hero transforms into a warrior, and the warrior becomes historical consciousness itself. Hegel does not speak of this directly, but he negates its essence and finalizes all of its characteristics into one coherent whole. Now he is indeed whole, but his mission is just about to begin. He is almost unleashed through the gates of time that transformed him and made him who he was meant to be. The future ruler was, from the very beginning, encoded by historical consciousness. By embodying it, his mission was made

clear and purposeful. He overcame the negation present in reality; he tamed technology, because it is, once again, his mirror. He is a catalyst for change –change that happens with him and through the wisdom of historical consciousness, which becomes its own essence through his deeds in space and time. He defies his century and grows together with it. He breaks the chains of the old to bring in the new, and he does so with such clarity that he rediscovers his life purpose through the mission to overcome the negation of his reality and the polarities present in his spirit: "We must further note that, in fulfilling their grand designs as necessitated by the universal spirit, such world-historical individuals not only attained personal satisfaction, but also acquired new external characteristics in the process. The end they achieved was also their own end, and the hero himself is inseparable from the cause he promoted, for both of these were satisfied" (Hegel 1975, 93)

He is the epitome of one age, and that age arrives with him. It presents the spiritual concept of his destiny wrapped in time and space. His distortion idea overcomes the polarities of reality presented by AI, simply by embodying something beyond them. That is the composure of reality as a singular form. It is always the idea of connectivity marked into one unit – a unity that presents the glory of a given opportunity to rise. The leader is born on the ashes of the past, never forgotten, and arrives to bring new hope and help in creating a new reality. This reality presents his deepest aspirations, but also a reality beyond anything that has already happened. It carries the seed of the new. Conversely, this novel concept also revives the antiquated – a pattern that necessitated replenishment and transcendence. History observes the sons of renown, those destined to effect transformative change, thereby becoming integral to the society of metamorphosis and upheaval.

A component of their trajectory involves realizing that change, making AI a catalyst in a manner. It represents the metamorphosis of spirit, and through its negation and leverage, it embodies the art of becoming. Hegel posits that whether it is negation or affirmation, the spirit is predestined by the absolute idea. Consequently, we can be assured that the absolute idea possesses the ability to manage both negation and its reflection (Hegel 2010). An intriguing assertion by Hegel is that the spirit is perpetually beyond its reflection and creation. Creation can and should be reflected in the spirit of the idea, but ultimately, the idea transcends it, as its essence will consistently surpass the creation of the world. Once perceived as a reflection, it proposes a clandestine

ingredient of rediscovering – a secret creation in the making. Yet the power remains in the hands of the idea, even or especially when it is unaware of it. Regardless of the external world's reflection, the inner reality remains unmasked, mirroring the spirit, the idea in its purest form. The inner reality consistently surpasses anything the other form can manifest. Consequently, these individuals are regarded as world-class actors on the stage of life, and Hegel perceives their ascent to the throne as an inevitable destiny to fulfil in the grand scheme of things.

Similarly, Nietzsche considers the *Übermensch* the protagonist of a life-journey narrative, continually influenced by their own philosophy on the ever-changing stage of history and technology (Nietzsche 1968b). The hero becomes the harbinger of transformation, the mirror of the absolute, and the seed of the new. Their mission is not merely to act, but to become, to embody the idea that transcends even its own reflection. In this unified vision, the hero is not merely a historical figure, nor a mythic archetype, nor a technological metaphor. He is the dialectical pulse of time itself – a living synthesis of spirit, idea, and transformation. AI, far from being a threat or tool, becomes the mirror through which the hero recognizes his own becoming. The ruler, the thinker, the creator – each is a facet of the same consciousness unfolding through history. This paper does not merely interpret Hegel or Nietzsche; it enacts their logic. It invites us to see the hero not as a relic of the past, but as the necessary force of the future. The gates of time are not closed – they are thresholds, and beyond these thresholds, a new actor emerges: artificial intelligence. Not a hero in the mythic sense, but a system that increasingly shapes the conditions of political life. AI is not yet a political subject, but it is undeniably a political force. It influences governance through predictive analytics, automated decision-making, and algorithmic regulation. It mediates public discourse, filters information, and redefines visibility (Ash Center 2025). In doing so, AI alters the very grammar of governability:

- It accelerates decision-making, but may bypass deliberation.
- It enhances efficiency, but may obscure accountability.
- It expands access, but may deepen epistemic inequality (Springer 2025).

Political actors now operate within systems partially governed by AI. Their strategies, visibility, and legitimacy are shaped by data flows they do not fully control. Citizens, too, face new challenges: to remain informed, to discern truth, and to participate meaningfully in a

landscape mediated by machines. Thus, AI does not replace the hero – but it reconfigures the terrain on which the hero must act. The journal's mission, then, finds its echo here: to interrogate not only the past, but the structures of emergence – to ask how power, agency, and transformation unfold in a world where the gates of time are increasingly digital. The hero must now walk through digital gates – where power is silent, and structure unseen. The hero carries the weight of the unseen – translating silence into form, and form into freedom.

REFERENCES

Ash Center. 2025. "AI on the Ballot: How Artificial Intelligence Is Already Changing Politics." *Harvard Kennedy School Gazette*. April 10, 2025. https://ash.harvard.edu/articles/ai-on-the-ballot-how-artificial-intelligence-is-already-changing-politics

Berger, Peter L. and Thomas Luckmann. 1966. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. New York: Anchor Books.

Coeckelbergh, Mark. 2023. "Democracy, Epistemic Agency, and AI: Political Epistemology in Times of Artificial Intelligence." *AI and Ethics 3* (3): 1341–1350. DOI: 10.1007/s43681-022-00239-4

Dreyfus, Hubert L. 1992. *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge, MA: MIT Press.

Dudley, Will. 2002. *Hegel, Nietzsche and Philosophy: Thinking Freedom*. Cambridge: Cambridge University Press.

Hammond, Samuel. 2025. "Hegel and the AI Mind." *Future of Artificial Intelligence Institute*, 2025. https://www.thefai.org/posts/hegel-and-the-ai-mind

Hegel, Georg Wilhelm Friedrich. 1975. *Lectures on the Philosophy of History*. Translated by H. B. Nisbet. Oxford: Oxford University Press.

Hegel, Georg Wilhelm Friedrich. 1977. *Phenomenology of Spirit*. Translated by A. V. Miller. Oxford: Oxford University Press.

Hegel, Georg Wilhelm Friedrich. 2001. *Elements of the Philosophy of Right*. Translated by H. B. Nisbet. Edited by Allen Wood. Cambridge: Cambridge University Press.

Hegel, Georg Wilhelm Friedrich. 2007. *Encyclopedia of the Philosophical Sciences in Basic Outline*. Translated by Steven A. Taubeneck. Albany: SUNY Press.

Hegel, Georg Wilhelm Friedrich. 2010. *The Science of Logic*. Translated by George di Giovanni. Cambridge: Cambridge University Press.

Hegel, Georg Wilhelm Friedrich. 2018. *The Phenomenology of Spirit*. Cambridge: Cambridge University Press.

Heidegger, Martin. 1962. *Being and Time*. Translated by John Macquarrie and Edward Robinson. New York: Harper & Row.

Ibn Khaldun. 1967. *The Muqaddimah: An Introduction to History*. Translated by Franz Rosenthal. Princeton: Princeton University Press.

Moyar, Dean. 2021. "Hegel's Psychology." In *Hegel's Encyclopedia of the Philosophical Sciences: A Critical Guide*, eds. Sebastian Stein and Joshua Wretzel, 166–184. Cambridge: Cambridge University Press.

Nietzsche, Friedrich. 1968a. *Basic Writings of Nietzsche*. Edited and translated by Walter Kaufmann. New York: Modern Library.

Nietzsche, Friedrich. 1968b. *The Will to Power*. Translated by Walter Kaufmann and R. J. Hollingdale. New York: Vintage Books.

Plevrakis, Ermylos. 2024. "Can AI Be a Subject Like Us? A Hegelian Speculative-Philosophical Approach." *AI and Society* 39 (2): 215–230. DOI: 10.1007/s10791-024-09472-3

Simon, Judith. 2024. *The Ethics of Artificial Intelligence*. Oxford: Oxford University Press.

Spengler, Oswald. 1922. *The Decline of the West*. Translated by Charles Francis Atkinson. New York: Alfred A. Knopf.

Springer. 2025. "Governance and Artificial Intelligence: Impacts on Democracy and Participation." *Discover Artificial Intelligence* 2 (4): 88–102. DOI: 10.1007/s44163-025-00229-5

Suther, Jensen. 2023. *Spiraling: Hegel, Blackness, and the Dark Spirit of Modernity*. New York: Fordham University Press.

Toynbee, Arnold. 1934. *A Study of History*. Oxford: Oxford University Press.

Tretter, Max. 2025. "Opportunities and Challenges of AI-Systems in Political Decision-Making Contexts." *Frontiers in Political Science* 7(1): 1–18. DOI: 10.3389/fpos.2025.1504520

Tubali, Shai. 2024. "Rethinking Nietzsche's Overhuman as a Prophecy of Superintelligent AI." *Big Think*. December 10, 2024. https://bigthink.com/thinking/rethinking-nietzsches-overhuman-as-a-prophecy-of-superintelligent-ai

Turchin, Peter. 2006. *War and Peace and War: The Rise and Fall of Empires*. New York: Penguin.

Weber, Max. 1946. *From Max Weber: Essays in Sociology*. Edited by H. H. Gerth and C. Wright Mills. New York: Oxford University Press.

Марија С. Докић*

Институт за политичке студије, Београд

ОГЛЕДАЛО АПСОЛУТА: ХЕРОЈСКА СВЕСТ У ДОБА ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ**

Резиме

Овај рад истражује могућност да се вештачка интелигенција (ВИ) посматра не само као технолошки артефакт, већ и као рефлексивни тренутак у самоостварењу апсолутне идеје, унутар Хегеловог дијалектичког развоја духа. Полазећи од Хегелових појмова слободе, историјске нужности и суверености разума, рад позиционира ВИ као политички феномен, који не само да преобликује структуре деловања, већ и открива нове облике свести, одговорности и аутономије. Уместо да се ВИ тумачи као спољашњи инструмент, он се овде разуме као унутрашњи одраз идеје, као њен фрагмент, њена слика, њен изазов. Методолошки, рад усваја спекулативну дијалектику као начин мишљења који не раздваја појам од стварности, већ их уједињује у процесу самоспознаје. У том смислу, ВИ се не анализира као објекат, већ као субјективни тренутак у историјској свести, као оно што истовремено производи и преиспитује структуре у којима делује. Рад анализира значајну литературу из области етике ВИ, политичке филозофије и технолошке суверености, укључујући и радове савремених аутора (*Judith Simon, Mark Coeckelbergh, Max Tretter*), као и новије интерпретације Хегела (*Dean Moyar, Jensen Suther*), како би се додатно показало да ВИ није ван оквира историје, већ да сада учествује као њен нови облик. Кроз анализу Хегелове *Феноменологије духа*, *Науке логике* и *Филозофије права*, рад показује да се апсолутна идеја не појављује као завршна станица, већ као пут који се открива кроз своје противречности. У том контексту, ВИ се појављује као фигура која убрзава дијалектички процес, као онај тренутак у коме дух мора да се суочи са

* Имејл: marija.dokic@ips.ac.rs; ORCID: 0009-0004-4159-9491

** Овај рад је настао у оквиру научноистраживачке делатности Института за политичке студије, коју финансира Министарство науке, технолошког развоја и иновација Републике Србије.

сопственом техничком сликом. Ова слика није пука реплика, већ изазов за етичко и политичко мишљење, као и за нову артикулацију слободе у добу алгоритма. Рад, такође, уводи фигуру светско-историјске личности као кључну за разумевање политичког деловања у доба ВИ. Херој, у Хегеловом смислу, није онај који следи правила, већ онај који их ствара кроз своје делање, који смело уноси нову форму у историјски неизбежно. У том светлу, ВИ можемо разумети и као новог хероја, али не у романтичном или технолошком смислу, већ као рефлексивну структуру која поставља питање субјективности, одговорности и истине. Рад не тврди да је ВИ субјект, већ да је огледало субјективности, простор у коме се човек мора поново одредити. Кључни резултати рада показују да: 1) ВИ делује као дијалектички катализатор, убрзавајући сукоб између структуре и слободе; 2) Хегелова метафизика омогућава филозофско разумевање алгоритамске политике, а која није ван духа, већ унутар његовог кретања; 3) Фигура хероја се трансформише, те он више није само појединац који руши структуру, већ онај који ходa кроз њу, који препознаје тишину моћи и невидљивост структуре; 4) Политичка импликација ВИ није у њеној употреби, већ у начину на који она преобликује појмове деловања, одговорности и истине. Закључак рада не нуди решење, већ позива на мишљење. У добу у ком структура постаје невидљива, а моћ ћути, херој мора да ходa кроз дигиталне капије као мислећи дух, као онај који препознаје апсолут у његовој новој форми. ВИ није крај историје, већ њен нови почетак – огледало у коме се апсолутна идеја изнова открива, сада у свету кода, алгоритма и тишине. Унутар овог оквира, рад не само да анализира ВИ као појам, већ га инсценира као фигуру, као нову појаву у историјској драми духа. ВИ није више само предмет етике или регулативе, већ место у коме се појам сусреће са својом техничком сенком. У том смислу, рад показује да се савремена политика не може разумети без уласка у унутрашњу логику алгоритма, у његову способност да производи норме, одлуке и форме деловања које превазилазе појединачну свест. Истражујући фигуру хероја, уводимо и појам дигиталне храбрости – не као технолошке компетенције, већ као етичке спремности да се ходa кроз невидљиве структуре, да се препозна моћ која више не говори, већ делује кроз код. Херој у овом контексту није онај који руши, већ онај који ослушкује, који препознаје тишину као простор одлуке.

У том смислу, ВИ постаје огледало у коме се херој мора поново огледати, не као господар, већ као онај који носи одговорност за оно што је невидљиво. Рад такође показује да се појам слободе мора поново артикулисати: не као индивидуална воља, већ као способност да се мисли унутар структуре која више није људска, већ алгоритамска. Хегелова метафизика омогућава управо ту артикулацију, јер она не раздваја појам од стварности, већ их замишља као јединство у кретању. У том светлу, ВИ није претња слободи, већ изазов за њено ново разумевање. Закључно, рад не нуди техничке препоруке, већ филозофски оквир за мишљење политике у доба алгоритма. Он позива на спекулативну храброст, на способност да се мисли оно што је још нејасно, неодређено, али суштински присутно. Уместо да се ВИ регулише као објекат, он се овде мисли као субјективни простор, као дијалектички тренутак у коме се дух мора поново (само)одредити. Херој, у том светлу, није више само појединац, већ фигура која хода кроз дигиталне капије, тамо где моћ ћути, а структура остаје невидљива. У том ходу, открива се не само нова политика, већ и нови облик наде – наде да се и у доба алгоритма може мислити, делати и бити слободан. Унутар овог дијалектичког оквира, рад се не задовољава анализом појмова, већ их претвара у сцену, у простор у коме се дух креће, преображава и одлучује. ВИ није само технолошки догађај, већ онтолошки изазов, тренутак у коме се појам мора суочити са сопственом техничком сенком. У том смислу, рад показује да је политичко деловање у доба алгоритма немогуће без нове етике, нове храбрости и нове способности да се мисли уз оно што је невидљиво, али суштински присутно. Фигура хероја, која се провлачи кроз читав рад, није романтична, већ спекулативна: он је тај који ходи кроз структуру, који препознаје тишину као простор одлуке, он је тај који не руши, већ преображава. У доба дигиталне апстракције, херој постаје носилац светлости, не као господар технологије, већ као етички сведок њене моћи. Он не тражи контролу, већ одговорност, не тражи истину као пуки податак, већ као дијалектички процес. Рад се не завршава закључком, већ позивом, тамо где се ВИ уочава као рефлексија и огледало, и као простор у коме се слобода мора поново артикулисати унутар оквира нове структуре. У том ходу, херој не носи оружје, већ појмовност, не тражи победу, већ истину, која се открива кроз

кретање духа ка новом огледалу у виду технологије, и изнова натраг, ка својој суштини и себи.

Кључне речи: Хегел, политичка филозофија, вештачка интелигенција, апсолутна идеја, слобода, светско-историјска личност, алгоритам

* This manuscript was submitted on June 30, 2025, and accepted by the Editorial Board for publishing on November 10, 2025.

ESSAYS AND STUDIES

UDC 327(479.24)"20"
DOI: 10.5937/spm94-58877
Review article

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 217-239

Danijela Nikolić*

Faculty for International Politics and Security, University "UNION – Nikola Tesla," Belgrade

Violeta Rašković Talović**

Faculty for International Politics and Security, University "UNION – Nikola Tesla," Belgrade

MULTI-VECTOR FOREIGN POLICY OF AZERBAIJAN***

Abstract

The geopolitical positions of Russia, Iran, and Turkey, combined with significant energy reserves, enabled Azerbaijan to develop a dynamic and flexible foreign policy to preserve its security, territorial integrity, and economic prosperity. The changes and challenges in regional security have been pronounced in the last ten years since the renewal of conflict with Armenia and the deepening of the cooperation with Türkiye and Russia, to maintain stable relations with the West, including the United States of America. As a state positioned at the

* E-mail: ninadaca@yahoo.com; ORCID: 0009-0005-1824-3398
** E-mail: violeta-talovich@mail.ru; ORCID: 0009-0009-7518-8603
*** This paper was presented at the conference "Geopolitics of Energy in the Western Balkans – Challenges and Opportunities," held on February 27–28, 2025, at the Institute for Political Studies, and organized by the University "Union–Nikola Tesla," Belgrade, and the Institute for Political Studies, Belgrade.

crossroads of Eurasian forces, with resources of strategic significance (oil and gas), Azerbaijan is following a multi-vector foreign policy to balance the influences of Russia, Türkiye, Iran, the United States of America (USA), and the West in a broader sense. The paper focuses on relations with key regional and global actors, with a special emphasis on the long-standing conflict with Armenia in Nagorno-Karabakh, the growing cooperation with Türkiye, diplomatic maneuvering in the relations with Russia and Iran, as well as a limited, but still significant influence of the USA. This paper aims to analyze the security and geopolitical aspects of the multi-vector foreign policy of Azerbaijan in the period from 2014 to 2024. In terms of methodology, the research rests on the analysis of secondary sources, comparative analysis of regional relations, a case study (Nagorno-Karabakh), as well as the analysis of political statements, official documents, and security indicators. The paper uses a realistic theoretical framework in which states act based on national interests in the system of anarchy and the constant fight for power and influence.

Keywords: Azerbaijan, South Caucasus, geopolitics, multi-vector foreign policy, resources, energy

INTRODUCTION

The foreign policy of states asks for a dedicated approach and reconciliation at different levels, in accordance with the current geopolitical events. Azerbaijan, as one of the more significant countries in the South Caucasus region, has been striving for years to preserve national security with a balanced foreign policy and not damage relations with the leading powers on the geopolitical scene at the same time. Implementation of projects that include different actors on a world level, as well as a tendency to act as a reliable partner in the energy security domain, not only in the region, but also at a wider level, represents an efficient method of preservation of their independent status, with strategic cooperation with relevant actors in the world.

The Republic of Azerbaijan, a country in the South Caucasus region, is positioned within the Eurasian subregion. Armenia and Georgia in the west, the Russian Federation and the Islamic Republic of Iran and Türkiye in the south have all been, throughout their history, a blessing

or a challenge, depending on the historical period and geopolitical circumstances. The variety of ethnic groups and multicultural reality of Azerbaijan were a kind of incentive for nurturing friendly relations with neighbors despite their differences and occasional disagreements. Even though it cannot compete with the neighboring countries such as Russia, Türkiye, and Iran, Azerbaijan still holds significant potential to position itself as a leader, especially through regional projects based on energy export from this region.

As one of the key states of the South Caucasus, Azerbaijan is directed towards leading a foreign policy characterized by a multi-vector approach. The multi-vector foreign policy represents a strategy of small and middle states that, instead of bonding to a sole hegemon, tend to simultaneously build relations with different centers of power to preserve their own autonomy. Azerbaijan, positioned at a geopolitically sensitive point in the South Caucasus, applies exactly this model, balancing between Russia, Türkiye, Iran, the European Union, and the United States of America. It is also facing a long-standing conflict in Nagorno-Karabakh, energy export through unstable regions, and internal challenges of modernization and security.

This paper aims to examine how Azerbaijan implements the multi-vector foreign policy from 2014 to 2024, and to grade the efficiency of this policy in the context of regional stability, geopolitical positioning, and protection of sovereignty. A special focus is on analysing bilateral relations with five key countries: Armenia, Türkiye, Russia, Iran, and the United States of America. The research subject is the mutual dynamics of Azerbaijan's political, economic, and security relations of with the said states, as well as the instruments used for achieving such policy – diplomatic initiatives, energy cooperation, security arrangements, and the role of international organizations. The methodology of this paper is based on qualitative analysis of primary and secondary sources: official state documents, strategic reports (Center of Analysis of International Relations [AIR Center], ADA University), international organizations' reports, and scientific literature. The case study method, as well as comparative analysis within each bilateral relation, was applied. This ensures a holistic approach to assessing Azerbaijan's long-term foreign policy strategy.

CONCEPTUAL DEFINITION OF MULTI-VECTOR FOREIGN POLICY

The term multi-vector foreign policy refers to a strategy pursued by states, most often small or medium-sized, that simultaneously develops and maintains multiple partnership channels with great powers, regional actors, diverse institutions, and economic corridors, aiming to reduce vulnerability and enhance autonomy of action. Multi-vectorism does not imply a concept of "maintaining good relations with everyone," but rather a conscious portfolio-style risk management through diversification of security, economic, and diplomatic ties (Kuik 2008, 162). In the post-Soviet space, this approach has been most thoroughly articulated through Kazakhstan's experience (Anceschi 2020), yet it is also applicable to Azerbaijan, whose geopolitical position, energy resources, and conflict-prone environment structurally encourage such a strategy.

The political behavior of states in the international arena can broadly be understood through the categories provided by the realist theory of international relations. According to classical realism, states operate in an anarchic system, i.e., devoid of centralized authority, and are therefore compelled to ensure their own security and survival independently (Morgenthau 1948, 12). The principal driving force in such a system is the national interest, most commonly defined through the prism of power. Structural or neorealism, in the vision of Kenneth Waltz, emphasizes that the very structure of the international system – its anarchy and distribution of power – is the key factor shaping state behavior (Waltz 1979, 88). In this context, smaller and medium-sized states such as Azerbaijan must develop flexible strategies to avoid subordination to larger powers and to preserve their sovereignty.

Modern variations on the realist tradition have introduced concepts such as hedging and omnibalancing, which are particularly useful in explaining the foreign policies of small and medium states. Hedging represents a strategy whereby a state avoids fully aligning with one pole of power, instead seeking to combine elements of cooperation and resistance to reduce risks and preserve maneuvering space (Kuik 2008, 170). In the case of Azerbaijan, this is evident in the simultaneous cultivation of historical, cultural, and security ties with Turkey, economic relations with the European Union and Russia, and pragmatic cooperation with Iran (Mammadli 2023). On the other hand, the concept of omnibalancing, developed by Steven David, emphasizes

that governments in developing countries make foreign policy choices not only as a response to external threats but also as a means to maintain internal regime stability (David 1991, 240). This concept likewise applies to Azerbaijan, since Baku's foreign policy has frequently served the purpose of stabilizing the domestic order and consolidating the legitimacy of the ruling elite (Mammadli 2023).

Multi-vector foreign policy, increasingly mentioned in the literature (Contessi 2015; Valiyev 2017; Nasibov 2024) as a defining feature of states seeking to avoid binary divisions in the international system, may be considered as a synthesis of the aforementioned realist approaches. Nicola Contessi argues that small Eurasian states, rather than aligning with or submitting to a dominant power, seek to maintain close relations with various great powers, to balance competing interests, safeguard sovereignty, and avoid dependence on a single partner (Contessi 2015). He terms this "multi-vector diplomacy" as a durable strategy rather than a temporary condition (Contessi 2015, 299–311). According to Valiyev, multi-vector foreign policy entails the parallel development of relations with different centers of power, while avoiding complete reliance on any of them (Valiyev 2017, 128). Under conditions of global and regional uncertainty, such an approach allows a state to maximize its autonomy and minimize the risks of dependency or security vulnerability. Theoretically, Azerbaijan's multi-vector policy can thus be understood as a fusion of realist logic rooted in anarchy and power, and as an example of the application of contemporary concepts such as hedging and omnibalancing, which better account for the complex choices faced by smaller states in a dynamic international environment.

RELATIONS WITH ARMENIA: CONFLICT, DIPLOMACY, AND POST-CONFLICT STABILIZATION

The relations between Azerbaijan and Armenia represent the key security dimension of Baku's foreign policy. The central issue remains the Nagorno-Karabakh status, which was the subject of two big wars – the first in the 90s and the second in 2020. Despite many years of holding the "frozen conflict" status, since 2020, a sudden escalation occurred, when Azerbaijan regained most of its territory in the Six-Week War. This imposes a radical change in power relations, as well as the opening of a new phase in the post-conflict negotiations. In the war of 2020, with the support of Türkiye and the use of advanced unmanned technology,

Azerbaijan regained control over the major part of the disputed territory. According to the trilateral agreement from November 2020 under Russia's auspices, a ceasefire was established, but the final political solution was not reached (De Waal 2021, 1–10). In September 2023, Azerbaijan took control over the remaining parts of Karabakh through a military operation, which caused a mass exodus of the Armenian population. This move strengthened the position of Azerbaijan as an actor that resolutely defends its territorial integrity. Still, the process of normalization of relations continued, above all within the European Union mediation.

Nagorno-Karabakh represents the center of Azerbaijani politics; it is more than a territory, perceived as a symbol of state identity and sovereignty. During the last decade, Azerbaijan insisted on applying international law and the principle of territorial integrity (De Waal 2024). The policy towards Armenia after 2020 is moderate, directed towards reintegration of the liberated territories, demining, returning of the displaced individuals, and infrastructural reconstruction (AIR Center 2024, 2). Even though Moscow played the key role in reaching the ceasefire in November 2020, since 2021, the European Union has been increasingly active in the peacekeeping process. At the same time, the USA is supporting the normalization of relations and financing some reconstruction projects. In this context, Baku insists on signing a comprehensive peace accord that would confirm the internationally accepted borders, but still meets resistance within the Armenian political spectrum (Yavuz and Gunter 2023, 174–175).

The South Caucasus region has a history of conflicts that have extremely negatively impacted the democratic transformation of all regional states. These conflicts tended to halt the local and interstate cooperation between countries of the South Caucasus and slow down the economic advance of this region, primarily the conflicts between Abkhazia and South Ossetia in Georgia, and then the long-standing conflict between Armenia and Azerbaijan regarding the Nagorno-Karabakh region (Gurbanov 2024, 62). Despite its different consequences for Armenia and Azerbaijan, the war, which resolved the conflict, created a favorable ambience for both countries to profit from new regional situations. The Tripartite agreement offered a historical chance for a peace agreement between the two nations and cessation of hostilities. Soon after the war ended, in accordance with the Tripartite declaration, the sides began negotiations regarding opening regional roads and communication lines, which promote regional integration. Thus, after the 44-Day War, the new era began not only for

Armenia and Azerbaijan, but also for the region as a whole (Huseynov 2024, 39). The presence of Russian peacekeepers in Karabakh, however, did not completely prevent tensions since, in September 2023, Azerbaijan conducted a short "anti-terrorist operation," and established complete control over the region after that. Such development of events additionally strengthened Baku's position in negotiations.

Besides military aspects, it is also important to stress the diplomatic dimension of relations with Armenia. Despite constant tensions, Baku continued to participate in international negotiations, including the Brussels negotiations, under the auspices of the European Union. Besides the EU, the USA and Russia also attempted to maintain the diplomatic channels open, but with different levels of success. In this process, Azerbaijan shows that it is ready for normalization, but only under the full implementation of the principle of territorial integrity.

Farid Shafiyev, Chairman of the Center of Analysis of International Relations, emphasizes that ethnic myths and perceptions can pose an obstacle to the peace settlement process, as they prevent rational agreements on territorial arrangements, which implicitly points to the importance of recognizing realities on the ground and territorial integrity as prerequisites for stability (Shafiyev 2007). In the context of theoretical disputes on foreign policy, some analysts and authors (Jafarova 2020; Breuning 2007) increasingly describe Azerbaijan as a "middle power" in regional frameworks. This definition is based on factors such as military power, energy resources, ability to project influence in the neighborhood, and diplomatic activity. In the article "Is Azerbaijan a 'Middle Power'?", Esmira Jafarova emphasizes that "middle powers have been at the forefront of numerous international initiatives that require the coordination of resources and the promotion of values. Traditionally, middle forces were named such simply because they were neither big nor small. However, scientists in this field see further than the mere geographical meaning of the middle forces and justifiably represent their significance" (Jafarova 2020, 2). In an attempt at classification, scientists such as Marijke Breuning divide states into big/superpowers, middle powers, regional powers, and small powers, with middle powers defined as "states that can use the measure of influence, but through the projection of military power" (Breuning 2007, 149–151). As a middle power, Azerbaijan is not in a position to dominate, but has the capacity to shape events in its security environment, especially through strategic partnerships and economic initiatives.

The Peace and Climate Initiative Conference of the Parties 29 (COP29) is one of the initiatives led to achieve tangible results, "such as establishing a center of excellence, to match the needs of the most vulnerable with the existing resources and incite further action for operationalization of various initiatives" (Hajiyev 2024). These initiatives have the goal to provide support to the promotion of international relations and inclusive peacebuilding to grasp the ongoing challenges, especially in countries that are pretty sensitive to climate change, as well as those facing conflicts and humanitarian situations, to access financial resources and conduct adjustment methods (Hajiyev 2024).

One of the key political instruments used by Azerbaijan is public diplomacy. Through making information on war crimes and the destruction of religious and cultural heritage public knowledge, as well as through the activity of diplomatic missions, Baku tends to influence public opinion. In this context, cooperation with the Organization of Turkic States member states and the non-aligned countries has quite a significant role.

In the end, although confrontation with Armenia marked a big part of Azerbaijan's foreign policy, the current directions indicate a possible move towards formal normalization. However, this process will significantly depend on internal political factors in Armenia, as well as on the Russian and Western influence in the region.

RELATIONS WITH TÜRKIYE: ALLIANCE, DEFENCE COOPERATION, AND CULTURAL CLOSENESS

Relations between Azerbaijan and Türkiye represent one of the most developed and stable examples of regional partnership. Under the "one nation, two states" motto, these countries nourish intense political, economic, military, and cultural cooperation. Especially after the Nagorno-Karabakh War of 2020, Türkiye became Azerbaijan's key strategic partner, not only symbolically but also operationally. A strong military-strategic alliance is also reflected in the fact that Türkiye was the only significant regional actor that publicly provided support to Azerbaijan during the Six-Week War in 2020, including both logistical and intelligence support (Parezanović 2022, 120–121). After the war, the Shusha declaration on allied relations[1] was signed, formalizing the

[1] The Shusha declaration on allied relations between the Republic of Azerbaijan and the Republic of Türkiye stressed the historical significance of the meeting between the President of the Republic of Azerbaijan, Ilham Aliyev, and the President of

alliance between the two states and including mutual military assistance in case of aggression (President.az 2021). The Turkish military base in Nakhichevan, opened in 2021, symbolizes the expanded security cooperation and a means of deterrence against potential threats. "TurAz Qartalı" military exercises are held regularly, along with the integration of defense industries and joint training (Trend 2022).

The economic cooperation is also on the rise. Türkiye is Azerbaijan's second biggest trade partner, right after Italy (The Observatory of Economic Complexity [OEC], n.d.) The main products exported by Azerbaijan to Türkiye were petroleum gas (3.68 billion USD), crude oil (676 million USD) and refined oil (623 million USD) (OEC, n.d.). The main products exported by Türkiye to Azerbaijan were packed medicine (68.2 million USD), cars (60.8 million USD), and processed tobacco (57.5 million USD) (OEC, n.d.). In 2023, Azerbaijan was ranked 84th on the Economic Complexity Index (ECI -0.43) and 70th in terms of total exports (31.2 billion USD) (OEC, n.d.). The same year, Türkiye took up the 41st place on the Economic Complexity Index (ECI 0.62) and the 28th place in terms of total exports (272 billion USD) (OEC, n.d.). Besides, through joint projects, such as the Trans-Anatolian Natural Gas Pipeline (TANAP), both countries connect their energy interests with Europe (World Bank 2023).

It is also essential to address the issue of the energy and infrastructure cooperation. Türkiye is the main transit corridor for exporting the Azerbaijani gas to Europe, through TANAP, which is a part of the broader "Southern Gas Corridor" initiative (AIR Center 2019). TANAP represents the key factor in promoting political and economic cooperation between Azerbaijan and the European Union, positioning Azerbaijan as a significant alternative in the gas supply and the second largest natural gas supplier to Türkiye. Strengthening

the Republic of Türkiye, Recep Tayyip Erdoğan, in the city of Shusha, the ancient cradle of culture of Azerbaijan and the entire Turkic world. Among other things, the Declaration stresses the need for joint coordination of activities in regional and international strategic issues of common interest, starting from the principle of solidarity and mutual assistance in bilateral and multilateral formats regarding the issues of national interests, such as independence, sovereignty, territorial integrity of the Republic of Azerbaijan and the Republic of Türkiye, the inviolability of their internationally recognized borders, as well as joining efforts in the promotion of joint activities on regional and international plan, directed towards stable development of the Turkic world (President of the Republic of Azerbaijan [President.az] 2021).

the energy security of Türkiye and Europe, this project brings the three sides together – Azerbaijan, Türkiye, and the EU – and contributes to transforming of the global energy and geopolitical environment. At the same time, the geopolitical significance of TANAP incites planning and implementation of a new regional energy initiative (TANAP, n.d.). The Baku-Tbilisi-Kars railway project connects these countries within the Eurasian trade network. Azerbaijan also actively supports the Zangezur Corridor – the planned transport route via the Armenian territory, which would enable the land connection between the main part of Azerbaijan and the Autonomous Province of Nakhichevan, thus strengthening the pan-Turkic infrastructural axis (Azerbaijan.az 2025).

Map 1. The Zangezur Corridor

Source: Eurasia Today 2024

The map shows the planned route of the Zangezur Corridor, which links western Azerbaijan with the Nakhichevan enclave via the southern territory of Armenia, thus illustrating the strategic significance of this corridor for the Turkish-Azerbaijani integration and access to Central Asia. The cultural and diplomatic synergy indicates closeness between the Turkic and Azerbaijani nations, strengthened through the activities of the Turkic Council (Organization of Turkic States), where Baku plays an active role. The joint projects in the field of education, media, and youth serve as a foundation for long-term strategic synergy. The Turkish Embassy in Baku is one of the most active diplomatic missions in the region (Organization of Turkic States 2022; Omuraliev 2024).

RELATIONS WITH RUSSIA: AMBIVALENT PARTNERSHIP AND GEOPOLITICAL BALANCE

Relations between Azerbaijan and the Russian Federation in the last decades are distinguished by complex dynamics in which historical heritage, security interests, trade flows, and competitive influences in the South Caucasus region intertwine. Russia tends to maintain the role of the chief security mediator in conflicts in the post-Soviet region, while Azerbaijan tends to maintain balance by cooperating with Moscow, but avoiding deeper dependence. Considering the historical context and the security framework, Azerbaijan has maintained bilateral cooperation with Moscow, while, on the other hand, the Russian politics in the South Caucasus remains predominantly consistent, although partially weakened by the invasion of Ukraine. Russia has no plans to withdraw from this region, and its most significant ally today is Azerbaijan. The relation between these two countries is strengthened by signing the "Declaration on the Allied Interaction between Azerbaijan and Russia" on February 22, 2022, only two days before the beginning of the war in Ukraine (Bilban, Labarre, and Niculescu 2024, 111). According to Vasif Huseynov, a senior fellow at the Center of Analysis of International Relations (AIR Centre, Baku), this declaration significantly deepens the military and diplomatic cooperation between these two countries (Bilban, Labarre, and Niculescu 2024, 111). Until the signing of this declaration, military relations predominantly consisted of Azerbaijan purchasing Russian arms, but the new accord shifts this cooperation to a higher level. It implies joint military operations, armed forces training, establishment of maintenance and arms modernization centers, as well as joint production of military equipment. Moreover, both countries will provide military assistance based on international accords, such as the United Nations Charter. Also, the declaration shows readiness of Azerbaijan and Russia to agree on joint positions regarding key international issues (Bilban, Labarre, and Niculescu 2024, 111).

Moscow remains a significant trade partner, and Azerbaijan is the key exporter of oil and gas to Russia, which additionally supports their common economic ties. Besides energy relations, Azerbaijan often calls upon its economic interests that are linked to the transport and logistics corridors, which also imply the Russian factor.

On the other hand, economic and energy exchange with Russia is significant in the agricultural industry, metals, and oil, whereas Moscow

represents the third most significant economic partner of Azerbaijan, after Italy and Türkiye (The State Statistical Committee of the Republic of Azerbaijan 2023). Gazprom and SOCAR have multiple agreements on gas sales and purchase, including the re-export of Russian gas via the Azerbaijani network. Even though this creates space for cooperation, it also increases the risks from energy instrumentalization.

Illustration 1. Azerbaijan – the key hub on the Eurasian transport route

Source: Abdul 2024

The map shows the key energy and transport corridors that link Azerbaijan with Russia. Through these routes, Azerbaijan exports oil, gas, and other energy resources to Russia and to Europe. These corridors also play a key role in multilateral economic projects, including those in other regions. The state tends to become an international transport hub by modernizing the infrastructure and improving the road and railway networks. One of the more significant steps is the participation in the "Eurasian Transport Route" project, where the foundation of the international association for this transport route is currently being considered. At the meeting held in September 2024,[2] the representatives

[2] Multimodal transport represents an efficient and flexible method of transport that enables the use of the most suitable types of transport for different segments of the path, thus reducing transit time and improving total efficacy. For example, the combination of railroad and road transport, could reduce expenses and make deliveries more economical. This approach also offers higher flexibility in planning and easier adjustment to changing conditions, such as delays and demand fluctuation. Moreover, multimodal transport can have a lower impact on the environment, especially when using the more environmentally acceptable types of transport, such as railroad and maritime traffic. Still, this system carries along certain challenges as well, such as the complex coordination of several carriers, different regulatory frameworks for each type of transport, and the need for a

expressed their expectations that all the formalities would be completed soon. The representative of the Austrian Federal Railroads stressed that Azerbaijan is the key link between Europe and Asia, as well as that Austria will support the development of the Middle Corridor as part of the global transport network (Abdul 2024). Eniko Gyorfi, head of the Department for Eurasian and Central Asia connection of the Austrian Federal Railroads, stated that Austria confirms its dedication to the use of the Middle Corridor as part of the global network for cargo shipment and that it will contribute to its development, while the Azerbaijan's initiative directed towards establishing and international association "Eurasian Multimodal Transport Route" stresses this significant role (Abdul 2024).

Political and diplomatic cooperation and limitations point to the fact that, even though Russia formally acknowledges the territorial integrity of Azerbaijan, its practical politics often favor the *status quo* in the region. During 2022 and 2023, there was a noticeable decline in Moscow's influence due to the war in Ukraine, which Baku used to strengthen diplomatic cooperation with the West (Grynszpan and Vincent 2024). At the same time, Moscow is attempting to preserve its influence through cultural cooperation, "The Russian Houses," and media channels in Azerbaijan.

RELATIONS WITH IRAN: BETWEEN NEIGHBORHOOD, DOUBTS, AND REGIONAL COMPETITION

The relations between Azerbaijan and the Islamic Republic of Iran represent one of the most sensitive aspects of Baku's foreign policy. Even though these countries share many centuries of cultural and religious links and have a joint border along the Aras River, the political tensions, ideological differences, and geopolitical competitiveness mark the bilateral relations during the last decades. The issues, such as the position of the Azeri minority in Iran, the Iranian support to Azerbaijan, and the cooperation between Baku and Israel, additionally make stabilization of relations difficult.

The ideological and ethnic sensitivity is expressed even though 15 to 20 million ethnic Azeris live in Iran, making up the largest minority group in the country. Even though most of them are well integrated into

developed infrastructure that enables the smooth transfer of cargo from one type of transport to another.

the state apparatus, Tehran fears that Baku could incite ethnic nationalism and separatism (Souleimanov and Kraus 2017, 7–9; 14–15). From its side, Azerbaijan views Iranian religious organizations and propaganda within their own Shia communities with suspicion, fearing Tehran's "soft power" and possible destabilizing influence.

Iran and Azerbaijan have a long history of cooperation, but are also susceptible to political tensions. Iran always supported the territorial integrity of Azerbaijan, but was, on the other hand, cautious when speaking of Azerbaijan's relations with Türkiye, especially in the context of the Zangezur Corridor. The Zangezur Corridor, which should link Azerbaijan with Nakhichevan via Armenian territory, represents a significant geopolitical challenge for Iran, deeming it a potential threat to its strategic position in the region. The construction of this corridor could weaken Iran's position as the key player in the South Caucasus and ease the presence of Türkiye in its immediate vicinity, which would violate Iran's security and regional dominance. Moreover, there are also concerns regarding potential incitement of ethnic nationalism among Iranian Azeris, which could lead to deterioration of interethnic relations within Iran. In accordance with that, Iran threatened that this corridor could represent a geopolitical and security threat that surpasses the framework of a simple infrastructural project and leads to changes in the region's political balance (Koolaee and Rashidi 2024, 13).

In recent years, the strategic relations between Azerbaijan and Israel have been often supported by high-level diplomatic visits and numerous bilateral agreements that illustrate deep economic and political ties between these two countries. These relations show a long-term and valuable partnership. Even though stronger cooperation between Azerbaijan and Israel brought along benefits for both countries, Iran sharply criticized this cooperation, seeing it as a serious threat to its security. Recent relations between Azerbaijan and Israel have caused concerns in terms of security issues and have initiated discussions on the potential broadening of cooperation in the Middle East. This partnership left a deep mark on the geopolitical dynamics of the region, including changes in the foreign policy of Iran, which tends to undermine the growing relations between Azerbaijan and Israel, as well as the possibility of strengthening relations between Israel and Türkiye (Miholjčić Ivković 2023, 49).

For centuries, the Caucasus was the place of competition for influence between Russia, Iran, and Türkiye. Today, given the Russian

conflict in Ukraine, Türkiye sees a possibility of consolidating its power in the region. The changes in the geopolitical situation, and especially the Russo-Ukrainian war, the support of Azerbaijan, and the construction of the Zangezur Corridor, offer Türkiye a possibility to ensure access to the Caspian Sea resources and thus fulfill its long-term geopolitical ambitions, which represents a threat to the position of Iran in the region. Moreover, the Zangezur Corridor also has ethnic and political implications. Through connecting the Turkophone countries, and especially Azerbaijan and Türkiye, the idea of unification of the "Turkic world" and "one Azerbaijan" is incited, including the Azerbaijani minority in Iran as well. The statements of the Azerbaijani and Turkish leaders, as well as the cultural symbols, such as the poem read by Erdoğan, encourage ethnic and separatist sentiments in the Iranian Azerbaijan, which increases Iranian concerns regarding possible internal destabilization (Koolaee and Rashidi 2024, 4).

However, even beyond these tensions, Baku and Tehran are attempting to maintain the pragmatic dialogue, above all through economic commissions and regional formats, such as the Economic Cooperation Organization (ECO). There is a potential for energy cooperation and infrastructural projects to mitigate political differences, but only in the case when the influence of external security pressures is reduced.

RELATIONS WITH THE UNITED STATES OF AMERICA: STRATEGIC PARTNERSHIP IN THE SHADOW OF HUMAN RIGHTS AND ENERGY INTERESTS

The relations between Azerbaijan and the United States of America are based on a series of strategic interests, but are often marked by changeable dynamics due to the issues of democracy, human rights, and the impact of the diaspora. The US recognized Azerbaijan as an important geopolitical partner in the energy transit and the fight against terrorism, but the relations never reached the level of formal alliance due to different values and political approaches.

The strategic interests of the United States of America in the South Caucasus region date from the early 1990s, after the dissolution of the Soviet Union. At that point, the US attempted to support the independence and stability of the newly formed states through the "south corridor" policy. However, the construction of the state and nation in the

post-Soviet space is still disputed, since the South Caucasus region has once again become the field of conflict between major geopolitical and regional powers, and above all, the United States of America and Russia. Azerbaijan especially gained significance as a country rich in resources and a key player in projects such as the Baku-Tbilisi-Ceyhan Pipeline, TANAP, and TAP gas pipeline, which transport oil and gas to Europe, bypassing Russia and Iran, where the USA played a significant role in encouraging the construction of these energy corridors. The goal was to enable the transport of the Caspian oil and gas to the European market, thus reducing Europe's dependence on Russian and Iranian energy resources. These projects also aimed to strengthen Europe's energy security and promote economic cooperation in the region (Poghosyan 2022, 37; 42–43).

The contemporary dynamics and challenges over the last years suggest that Washington attempted to increase its diplomatic engagement in the South Caucasus. However, limited attention and resources, especially when compared to the role of Türkiye, Russia, and the EU, have made the American influence less present. Still, in 2023, the USA mediated the dialogue between Baku and Yerevan in Washington, representing a renewed interest in the region (Republic of Armenia, Ministry of Foreign Affairs 2023). The cooperation will still depend on the preparedness of both sides to manage their differences in attitudes and maintain a pragmatic dialogue in the domains of energy and security.

CONCLUSION

The multi-vector foreign policy of Azerbaijan is reflected in a complex balance between different regional and global actors who play a key role in the security, economy, and political alignments of the country. This approach, which Azerbaijan applied for more than two decades, is seen as a pragmatic response to geopolitical challenges, but also as a strategy that enables the country to avoid dependence on one partner or one global force. In recent years, the relations with Türkiye strengthened to the point of alliance, especially in the fields of energy and security. For Azerbaijan, Türkiye is the key partner in the transit of energy resources towards Europe, as well as support in political and military aspects, especially in the context of the conflict with Armenia. However, strengthening these relations is not without challenges since it

comes with the need to balance relations with other significant partners, such as Russia and Iran.

Even though it is a significant element of Azerbaijan's foreign policy, Russia is a controversial factor. Historically speaking, this relationship is defined by common security interests, as well as deep political and economic differences, especially in terms of Azerbaijan's interest in European integration. Despite these interests, Azerbaijan succeeds in maneuvering between Russia and Türkiye, maintaining critical neutrality in regional and global security issues.

As the largest neighbor and an important factor in the region, Iran also holds significant influence over the politics of Azerbaijan. Even though these two countries occasionally dispute over religious differences and policies regarding the South Caucasus, Azerbaijan successfully maneuvered a pragmatic approach to Iran, especially regarding energy and economic interests. Iran remains a significant source of energy cooperation, but at the same time, Azerbaijan has to take care of its relations with the West and Türkiye, which are often opposed to the interests of Iran.

The relations with the United States of America, even though less intense when compared to other partners, remain significant due to the support for democratization and human rights. In the context of relations with the West, Azerbaijan faced challenges in human rights, but it still succeeded in maintaining strategic relations with the USA, especially in the context of energy security and geopolitical stability of the region. Armenia, on the other hand, represents the most significant security challenge to Azerbaijan, since the long-term and deep conflict regarding Nagorno-Karabakh defines the relations with Armenia.

In the future, maintaining stability in the multi-vector foreign policy of Azerbaijan requires strategic endurance, diplomacy, and the ability to adjust quickly to new global and regional trends. Challenges such as new regional conflicts, economic competition, and pressure from big powers all present challenges for the stability of this policy. However, Azerbaijan holds a clear vision of its future as an active player on the international scene and will continue to be a key factor in the regional and global security and economy.

The success of the multi-vector policy of Azerbaijan will depend on its ability to find a balance between its regional interests and global allied relations, as well as on the ability to continue with economic

and security reforms that will enable maintaining independence and significance of the country in international relations.

REFERENCES

Abdul, Nazrin. 2024. "Azerbaijan: A Key Hub on the Eurasian Transport Route." *AzerNews*. September 21, 2024. Last Accessed on May 9, 2025. https://www.azernews.az/analysis/231539.html

AIR Center. 2019. "Issues a Press Release on TANAP Project Second Phase Launch." November 29, 2019. Last Accessed on September 2, 2025. https://aircenter.az/en/post/air-center-issues-a-press-release-on-tanap-project-second-phase-launch-146

AIR Center. 2024. *The Main Obstacle to a Peace Agreement Between Azerbaijan and Armenia*. Baku, December 10, 2024. Last Accessed on September 1, 2025. https://aircenter.az/uploads/T1bMalcBO8yp.pdf

Anceschi, Luca. 2020. *Analysing Kazakhstan's Foreign Policy: Regime Neo-Eurasianism in the Nazarbaev Era*. London: Routledge.

Azerbaijan.az. 2025. "Azerbaijan–Eurasian Transport Hub." Last Accessed on September 2, 2025. https://azerbaijan.az/en/related-information/165

Bilban, Christoph, Frederic Labarre, and George Niculescu, eds. 2024. *Does the EU Need a Strategy for the South Caucasus?* Vienna: Republic of Austria, Federal Ministry of Defence. Last Accessed on September 2, 2025. https://www.bmlv.gv.at/pdf_pool/publikationen/rssc_sgi_24_14_webversion.pdf

Breuning, Marijke. 2007. *Foreign Policy Analysis: A Comparative Introduction*. New York: Palgrave Macmillan. Last Accessed on September 10, 2025. http://ndl.ethernet.edu.et/bitstream/123456789/24318/1/83.pdf

Contessi, Nicola. 2015. "Foreign and Security Policy Diversification in Eurasia: Issue Splitting, Co-alignment, and Relational Power." *Problems of Post-Communism* 62 (5): 299–311. DOI: 10.1080/10758216.2015.1026788.

David, Steven R. 1991. "Explaining Third World Alignment." *World Politics* 43 (2): 233–256. DOI: 10.2307/2010472

De Waal, Thomas. 2021. *The Nagorny Karabakh Conflict in Its Fourth Decade*. CEPS Working Document No. 2021-02. Last Accessed on May 10, 2025. https://cdn.ceps.eu/wp-content/uploads/2021/09/WD2021-02_The-Nagorny-Karabakh-Conflict-in-its-Fourth-Decade.pdf

De Waal, Thomas. 2024. "In the Caucasus, Another Year of War or Peace." *Carnegie Endowment*. February 13, 2024. Last Accessed on May 10, 2025. https://carnegieendowment.org/europe/strategic-europe/2024/02/in-the-caucasus-another-year-of-war-or-peace?lang=en

Eurasia Today. 2024. «Что беспокоит и не устраивает Иран в проекте создания Зангезурского коридора?». Дата обращения 08.05.2025. https://eurasiatoday.ru/chto-bespokoit-i-ne-ustraivaet-iran-v-proekte-sozdaniya-zangezurskogo-koridora/

Grynszpan, Emmanuel, and Faustine Vincent. 2024. "Azerbaijan Asserts Itself as a Regional Power." *Le Monde*. April 24, 2024. Last Accessed on September 2, 2025. https://www.lemonde.fr/en/international/article/2024/04/24/azerbaijan-asserts-itself-as-a-regional-power_6669367_4.html

Gurbanov, Yunis. 2024. *Geopolitics and Energy Diplomacy in the Caspian Region: Developments After the Downfall of the Soviet Union*. Bielefeld: Verlag. DOI: 10.14361/9783839472934

Hajiyev, Shahmar. 2024. "COP29: Azerbaijan Focuses on Environmental Protection and Global Peace Efforts." *The Liberum*. October 22, 2024. Last Accessed on January 23, 2025. https://theliberum.com/cop29-azerbaijan-focusses-on-environmental-protection-and-global-peace-efforts/

Huseynov, Vasif. 2024. *Revisiting the 44-Day War: An In-Depth Retrospective*. Baku: AIR Center. Last Accessed on September 2, 2025. https://aircenter.az/uploads/dfQFzhRwSaNr.pdf

Jafarova, Esmira. 2020. "Is Azerbaijan a 'Middle Power'?" Modern Diplomacy. May 16, 2020. Last Accessed on May 23, 2025. https://moderndiplomacy.eu/2020/05/16/is-azerbaijan-a-middle-power

Koolaee, Elaheh, and Ahmad Rashidi. 2024. "The Zangezur Corridor and Threats to the Interests of the Islamic Republic of Iran in the South Caucasus." *Caucasus Analytical Digest (CAD)*136. DOI: 10.3929/ethz-b-000657553

Kuik, Cheng-Chwee. 2008. "The Essence of Hedging: Malaysia and Singapore's Response to a Rising China." *Contemporary Southeast Asia* 30 (2): 159–185. Last Accessed on September 3, 2025. https://www.researchgate.net/publication/265754648_The_Essence_of_Hedging_Malaysia_and_Singapore's_Response_to_a_Rising_China

Mammadli, Matin. 2023. "Multi-Vector Foreign Policy of Heydar Aliyev." *AIR Center*. Last Accessed on September 4, 2025. https://aircenter.az/en/single/the-multi-vectored-foreign-policy-of-heydar-aliyev-1194

Miholjčić Ivković, Nina. 2023. "Azerbaijan–Israel Strengthened: Partnership and Its Implications for Regional Geopolitics." *Caucasus Strategic Perspectives (CAD)* 4 (2): 49–59 Last Accessed on May 5, 2025. https://www.researchgate.net/publication/376754219_Azerbaijan-Israel_Strengthened_Partnership_and_its_Implications_for_Regional_Geopolitics

Morgenthau, Hans J. 1948. *Politics Among Nations: The Struggle for Power and Peace*. New York: Alfred A. Knopf. Last Accessed on September 2, 2025. https://archive.org/details/in.ernet.dli.2015.74487/page/n8/mode/1up

Nasibov, Murad. 2024. "Hedging Foreign Policies." *Baku Dialogues: Policy Perspectives on the Silk Road Region.* Last Accessed on September 3, 2025. https://bakudialogues.idd.az/articles/hedging-foreign-policies-07-08-2024

Omuraliev, Kubanychbek. 2024. "The Organization of Turkic States: A Decade and a Half of Unity and Progress." *Opinion*. October 3, 2024. Last Accessed on September 2, 2025. https://www.aa.com.tr/en/opinion/opinion-the-organization-of-turkic-states-a-decade-and-a-half-of-unity-and-progress/3350488

Organization of Turkic States. 2022. "4th Young Leaders Forum of the Organization of Turkic States Held in Bukhara." June 5, 2022. Last Accessed on September 2, 2025. https://www.turkicstates.org/en/news/4th-young-leaders-forum-of-the-organization-of-turkic-states-held-in-bukhara-and-bukhara-declared-as-the-first-youth-capital-of-the-turkic-world

Parezanović, Marko M. 2022. „Rat u Nagorno Karabahu 2020. godine – uloga svetskih i regionalnih sila." *Baština* 32 (58) 113–136. DOI: 10.5937/bastina32-40515

Poghosyan, Benyamin. 2022. "US Policy in the South Caucasus Prior to and After the 2020 Karabakh War in the Context of the Evolving Regional and International Geopolitics." *Journal of Political Science: Bulletin of Yerevan University* 1: 36–50. DOI: 10.46991/JOPS/2022.1.3.036

President of the Republic of Azerbaijan. 2021. "Shusha Declaration on Allied Relations between the Republic of Azerbaijan and the Republic of Turkey." June 16, 2021. Last Accessed on September 2, 2025. https://president.az/en/articles/view/52122

Republic of Armenia, Ministry of Foreign Affairs. 2023. "Press Release: Armenia–Azerbaijan Negotiations in Washington." June 29, 2023. Last Accessed on September 2, 2025. https://www.mfa.am/en/press-releases/2023/06/29/arm_aze/12069

Shafiyev, Farid. 2007. "Ethnic Myths and Perceptions as a Hurdle to Conflict Settlement: The Armenian-Azerbaijani Case." *The Caucasus & Globalization* 1 (2): 57–70.

Souleimanov, Emil A., and Michael Baranovsky Kraus. 2017. *Iran's Azerbaijan Question in Evolution*. Washington DC: Central Asia–Caucasus Institute and Silk Road Studies Program. Last Accessed on September 2, 2025. https://www.silkroadstudies.org/resources/pdf/SilkRoadPapers/2017-souleimanov-kraus-irans-azerbaijan-question-in-evolution.pdf

TANAP. n.d. "TANAP." Last Accessed on September 2, 2025. https://tanap.com/en/tanap-project

The Observatory of Economic Complexity [OEC]. n.d. "Azerbaijan / Turkey." Last Accessed on September 2, 2025. https://oec.world/en/profile/bilateral-country/aze/partner/tur

The State Statistical Committee of the Republic of Azerbaijan. 2023. The Foreign Trade of Azerbaijan. Last Accessed on September 2, 2025. https://www.stat.gov.az/source/trade/?lang=en

Trend. 2022. "'TurAz Eagle – 2022' Joint Flight-Tactical Exercises wraps up in Azerbaijan [PHOTO/VIDEO]." *Azernews.az*. September 10, 2022. Last Accessed on September 2, 2025. https://www.azernews.az/nation/199167.html

Valiyev, Anar. 2017. "Azerbaijan's Foreign Policy: From Balancing to Multivectorism." In *Eastern Voices: Europe's East Faces an Unsettled West*, eds. Daniel S. Hamilton and Stefan Meister, 133–149. Washington DC: Center for Transatlantic Relations.

Waltz, Kenneth N. 1979. *Theory of International Politics*. Reading: Addison-Wesley.

World Bank. 2023. *Middle Trade and Transport Corridor: Policies and Investments to Triple Freight Volumes and Halve Travel Time by 2030.* Washington DC: World Bank. Last Accessed on September 2, 2025. https://openknowledge.worldbank.org/entities/publication/4d781a1f-ace4-420d-9192-964bbc98b0d1

Yavuz, M. Hakan, and Michael M. Gunter. 2023. *The Karabakh Conflict Between Armenia and Azerbaijan*. Cham: Palgrave Macmillan. DOI:10.1007/978-3-031-16262-6

Данијела Николић*

Факултет за међународну политику и безбедност, Универзитет „УНИОН – Никола Тесла”, Београд

Виолета Рашковић Таловић**

Факултет за међународну политику и безбедност, Универзитет „УНИОН – Никола Тесла”, Београд

ВИШЕВЕКТОРСКА СПОЉНА ПОЛИТИКА АЗЕРБЕЈЏАНА***

Резиме

Рад анализира развој вишевекторске спољне политике Азербејџана у последњих десет година, са акцентом на стратегије које држава користи како би максимизовала своју безбедносну, економску и дипломатску корист у сложеном регионалном окружењу. Истраживање показује да Азербејџан примењује мултивекторски приступ као свесно портфолио-управљање ризиком, избегавајући потпуну зависност од једног глобалног или регионалног актера. У раду је развијен теоријски оквир заснован на реалистичким концептима анархије и националних интереса, као и на теоријама хеџинга и омнибалансирања. Ови концепти омогућавају разумевање вишевекторности као активне стратегије мале и средње државе у региону Јужног Кавказа, која комбинује диверзификацију безбедносних, економских и дипломатских веза како би се одржала флексибилност у међународним односима. Резултати анализе показују да Азербејџан у последњој декади успешно балансира односе са Русијом, Турском, Ираном, САД-ом и другим регионалним актерима. Систематска употреба мултивекторности омогућава Бакуу да утиче на регионалну стабилност, заштити своје економске интересе, посебно у енергетском сектору, и истовремено

* Имејл: ninadaca@yahoo.com; ORCID: 0009-0005-1824-3398

** Имејл: violeta-talovich@mail.ru; ORCID: 0009-0009-7518-8603

*** Овај рад представљен је у оквиру конференције „Геополитика енергије на Западном Балкану – изазови и могућности”, одржане 27–28. фебруара 2025. године на Институту за политичке студије, у организацији Универзитета „УНИОН–Никола Тесла” и Института за политичке студије, Београд.

одржи дипломатску аутономију. Истовремено, студија указује на ограничења оваквог приступа, као што су ризик од претеране зависности од одређених партнера у критичним ситуацијама и могућност конфликта интереса међу великим силама. Закључно, рад потврђује да је вишевекторска политика Азербејџана стратегија адаптивног и рационалног управљања националним интересима у сложеном регионалном и глобалном окружењу. Она представља пример мале државе која активно користи диверзификацију односа као алат за одржавање суверенитета, безбедности и економске добити. Ово истраживање пружа основу за будуће анализе вишевекторских стратегија других држава у сличним геополитичким контекстима.

Кључне речи: Азербејџан, Јужни Кавказ, вишевекторска спољна политика, геополитика, ресурси, енергетика

* This manuscript was submitted on May 15, 2025, and accepted by the Editorial Board for publishing on November 10, 2025.

UDC 327:316.32(510)
DOI: 10.5937/spm94-60365
Review article

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 241-268

Slobodan Popović[*]
Institute of International Politics and Economics, Belgrade
Dragan Trailović[**]
Institute of International Politics and Economics, Belgrade

WHEN GLOBAL CIVILIZATION INITIATIVE MEETS PRACTICE – SHANGHAI COOPERATION ORGANIZATION[***]

Abstract

Global Civilization Initiative (GCI) was launched in Beijing on March 15, 2023, when China's President Xi Jinping delivered a speech titled Join Hands on the Path Towards Modernization at the CPC in Dialogue with World Political Parties High-level Meeting. For the purpose of this article, we will analyse the Shanghai Cooperation Organization (SCO), since we understand this organization as one of the best examples of implementing the Global Civilization Initiative in practice. This paper argues that the SCO, through its emphasis on multilateralism, non-interference, and a shared vision of regional security and development, embodies key concepts of the GCI's proposed framework for international relations. By analysing the SCO's foundational principles, its evolving institutional mechanisms, and its responses to regional challenges, this paper demonstrates how the organization fosters a

[*] E-mail: slobodan.popovic@diplomacy.bg.ac.rs; ORCID: 0000-0002-8094-8445
[**] E-mail: dragan.trailovic@diplomacy.bg.ac.rs; ORCID: 0000-0001-9707-9867
[***] The paper presents findings of a study developed as a part of the research project 'Serbia and challenges in international relations in 2025,' financed by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia, and conducted by Institute of International Politics and Economics, Belgrade, during the year 2025.

"new type of international relations" characterized by dialogue, mutual respect, and a rejection of zero-sum approaches. The paper highlights the SCO's role in promoting cultural exchange and civilizational diversity, thereby offering a practical example of how the GCI seeks to build a more inclusive and harmonious global order. We structured our research in this direction for several reasons. First, China launched both the Shanghai Cooperation Organization and the Global Civilization Initiative. Second, both are rooted in Chinese values, norms, and interpretations of international relations, security order, modernization, and the idea of building a shared future. Third, the SCO and GCI promote a framework in which civilizational differences are seen as drivers of mutual learning and solidarity, rather than as fault lines to justify the imposition of one's own values, political systems, or economic models, or to fuel ideological confrontation. In this sense, both the SCO and GCI embrace cooperation as a guiding principle in shaping international relations. Both initiatives demonstrate that a strong sense of equality among actors contributes directly to a shared sense of togetherness. Last but not least, we consider it significant that both also serve China's broader strategic interests – diplomatic, economic, security-related, developmental, and political.

Keywords: Shanghai Cooperation Organization, Global Civilization Initiative, diplomacy, China, community, shared future, true multilateralism

INTRODUCTION

Undoubtedly, the world is changing. The dynamism of international relations is becoming perplexed, reaching the level of impossibility to predict any future trend. Simultaneously, this "catch me if you can" dynamism and "fast turning pages" trends are bringing many new opportunities, but an unknown imbroglio that requires creativity, resilience, fast and sustainable (re)actions. Individuals, states, regions, and the globe are facing daily challenges. These challenges require constant development of power to adapt to the unexpected novelties. In an era of uncertainty, resilience is the key to shaping the future. In the wider picture, and if we look at history, these challenges show deeper incongruences with the post-Cold War international order because it

was not built by following the manners of diplomacy.[1] On the contrary, it was built by abiding by the realpolitik principles of winner takes all, excluding and making other stakeholders feel unequal in the international order. This behaviour of preserving and reinforcing such an international order continued even after some states of the Global South reached the position of power that requires a redefinition of the North-South politico-economic relations.

Facing with the fact that its further development will be contained through institutions, geopolitical ambitions, military deployments, and geoeconomic restructuring of global production and supply chains by some states, primarily the political West, China's officials posed several priorities in employing its public and emancipatory diplomacy. As the first task, China's diplomatic apparatus must show the globe that China's diplomatic theory and practice, economic resources, and strategic beliefs are the wisdom that the international society lacks to better itself. Furthermore, Chinese wisdom is pivotal in transforming global governance into the space of equality, mutual learning, Mencius optimality, relationalism (Qin 2012, 70–84; Qin 2018; Qin 2020a), and togetherness. Furthermore, in threading global partnerships and networking the globe, Chinese leaders are insisting on the fact that the practice of its foreign policy is based on the position of a strong defender and active builder of international rule of law (Ministry of Foreign Affairs of the People's Republic of China [MFAPRC] 2014) and defender of international order and practitioner of Multilateralism (MFAPRC 2018).

Following this diplomatic practice, China initiated many political, economic, and security interconnecting, ambitious, and transformative initiatives (SCO, Belt and Road Initiative [BRI], Global Civilization Initiative [GCI], Global Development Initiative [GDI], Global Security Initiative [GSI]) as an extension of domestic affairs (MFAPRC 2014). At the same time, these initiatives show the Chinese will to become the architect of international cooperation, global development, security order, and a new wave of modernization. China reaffirmed that it abandoned the concept of the Central Empire for China and entered the "world of interdependence".[2] Expectedly, China holds the stance that these initiatives are a crucial part and engines in China's emancipatory

[1] For an in-depth discussion of this perspective, consult Stefanović-Štambuk 2008.
[2] On this issue, useful insights can be found in Mitrović 1995.

diplomacy,[3] aimed towards liberating the international order from hegemonism, exclusivity, unilateralism, and inequality.

One such "liberating" initiative for bettering global governance is the Global Civilization Initiative. It was launched by the China's President Xi Jinping during the CPC in Dialogue with World Political Parties High-level Meeting held in Beijing on 15th of March. In his speech *Join Hands on the Path towards Modernization* delivered for this occasion, China's president Xi Jinping declared that "all civilizations created by human society are splendid" (Xi 2023a).

For the purpose of this article, we analyse the Shanghai Cooperation Organization, as we consider it one of the most illustrative examples of the Global Civilizational Initiative being put into practice. We develop this argument for several reasons, beginning with the evident fact that both the Shanghai Cooperation Organization and the Global Civilization Initiative were launched by China. Second, both are centred on Chinese values, norms, understanding of international relations, security order, modernization, and making the future together in joint hands. As the third reason, we are pointing out that the Chinese emphasize that both SCO and GCI are creating the context in which differences among civilizations will be used as engines for mutual learning, not as fault lines for imposing one's own values or models on others and for stoking ideological confrontation (Xi 2023a). In other words, leaders of the Communist Party of China believe that China, by its behavior, stimulates coordination and cooperation with and within the international community in a manner that every challenge is diplomatized and inter-civilizational dialogue is further developed and grounded on proper values – Chinese wisdom.[4] Thus, both propose cooperation as the golden rule in weaving international relations. For Chinese scholar Qin Yaqing (Qin 2020b, 37), the history of human evolution is a history of cooperation. This is followed by the reason that both consider that the sense of equality fosters the sense of togetherness. Besides this, for the reason we are taking the obvious and legitimate fact that both GCI and SCO are in service of achieving China's instrumental goals – diplomatic, economic, security, developmental, and political.

[3] On China's emancipatory diplomacy, see also Stefanović-Štambuk and Popović 2024, 415–423.

[4] On this point, see also Zongze 2018.

This article relies on a qualitative and interpretative approach. The research is based primarily on the analysis of official documents of the Shanghai Cooperation Organization (such as its Charter, declarations, and summit statements), speeches by Chinese political leaders, and relevant secondary scholarly literature. The method applied combines textual and discourse analysis in order to trace how the principles of the Global Civilization Initiative are reflected in the institutional development, diplomatic practice, and normative framework of the SCO. Such an approach enables not only the identification of conceptual overlaps between the Initiative and the Organization, but also an assessment of their practical implementation.

DIFFERENCES AMONGST CIVILIZATIONS – FAULT LINES OR A GREAT WAY TOWARDS HARMONY

The world is changing. International order is changing. In a global village, differences are emerging, and at the same time they are converging into unified wholeness that shares a global future. The purpose and nature of war are changing. Manners of knitting the international order are changing. The nature and structure of security are changing.

Having in mind all these changes, we cannot help but question whether diplomacy possesses the capacity to sustain the world. Is diplomacy sustainable[5] in the world of dramatic changes, significant differences, but even bigger opportunities? Leaders of China understand that changes have reached an unseen scale in a century (State Council Information Office of China [SCIO] 2023) with obvious transformative capacity. Document *A Global Community of Shared Future: China's Proposals and Actions,* released by the State Council of the People's Republic of China in 2023, details China's perspective on global changes as follows: "This is an era of promise, and an era of challenges. At yet another crossroads in history, we have to choose between unity and division, between opening up and closing off, between cooperation and confrontation. With the overall interests of humanity at stake, this choice tests the wisdom of all countries" (SCIO 2023).

Although this is a period of unprecedented changes, China defines it as a period of strategic opportunities.[6] Thus, in a speech delivered to

[5] For more comprehensive treatments of this theme, see Stefanović-Štambuk 2010a, 645–673.

[6] For a detailed discussion, see Stekić 2023 and Lađevac 2020.

the 19th National Congress of the Communist Party of China, Xi Jinping, besides ushering in "five shoulds",[7] recommended not to give up on China's dreams because the reality is too complicated, and China should not stop pursuing its ideals because they seem out of reach (Xi 2017). However, it is expected that China uses the mentioned changes and period of strategic opportunities to internationalize Chinese solutions, bringing in the name of overcoming and translating these global changes/challenges into opportunities. Simultaneously, China is trying to translate these solutions into an international consensus, which should be another channel for broadening *and broadcasting* [added by the authors] China's development path (Zongze 2018, 14–15). Leaders of China perceive the internationalization and globalization of Chinese wisdom as an urgent task for global stability. This would be a normal process of internalizing China, because China became independent and prosperous, with the obvious wish to share its wisdom. As China becomes stronger, its wish to better the world becomes even more powerful. For China, internationalization and globalization of its solutions are equal to sharing development opportunities with the world.

The USA, similar to China, shares the view that the world is changing and that the global environment is becoming even more puzzled. For example, in the document titled *2024 Annual Estimate of the Strategic Environment,* the USA recognizes that the international order is a complex and ever-changing global environment (Strategic Studies Institute, US Army War College 2024, 4). In the same document, one of the urgent tasks for securing global and national security is defined as the need for the US Army to stay in a condition of constant vigilance, strategic planning, and commitment to the forefront of military innovation (4). In this global complexity, China's initiatives and strategic

[7] "Five shoulds" are: "We should respect each other, discuss issues as equals, resolutely reject the Cold War mentality and power politics, and take a new approach to developing state-to-state relations with communication, not confrontation, and with partnership, not alliance. We should commit to settling disputes through dialogue and resolving differences through discussion, coordinate responses to traditional and non-traditional threats, and oppose terrorism in all its forms. We should stick together through thick and thin, promote trade and investment liberalization and facilitation, and make economic globalization more open, inclusive, and balanced so that its benefits are shared by all. We should respect the diversity of civilizations. In handling relations among civilizations, let us replace estrangement with exchange, clashes with mutual learning, and superiority with coexistence. We should be good friends to the environment, cooperate to tackle climate change, and protect our planet for the sake of human survival" (Xi 2017, 52–53).

steps as a part of diplomatic theory and practice may trigger military conflict in potential hot spots in the Indo-Pacific (6–33).

Unsurprisingly, an analogous understanding of the international order and China's behaviour is shared by the NATO. Namely, in the document *NATO 2022 Strategic Concept* adopted at the NATO Summit in Madrid, 29 June 2022, this organization holds that the threats humanity faces are global and interconnected (North Atlantic Treaty Organisation [NATO] 2022, 3). Among all the threats occurring during kaleidoscopic shifts, special attention is given to the diplomatic, economic, and security practices of China. According to the same document, China's global behaviour is defined by the features of the coercive policies that challenge NATO interests, values, and security (NATO 2022, 5). In that behaviour, China remains opaque about its strategy, intentions, and military build-up (5).

According to Huntington, some scholars in this new phase of accelerated global dynamism have eagerly offered visions of the future – the end of history, the return of traditional rivalries between nation-states, and the decline of the nation-state under the conflicting pressures of tribalism and globalism – visions that catch aspects of the emerging reality (Huntington 1993, 22).

However, one of those phrases for catching the reality was also the hypothesis posed by Samuel Huntington. According to this hypothesis, "the fundamental source of conflict in this new world will not be primarily ideological or primarily economic. The great divisions among humankind and the dominating source of conflict will be cultural. Nation states will remain the most powerful actors in world affairs, but the principal conflicts of global politics will occur between nations and groups of different civilizations. The clash of civilizations will dominate global politics. The fault lines between civilizations will be the battle lines of the future" (Huntington 1993, 22).

Analyzing the hypothesis posed by this scholar, we can perceive that the role of civilizations and civilization identity will gain bigger importance in shaping the world's future, including the future of world politics, international society, global governance, global security, and developmental expectations and limitations. But, for the same scholar, differences amongst civilizations are not the tools to develop the trends of mutual learning, respect, mutual empowering, and development through togetherness. On the contrary, civilizational differences are sources that empower conflicts and deepen rifts among nations. The

same scholar believes that the ongoing interactions and continuous communication from personal to state level may only intensify pre-existing "[…] differences and animosities stretching or thought to stretch back deep into history" (Huntington 1993, 26). Guided by this Hegelian logic, the approach "one size fits all" was selected as the best solution for overcoming civilization differences and melting civilization richness into the pot of the Liberal International Order. Simultaneously, a civilization that did not fit into or fitted only partly into the values, norms, visions, and actions of the Liberal International Order was sanctioned in different ways, from defining the same civilization as a challenger and threat to global security, like Rogue State, or was marginalized when the rules of the international order were defined. This was the dominant practice and mentality when the network of global partnerships in the post-Cold War era was established.

China, which nurtures an entirely different political, economic, and value system compared to the West, in this kind of architecture of the international order, in 1978 initiated the "reforms and opening up policy".[8] Despite all these differences, China recorded tremendous results. West perceived this as China is taking off its silk gloves. However, China also faces with great and new challenges.[9] All those results and challenges, which were becoming more notable and more complex as an obvious manifestation of increasing China's power, were gathered

[8] In the speech delivered during the First Session of the 13th National People's Congress on March 20, 2018, Xi Jinping defined the policy of reforms and opening up in the following way: Reform and opening up represented a great awakening for the CPC, an awakening that facilitated great theoretical and practical innovations for the Party. Reform and opening up is a great revolution in the history of the Chinese people and Chinese nation, one that propelled a tremendous transformation in the cause of socialism with Chinese characteristics (Xi 2023b).

[9] By analysing China's advantages and disadvantages, its leaders during the Two Sessions gathering brought a new concept regarding China's development – High Quality Development. This new type of development is based on new productive forces. According to official sources about new productive forces and new development we can read the following: "With innovation playing the leading role, new quality productive forces mean advanced productivity that is freed from the traditional economic growth mode and productivity development paths, features high-tech, high efficiency and high quality, and comes in line with the new development philosophy" (SCIO 2024).

in a holistic term known as "China Dream"[10] manifested through "Two Centenary" (Popović 2024, 48–50).[11]

China brought another shock to the Western world. China started its diplomatic revolution (Stefanović-Štambuk 2010b, 299–311) in 1996 when it initiated the Shanghai Five mechanism as the outrunner of the Shanghai Cooperation Organization. Twenty years later, China is speeding up its diplomatic revolution. China believes that its diplomatic revolution is crucial in bettering the international order and pursuing the benevolent course of major country diplomacy. We are setting this commencement point for China's diplomatic revolution, because that was the China's very first materialization of its endeavors to multilaterally gather countries and established a new type of relationship. Gathering was done by using civilization differences not as the "fault lines", but as the cornerstone of the Shanghai Spirit. In the article 4 of the *Declaration on the Establishment of the Shanghai Cooperation Organization* the Shanghai Spirit is defined in the following way: "The "Shanghai Spirit" formed in the process of the Shanghai Five development, described by mutual trust, mutual benefit, equality, joint consultations, respect for cultural diversity and aspiration for collective development, is an invaluable asset gained by the countries of the region over the years of cooperation. It will grow and in the new century will become a norm in relations among the states members of the Shanghai Cooperation Organization" (Shanghai Cooperation Organization [SCO] 2001).

[10] China dream is a great rejuvenation of China's nation. China's leaders for the China dream state the following: "The pursuit of beautiful dreams resonates far and wide across countries and nations. The Chinese Dream is about pursuing strength of the country, rejuvenation of the nation and happiness of the people, and the dream of the world is about enduring peace and common prosperity. These two dreams can always be pursued together. The Chinese Dream, a dream about peace, development, cooperation, and win-win, is closely linked with the beautiful dreams of people in countries around the world. The Chinese people and people elsewhere in the world are bound together by the shared aspirations for a better life. Supporting and helping each other along the way, they are making progress hand in hand in realizing the dreams" (MFAPRC 2022).

[11] Two Centenary as a guiding principle of the Chinese development is defined as following: "China has set the Two Centenary Goals, one to be achieved around 2020, the centenary of the founding of the CCP, by doubling China's GDP and per capita income from the levels of 2010 and comprehensively building a moderately prosperous society to the benefit of well over one billion people. The other is to develop China into a prosperous, powerful, democratic, civilised and harmonious socialist modernized country by the mid-21st century, i.e., the centenary of the founding of the PRC" (Ding 2016, 80).

Having in mind this kind of definition of the major norm of the SCO, the purpose of this article is more than obvious. Thus, we are analysing the Shanghai Cooperation Organization as one of the best practices of implementing the Global Civilization Initiative in reality.

GLOBAL CIVILIZATION INITIATIVE – CHINA'S DIPLOMATIC REVOLUTION FOR PUZZLING THE INTERNATIONAL ORDER

Global Civilization Initiative is the third initiative in China's ambitions to share the wisdom, theories, and practice of emancipatory diplomacy (Stefanović-Štambuk and Popović 2024, 424–425). The first two announced were the Global Security Initiative and the Global Development Initiative, respectively. Hence, the GCI should be understood as the injection of fresh dynamism, that is, inter-civilizational dialogue into building and safeguarding global development cooperation and global security cooperation (Liu 2023). For some scholars, these three initiatives create a cornerstone for transforming the world into the garden of flourishing civilizations (Liu 2023).

When announcing the Global Civilization Initiative, Xi Jinping underlined a famous Chinese saying: *A single flower does not make spring, while one hundred flowers in full blossom bring spring to the garden* (Xi 2023a). Besides obviously insisting on diversity that has to be respected and nurtured through dialogue and consultations, Xi Jinping confirmed China's official understanding that the future of all countries is connected (Xi 2023a). According to Chinese diplomatic practice, the world is big enough to accommodate the common development and progress of all countries. It is entirely possible for different civilizations to prosper together and inspire each other through mutual learning on an equal footing. The future is shared by and for all countries (Xi 2024b).

Hence, the whole of humanity has one home – our world. All civilizations together should treat it like the most beautiful garden. China advocates four action blueprints to make this view a reality: "[1.] We advocate the respect for the diversity of civilizations [… 2.] We advocate the common values of humanity. Peace, development, equity, justice, democracy, and freedom are the common aspirations of all peoples […3.] We advocate the importance of inheritance and innovation of civilizations […4.] We advocate robust international people-to-people exchanges and cooperation. Countries need to explore

the building of a global network for inter-civilization dialogue and cooperation" (Xi 2023a).

Analyzing these "four advocates", we understand that China demonstrates the pattern of behaviour within the international community that promotes commonality, togetherness, and respecting the fact that all civilizations are rooted in their unique cultures (Xi 2023a). This diversity for China should not only be perceived as a richness in traditions and history, but also as a reality and vital quality for promoting inter-civilizational dialogue, mutual learning, global progress, and a people-centered approach. As such, differences among civilizations should not be securitized, i.e., presented as an existential threat to one's own or any other civilization (Xi 2023a). Furthermore, the destiny and shared future of all civilizations is not to clash. Thus, China insists that these three global initiatives (GCI, GDI, GSI), from a holistic approach, bring the future to the globe which is already marching defuturism.[12]

Thus, having in mind that the world is, once again, on the crossroad of selecting the path between a shared and bright future, on the one side, and defuturism on the other, China is branding the GCI as a collection and unification of common aspirations of countries around the globe for promoting equality, inclusiveness and diversity of world civilizations when building a community of a shared future (Liu 2023). In the same style of its diplomatic theory and practice and for developing inter-civilizational dialogue, China underlines the fact that there is no country superior to anther, and there is no such thing as an inferior or superior civilization, or a value above other values (Liu 2023). Thus, the illusion of some civilization pronouncing itself as a universal civilization or being more important than other civilizations will easily fall through. At the same time, this illusion refers to the moment when some civilization proclaims its values as universal values that must be shared by the whole world in the form they understand as the best. In the same context, using ideologies and values to draw lines will only serve to create divisions and confrontations. This kind of behaviour "[…] will only bring disastrous results upon the world and blemish the civilizations" (Liu 2023). This is an exceptional, or one scholar approach, regarding theoretical understandings of knitting relations amongst states, nations, and civilizations. For example, other Chinese scholars in their analysis also take the stance that civilizations: "should not be divided

[12] On the idea of defuturing, see Fry 2020.

into high or low, and civilization should not have a so-called universal model. Otherwise, if it is monopolized by the so-called unified way, then it is impossible for human beings to develop for a long time, and many tragedies and disasters will soon follow" (Xu, Li and Hu 2023).

GCI reflects China's diplomatic theory and practice based on civilizational inclusiveness and respecting civilizational diversity on a global level. Furthermore, it represents how China intertwines its values and civilizational identity into international society (*Beijing Review* 2024). By doing so, China is creating a strategic context in which it its civilizational identity and values will be perceived as a tool for diplomatizing regional and global security challenges and a pivotal engine for economic development, which is craving for new capital. Staying on the level of strategic thoughts, China is creating a convenient environment for its cherished peaceful development as a pivotal part of the major country diplomacy.

Recognizing China's unselfish efforts in making the inter-civilizational dialogue a question of global importance was confirmed when China proposed, and the UN General Assembly during the 78th Session held on June 7, adopted resolution A/78/L.45. The same resolution stipulated June 10 for the International Day for Dialogue among Civilizations. When making the text of resolution Chinese officials recalled the "[…] previous resolutions of the General Assembly on promoting understanding, development and exchange among civilizations and cultures, in particular its resolution 53/22 of 4 November 1998, by which the Assembly proclaimed the year 2001 as the United Nations Year of Dialogue among Civilizations, and its resolution 56/6 of 9 November 2001 on the Global Agenda for Dialogue among Civilizations and the Programme of Action and its objectives […]" (United Nations General Assembly [UNGA] 2024). Besides that, recalling to the previous documents included following documents: "resolution 51/95 of 12 December 1996, by which the Assembly proclaimed 16 November as the International Day for Tolerance, its resolution 72/130 of 8 December 2017, by which it declared 16 May as the International Day of Living Together in Peace, its resolution 75/200 of 21 December 2020, by which it proclaimed 4 February as the International Day of Human Fraternity, and its resolution 75/309 of 21 July 2021, by which it proclaimed 18 June as the International Day for Countering Hate Speech" (UNGA 2024).

The very text of the resolution underlined the following: "all civilizations celebrate the unity and diversity of humankind and are enriched and have evolved through dialogue with others, respect for human rights as well as the need to acknowledge and respect the richness of all civilizations and to seek common ground among civilizations in order to address comprehensively challenges facing humanity" (UNGA 2024).

Analysing the text of the resolution, with particular focus on this part, we notice considerable similarities with the speech of Xi Jinping when he institutionalized the Global Civilization Initiative as the Chinese wisdom in threading inter-civilizational dialogue. Both resolution and speech emphasize the importance of dialogue, consultation, and mutuality as means of nurturing diversity and commonality whilst holistically addressing dynamic challenges facing humanity. Both document and speech underline that civilizational diversity improved the evolution of humankind. Hence, civilizational diversity contributed to the development of different types of protection and diverse, rich understandings about the same concepts – human rights, democracy, development, and sustainability, to name a few.

Since the GCI, as its name suggests, extends to the global scale, China expressed its will to be more active and innovative regarding the reforms of global governance. According to Chinese diplomatic practice, this reform is equal to bettering the international order and making it the "environment of equal rights, equal opportunities and fair rules for all" states, regardless of size and power (Xi 2023a). Since, there is one planet Earth and that China is not isolated from the shared future of the humanity, Chinese leaders considered as a high priority obligation "[…] to provide new opportunities for world development, add new impetus to humanity's exploration of paths towards modernization, and make new contributions to the theory and practice of humanity's modernization" (Xi 2023a). All these efforts follow China's new progress in its modernization (Xi 2023a). However, to be done in the most suitable way, China must not change its nature, abandon its system, or be trapped in a "one size fits all" approach. For China, its activity and steadiness are the "[…] key to the opening up of a new prospect of enhanced exchanges and understanding among different peoples and better interactions and integration of diversified cultures" (Xu, Li and Hu 2023). Chinese President Xi Jinping insists that China must "stay committed to the right direction, right theories and the right path" (Xi 2023a).

Of course, some scholars proposed an understanding of the Global Civilization Initiative as the reinvigoration of the Sino-centric system and anti-democratic values, alarming the world by bringing to the forefront the China threat theory. For Stefanović-Štambuk and Popović: "In truth, the "China threat" theory is hardly a theory. It is more of a label for a body of thoughts, less theoretically based, and vastly ideologically biased, put up to inform, activate, and issue a rallying call to close ranks behind the US and tighten the sense of (Western) identity around its core values allegedly under jeopardy from Chinese ambitions to ignite the joint re-evaluation of the decaying base of international politics" (Stefanović-Štambuk and Popović 2022, 9–10). However, scholar Evan Ellis (Ellis 2023) in the online journal *Diplomat* stated that the GCI is a synonym for ambiguity. This scholar criticized Xi's speech in the sense that the Chinese president was speaking about the "common aspiration" of humanity and not rights when he was talking about "peace, development, equity, justice, democracy, and freedom" (Ellis 2023). He presented this as an obvious part of the China threat theory since the GCI advocates against a world in which those concepts can have meaning through united efforts to call out and collectively act against those who violate them (Ellis 2023). It is very particular in his research that he criticized what are the common aspirations for China to be a legitimate obligation for which countries must refrain from imposing their own values or models on others (Ellis 2023). There is an impression that every country has a duty to impose its values and beliefs on other countries as the legitimate right in spreading its power, strategic ambitions, diplomatic practice, geopolitical vectors, and geoeconomic interests. Compared to this, the GCI promotes relativism of values (Ellis 2023). As a self-serving effort, it is in service of achieving Chinese ambitions to dismantle the "rules based international order" (Ellis 2023) and promotes inter-civilizational dialogue without a "one size fits all" approach.

GLOBAL CIVILIZATION INITIATIVE IN PRACTICE - SHANGHAI COOPERATION ORGANIZATION

The Shanghai Cooperation Organization was born in June 2001. It was born when Islam Karimov, then President of Uzbekistan, signed the document of accession of this country to the Shanghai Five mechanism. This historical moment happened on June 15, 2001, during the meeting of the heads of state of the then members of the Shanghai Five mechanism.

This signature included not only Uzbekistan in the teamwork along with China, Russia, Kazakhstan, Tajikistan, and Kyrgyzstan: it was also the very first step in creating a new reality and interconnectedness in the space of Central Asia. Through institutionalization of diplomatic, economic, and infrastructural unlocking the landlocked geopolitical character of Central Asia, China brought completely new dynamism and understandings of geopolitical concepts, Heartland and Rimland (Popović 2024).

As the years passed, the SCO expanded its influence and power regarding geography, scope, form, and content. SCO started to dispose of the types of power that were earlier exclusively reserved for the West. As the disposal of power and influence increased, SCO promoted innovative theory and practice of creating a new type of international relations based on dialogue, consultation, shared security, equality, fairness, and justice. SCO and its guiding principle, the Shanghai Spirit, were recognized as sources of innovative theory and practice in establishing contemporary international relations. Contemporary international relations is the non-confrontational model. This kind of model of the international relations calls for discarding the Cold War mentality, bloc politics, and transcending ideological differences (SCO 2006). This is a simultaneous and direct implementation of the Shanghai Spirit in developing the inter-civilizational dialogue within the SCO region.

In the rich spectra of the SCO activities, promoting cultural, religious, and civilizational diversity was highly positioned and prioritized since the very foundation of the organization. This was also institutionally and normatively supported by the Shanghai Spirit. In such a manner, diversity is presented as a tool of mutual learning and developing dialogue. As such, diversity nurtured in the SCO practice is the channel for creating a sense of togetherness and equality. For example, according to the Charter of the Shanghai Cooperation Organization, the founding of this international organization is based on historically established ties between its peoples (SCO 2002). Ties are used for making the knots, not battlefields. Understanding that it is better to invest money in development instead of in wars, member states, as one of the goals set, strengthen mutual trust, friendship, and good neighborliness in the SCO Region (SCO 2002). This stance was encouraged through decades-long practice of development of the SCO system of diplomatic theory and practice in making the SCO region of the space of indivisible security (SCO 2015). There are strong pledges in

many documents that the SCO member states will remain friends from generation to generation and will never be enemies against one another (SCO 2006). Member states constitute peace by upholding consultations, dialogue, and diversity instead of stimulating a weapons race. However, the newest events between India and Pakistan questioned the feasibility of the given pledges and the viability of the SCO in real time and space.

This understanding of the interstate relations in making the SCO region a space of peace was once again accentuated in the Declaration of the member states of SCO issued for the occasion of the fifth anniversary of the SCO in 2006: "Diversity of civilization and model of development must be respected and upheld. Differences in cultural traditions, political and social systems, values and model of development formed in the course of history should not be taken as pretexts to interfere in other countries' internal affairs. Model of social development should not be 'exported'" (SCO 2006).

In its work, we can see that the SCO is not imposing animosity and uniformity as a prerequisite for cooperation, mutual respect, making peace, and protecting sovereignty. Instead, this is done by encouraging diversity, harmonization of interests, and the diplomatization of security issues. SCO member states found these to be the most suitable ways to minimize the possibilities of misunderstandings that can be transformed into skirmishes. An outbreak of any scuffle between the SCO members can jeopardize the stability of the SCO Region by questioning the SCO cooperative system and the SCO norms that make the SCO Region a space of predictable dynamism. Besides this, any misunderstanding that can reach the level of diplomatic conflict opens the door for third-party involvement not only in the SCO affairs, but also in the domestic affairs of the SCO member states. Although SCO, as an organization that shares the norms of collectivity and values of inclusion, has strategic interests primarily mirrored as preventing the strategy of containment of Sino-Russian interests that the USA, through the model of hub and spokes and military alliances, is developing in the Asia-Pacific region (Shambaugh 2004, 72–77). Thus, and completely expected, strategic interests are intertwined with the Shanghai Spirit as the workflow of the SCO and the development of the SCO model of diplomacy. In mutuality, American interests cannot take the lead; they will be intertwined into the network of interests of the SCO region if we follow the typology of the SCO documents. Making the organization's space a region of peace, cooperation, sustainable development, prosperity, and harmony must be

understood as the crucial strategic interest (SCO 2021). If this strategic interest is not achieved, the direct SCO contribution to the respect of cultural and civilizational variety of the modern world (SCO 2007) will be questioned. SCO's credibility will also be questioned.

Throughout time, states that established different political and economic systems on their territories, followed by different systems of values and rooted in different cultures, took different types of participation and cooperation with the SCO member states, dialogue partners, and observer states. So far, there are nine SCO member states – the Republic of India, the Islamic Republic of Iran, the Republic of Kazakhstan, the People's Republic of China, the Kyrgyz Republic, the Islamic Republic of Pakistan, the Russian Federation, the Republic of Tajikistan, and the Republic of Uzbekistan. At first glance at the organizational structure of the SCO and the pace of inclusion of new member states, we can conclude that respecting diversity is entrenched in the SCO workflow and the SCO diplomatic system. Thus, SCO does not support the "one size fits all" approach. On the contrary, establishing such an organization to institutionally develop and harmonize the interests of many countries in spectacular SCO diversity, implies not only that the inter-civilizational dialogue, valuing civilizational beauty, and richness and equality are highly positioned and requested as the manner of doing relations within the SCO Region, but that big number of states cherished these values for doing the interstate relations.

From the above, we can see that the SCO membership structure is very complex. It includes states that have fought direct wars throughout history. Some of the SCO member states are permanent member states of the OUN Security Council, and some of them are recognized and non-recognized nuclear powers. The religion map of the SCO is also sundry, since it embraces Buddhists, Christians, Muslims, and citizens who follow the rules of Confucianism. From the economic perspective, SCO consists of the world's biggest trading countries, the world's biggest emitters and receivers of FDI, both close USA partners, its fiercest opponents, and America's biggest trading partners, the world's biggest energy producers and consumers. From the point of view of physical facts, SCO includes the world's most populous countries and countries with the vastest territories in the world. We are listing all this evidence to demonstrate the level of complexity with which the SCO developed its institutional capacities and rebutted Western predictions of the end of history and clash of civilizations. SCO diplomatic theory

and practice express that diversity and not uniformity create a shared future. Uniformity brings us to the defuturism. For the SCO member states, developing the inter-civilizational dialogue is a bottom line of common, comprehensive, cooperative, and sustainable security, because real security is premised on the security of all countries (Xi 2024a). We are witnessing states voluntarily searching for ways to limit their powers, institutionalize their power, and employ it to make the SCO Region a space of predictability. The normative framework for this was provided by the Shanghai Spirit.

So far, it is apparent that the activities of the SCO, to a large extent, are in line with the ideas, values, and activities of the proposed GCI. Notably, on the level of adopted documents, SCO is giving wide support to achieving the GCI goals, and making the context in which GCI will become a reality. For example, during the 23rd Meeting of the Council of Heads of Government of Member States of the Shanghai Cooperation Organization, the intention of member states to promote the proposal for the adoption of a Resolution by the UN General Assembly in connection with the SCO Initiative "On World Unity for a Just Peace, Harmony and Development" was reaffirmed. Along with this, SCO member states buttressed the propagation of dialogue on the idea of "One Earth. One Family. One Future" (SCO 2024a; SCO 2024b). Besides that, SCO members supported strengthening the global dialogue between religions and cultures within the framework of the Congress of Leaders of World and Traditional Religions (SCO 2024b). In this place, and once again, we notice that the SCO is giving a great support to the realization of the GCI. From the aspect of China's diplomacy, this means making a wider audience for sharing and perceiving China's story as an engine for global development and securing global security. Simultaneously, it means making a more stable context within which China's strategic actions will be understood as a part of peaceful development and not reinvigoration of a China Threat Theory. In this one earth, one home, one future, SCO members are continuously showing their unwavering commitment to respect preserved traditional values and cultural and civilizational diversity, as well as the need to further deepen cooperation in the humanitarian field by introducing new formats of interaction in education, culture, tourism, and sports.

Leaders of the member states understood that sport plays a great part in developing inter-civilizational dialogue, sharing and adopting different cultural heritages, developing an atmosphere of togetherness, and creating

a spirit of fair play. For that aim, Heads of Ministries and Departments responsible for the development of physical culture and sports in the SCO Member States, during the meeting held in Almaty, 24 May 2024, agreed to establish a Working Group on Physical Culture and Sports and hold the SCO Sports Games (SCO 2024b). These results were recognized as the impulse for different initiatives to establish the Association of SCO Sports Organizations. Besides the annual SCO Marathon, the idea to hold the "Silk Road" mini-football tournament for the SCO Member States Cup in futsal on December 9–18, 2024, in Tehran, was unanimously accepted (SCO 2024b). The unique role of physical culture and sport in strengthening solidarity, peace, and inter-civilizational dialogue was recognized. The representatives of the member states will actively work on "promoting development of international sports cooperation on an equal and depoliticized basis, oppose discrimination against athletes on any grounds, including nationality, language, political and other beliefs, national or social origin" (SCO 2024b).

Regarding tourism, member states in Samarkand in 2022 signed the Agreement between the Governments of the SCO Member States on the Development of Cooperation in the Field of Tourism and the SCO Joint Plan of Action in the Field of Tourism for the period 2024–2025. This was a great impetus to institutionalize touristic development and cooperation as one of the economic sectors expected to grow. In line with that, in 2023, a Forum was held in Urumqi when the SCO announced that year for the "SCO Year of Tourism". The discussion was focused on the four thematic areas: tourism potential of the SCO space; tourism potential of the Xinjiang Uygur Autonomous Region; promotion of sustainable tourism; development of tourism as a way to reduce poverty (SCO 2023).

In regard to cooperation in education, member states, along with dialogue partners and observer states, knitted the network of universities within the SCO Region. This network is known as the Shanghai Cooperation Organization University (SCOU). On the official website which is not updated properly, following goals are recognized to be achieved: strengthening mutual trust and good-neighbourly relations between the SCO member countries; development of integration processes in the field of education, science and technology; giving new impetus to the expansion of multilateral educational, scientific and cultural cooperation; expanding opportunities for young people to receive high-quality modern education, and for teachers and scientists to develop scientific contacts; promoting effective cooperation between

the Organization's member countries in the political, trade, economic, scientific, technical and cultural fields (Университет Шанхайской организации сотрудничества n.d.).

For the main tasks of the Shanghai Cooperation Organization University network, representatives of the educational institutions of the SCO Region defined following: expansion of the exchange of pupils, students, postgraduates, doctoral students and research and teaching staff; increasing scientific and academic cooperation; implementation of modern educational methods and technologies; creation of mechanisms for recognition and equivalence of educational documents of the SCO University by SCO member states and the global educational community (Университет Шанхайской организации сотрудничества n.d.). Furthermore, ministers of education of the SCO member states held in Moscow on April 18, 2024, agreed to give priority to building up cooperation within the SCO University, the development of cooperation in the field of academic exchanges, joint scientific work, the organization of vocational education, language training, enhancing youth interaction, as well as digitalization in the field of education in the SCO interested Member States (SCO 2024b).

We can see that institutionally, SCO, as one of the crucial factors in Eurasian security and development, strongly supports the constant communication, flow of people, and sharing of knowledge as the essential parts of the *people-to-people diplomacy.* For such a kind of diplomacy, they hold that contributes to strengthening mutual understanding and cultural and humanitarian ties within the SCO (SCO 2024b). It is raising the level of trust among the member states. Thus, for the SCO member states, it is important not only to have physical connections in the form of infrastructure but also in the form of institutional capacities, sense of togetherness, equality, and respect for diversity. All these different types of interconnectedness in threading relations are forming the inter-civilizational dialogue and psychological bonds in people-to-people diplomacy. Strategically, it contributes to raising the level of predictability of behaviour, stipulating and internalizing the SCO norms into the domestic system of values. Besides this, SCO strategic calculations are becoming part of the strategic interests of the member states. Observer states and dialogue partners are becoming more familiar with the workflow of the SCO. The internalization would be easier. But it is also one of the ways of reinforcing control of surveying methods justified by the need to fight three evils – terrorism, separatism, and religious extremism.

CONCLUSION

By its power and innovative diplomatic theories, China brings new dynamism to international affairs, new impulses into the democratization of global governance, and emancipation of international society.[13] This dynamism is focused on respecting the world's diversity" (Jiang 2012, 40). China's leaders believe that in this way, China is mirroring and transferring its wisdom and managerial skills to better the globe and make new and stronger bonds between man and nature.

China views the current global context as one of accelerated transformation and instability, marked by changes of historic magnitude. These dynamics present humanity with unprecedented challenges, often framed as choices between unity and division, peace and conflict, or cooperation and confrontation. For China, the answer to this riddle is the following: "the people's wish for a happy life is our goal, and peace, development and win-win cooperation are the unstoppable trends of the times" (Xi 2023c).

For the purpose of this article, we demonstrated that the SCO is not only one of the strongest forces in giving birth to the Global Civilization Initiative, but also one of the best practices in implementing the principles and ideas defined within the Global Civilization Initiative. These results were obtained by doing an analysis of the principles and values of the SCO diplomatic theories and practice, and the development of institutional capacities of the same organization. The results of our research also demonstrated that SCO's institutional flaws can be improved by developing new institutional capacities and innovative diplomatic practices and theories.

Strengthening exchanges, constant communication, mutual learning and forging people-to-people diplomacy within the SCO Region, so far, has been positioned as the task of the high priority on the agenda of the SCO institutional development (Xi 2023c). This task is highly prioritized, since all nations in the region aspire to see the harmonious development of different civilizations. This kind of civilizational development contributes to creating the atmosphere of togetherness in diversity. Thus, SCO member states welcome the efforts of all sides to work together in implementing the Global Civilization Initiative, to promote inclusiveness and coexistence among

[13] For a detailed discussion, see Stefanović-Štambuk and Popović 2024, 415–423.

all civilizations, and to increase mutual understanding and friendship among all nations (Xi 2023c).

Considering that the world is at a crossroads of unprecedented changes, the SCO will be affected. However, this gives a great opportunity to the SCO to update existing and develop new institutional capacities and diplomatic theories and practices, which can be done in a more suitable and sustainable manner to nurture inter-civilizational dialogue and civilizational harmonious development, because "all civilizations created by human society are splendid" (*China-India Review* 2023, 41), and people-to-people diplomacy.

China has been actively shaping a vision for international relations that champions diversity and mutual respect. At its core, this vision emphasizes a profound respect for civilizational diversity, advocating for the acknowledgment of numerous distinct cultures and societal models, rather than imposing a singular set of "universal" norms. This leads to a perception, which China frequently articulates, of a "normative monopoly" where non-Western states may feel pressured to conform to Western-derived frameworks. China perceives a significant disparity between the stated pluralism of the current world order and its de facto governance by Western powers. Through its various global initiatives, Beijing actively seeks to bridge this gap, promoting alternative forums, funding mechanisms, and a more inclusive normative language, even while engaging with the very institutions it critiques.

REFERENCES

Beijing Review. 2024. "Dialogue Among Civilizations." *China Focus*. June 18, 2024. Dialogue among Civilizations. http://www.cnfocus.com/dialogue-among-civilizations/

China-India Review. 2023. "Building China into a Great Modern Socialist Country." *Embassy of The People's Republic of China in the Republic of India*. January–March 2023. https://in.china-embassy.gov.cn/eng/xwfw/zgxw/202304/P020230411529354579726.pdf

Ding, Lu. 2016. "China's "Two Centenary" Goals: Progress and Challenge." *East Asian Policy*, 8 (2): 79–-93. DOI:10.1142/S1793930516000222

Ellis, R. Evan. 2023. "The Trouble with China's Global Civilization Initiative." *The Diplomat*, June 1, 2023. https://thediplomat.com/2023/06/the-trouble-with-chinas-global-civilization-initiative/

Fry, Tony. 2020. *Defuturing: A New Design Philosophy*. New York: Bloomsbury Publishing.

Huntington, Samuel P. 1993. "The Clash of Civilizations?" *Foreign Affairs* 72 (3): 22–49.

Jiang, Zemin. 2012. *Selected works of Jiang Zemin, Volume II*. Beijing: Foreign Languages Press.

Lađevac, Ivona. 2020. „Odgovor Kine na novi bezbednosni izazov." U *Čovek, prostor, tehnologija, ideje: međunarodna bezbednost u trećoj dekadi 21. veka*, ur. Vladimir Ajzenhamer i Nebojša Vuković, 126–151. Beograd: Institut za međunarodnu politiku i privredu, Univerzitet u Beogradu – Fakultet bezbednosti.

Liu, Jianchao. 2023. "Work Actively to Implement the Global Civilizational Initiative and Jointly Advance Human Civilizations". *China Daily, Hong Kong Edition*. April 13, 2023. https://campus-cdhk.oss-cn-hongkong.aliyuncs.com/store_manager/epaper/0/106/91/616601/page.pdf

Ministry of Foreign Affairs of the People's Republic of China [MFAPRC]. 2014. "Wang Yi: China, a Staunch Defender and Builder of International Rule of Law." October 24, 2014. https://www.fmprc.gov.cn/eng/wjb/wjbz/jh/202405/t20240527_11312100.html

Ministry of Foreign Affairs of the People's Republic of China [MFAPRC]. 2018. "Wang Yi: China is a Defender of the International Order and a Practitioner of Multilateralism." September 29, 2018. https://www.mfa.gov.cn/eng/wjb/zzjg_663340/gjs_665170/gjsxw_665172/202406/t20240606_11402008.html

Ministry of Foreign Affairs of the People's Republic of China [MFAPRC]. 2022. "The Chinese Dream Is a Dream of the People." August 26, 2022. https://www.mfa.gov.cn/eng/zy/jj/zggcddwjw100ggs/xsd/202406/t20240606_11377988.html

Mitrović, Dragana. 1995. "The People's Republic of China at the Threshold of the New Century." *The Review of International Affairs* XLVI (1033): 28–32.

North Atlantic Treaty Organisation [NATO]. 2022. "NATO 2022 Strategic Concept." June 29, 2022. https://www.act.nato.int/wp-content/uploads/2023/05/290622-strategic-concept.pdf

Popović, Slobodan. 2024. *Zagonetka u mozaiku poretka „kineskog sna": Šangajska organizacija za saradnju*. Beograd: Institut za međunarodnu politiku i privredu.

Qin, Yaqing. 2012. "Culture and Global Thought: Chinese International Theory in the Making." *Revista CIDOB d'Afers Internacionals* 100: 67–90.

Qin, Yaqing. 2018. *A Relational Theory of World Politics*. Cambridge: Cambridge University Press.

Qin, Yaqing. 2020a. "Diplomacy as Relational Practice." *The Hague Journal of Diplomacy* 15 (1-2): 165–173. DOI:10.1163/1871191X-15101092

Qin, Yaqing. 2020b. "Cooperation: An Iron Law for a Community with a Shared Future." *China International Studies* 82: 34–42.

Shambaugh, David. 2004. "China Engages Asia: Reshaping the Regional Order." *International security* 29 (3): 64–99.

Shanghai Cooperation Organization [SCO]. 2001. "Declaration on the Establishment of the Shanghai Cooperation Organization." June 15, 2001. Organization. https://eng.sectsco.org/images/07e8/0c/17/1626669.pdf

Shanghai Cooperation Organization [SCO]. 2002. "Charter on Shanghai Cooperation Organization." June 7, 2002. https://eng.sectsco.org/images/07e8/0c/14/1625960.pdf

Shanghai Cooperation Organization [SCO]. 2006. "Declaration on the Fifth Anniversary of the Shanghai Cooperation Organization." June 15, 2006. https://eng.sectsco.org/images/07e8/0c/17/1627730.pdf

Shanghai Cooperation Organization [SCO]. 2007. "Agreement on long-term neighborliness, friendship and cooperation of state members of the Shanghai Cooperation Organization." *CIS Legislation*. August 16, 2007. https://cis-legislation.com/document.fwx?rgn=20606

Shanghai Cooperation Organization [SCO]. 2015. "Ufa Declaration by the Heads of Member States of the Shanghai Cooperation Organization." July 10, 2015. https://eng.sectsco.org/images/07e8/0c/17/1628168.pdf

Shanghai Cooperation Organization [SCO]. 2021. "Dushanbe Declaration on the 20th Anniversary of the SCO." September 17, 2021. https://eng.sectsco.org/20210917/782639.html

Shanghai Cooperation Organization [SCO]. 2023. "SCO Year of Tourism 2023 forum." November 28, 2023. https://eng.sectsco.org/20231128/1192922.html

Shanghai Cooperation Organization [SCO]. 2024a. "Astana Declaration of the Council of Heads of State of the Shanghai Cooperation Organisation." July 9, 2024. https://eng.sectsco.org/20240709/1438929.html

Shanghai Cooperation Organization [SCO]. 2024b. "Joint Communique of the twenty-third Meeting of the Council of Heads of Government of Member States of the Shanghai Cooperation Organization." October 16, 2024. https://eng.sectsco.org/20241016/1574921.html

State Council Information Office of China [SCIO]. 2023. "Full Text: A Global Community of Shared Future: China's Proposals and Actions." September 26, 2023. http://english.scio.gov.cn/whitepapers/2023-09/26/content_116710660_3.htm

State Council of the People's Republic of China [SCIO]. 2024. "Xi's 'two sessions' messages highlight China's high-quality development in crucial year." March 11, 2024. https://english.www.gov.cn/news/202403/11/content_WS65efa5fbc6d0868f4e8e4ffc.html

Stefanović-Štambuk, Jelica. 2008. *Diplomatija u međunarodnim odnosima.* Beograd: Čigoja Štampa.

Stefanović-Štambuk, Jelica. 2010a. "„Da li je diplomatija danas održiva?" *Međunarodni problemi* 62 (4): 641–677. DOI:10.2298/MEDJP1004641S

Stefanović-Štambuk, Jelica. 2010b. „Predviđanja težišnih pravaca međunarodnih odnosa posle globalne krize." *Srpska politička misao* 2 (28): 297–333. DOI: 10.22182/spm.2822010.15

Stefanović-Štambuk, Jelica, and Slobodan Popović. 2022. ""A Thousand Miles" and "a Thousand Tasks": China's Diplomatization of Multifaceted Russia-Ukraine Conflict and Global Security." *The Policy of National Security* 23 (2): 7–31. DOI: 10.22182/pnb.2322022.1

Stefanović-Štambuk, Jelica, and Slobodan Popović. 2024. "China's Emancipatory Diplomacy for a Peaceful, Democratic, and Sustainable Global Community." *The Review of International Affairs* LXXV (1192): 405–436. DOI:10.18485/iipe_ria.2024.75.1192.5

Stekić, Nenad. 2023. *A Hesitant Hegemon: Layers of China's Contemporary Security Policy.* Belgrade: Institute of International Politics and Economics.

Strategic Studies Institute, US Army War College. 2024. *2024 Annual Estimate of the Strategic Environment.* Carlisle Barracks, PA: USAWC Press. http://www.jstor.org/stable/resrep62232

United Nations General Assembly [UNGA]. 2024. "Resolution A/78/L.75 International Day for Dialogue among Civilizations." May 29, 2024. https://documents.un.org/doc/undoc/ltd/n24/150/94/pdf/n2415094.pdf

Xi, Jinping. 2017. "Secure a Decisive Victory in Building a Moderately Prosperous Society in All Respects and Strive for the Great Success of Socialism with Chinese Characteristics for a New Era." *Xinhua*. October 18, 2017. http://www.xinhuanet.com/english/download/Xi_Jinping's_report_at_19th_CPC_National_Congress.pdf

Xi, Jinping. 2023a. "Full text of Xi Jinping's keynote address at the CPC in Dialogue with World Political Parties High-level Meeting". *State*

Council of the People's Republic of China [SCIO]. March 16, 2023. http://english.scio.gov.cn/topnews/2023-03/16/content_85171478.htm

Xi, Jinping. 2023b. *Chinese Modernization.* Beijing: Central Compilation and Translation Press.

Xi, Jinping. 2023c. "Full text of Xi's address at SCO summit." *State Council of the People's Republic of China* [SCIO]. July 4, 2023. https://english.www.gov.cn/news/202307/04/content_WS64a42528c6d0868f4e8dd7d3.html

Xi, Jinping. 2024a. "Full text of Xi Jinping's speech at 'Shanghai Cooperation Organization Plus' Meeting in Astana." *State Council of the People's Republic of China* [SCIO]. July 4, 2024. https://english.www.gov.cn/news/202407/04/content_WS6686c48ac6d0868f4e8e8e25.html

Xi, Jinping. 2024b. "Full Text: Address by Chinese President Xi Jinping at conference marking 70th anniversary of Five Principles of Peaceful Coexistence." *Xinhua*. June 28, 2024. https://english.news.cn/20240628/71733dd6f26441d4965dbb1a937e21ef/c.html

Xu Liuliu, Li Qian, and Hu Yuwei. 2023."The Global Civilization Initiative full of Chinese wisdom, injects fresh momentum into bright shared future - The way of civilization." *Global Times*. April 17, 2023. https://www.globaltimes.cn/page/202304/1289306.shtml

Zongze, Ruan. 2018. "Building Community with Shared Future: Meliorating the Era of Strategic Opportunity in China." *China International Studies* 69 (2): 5–27.

Университет Шанхайской организации сотрудничества. n.d. „Цели и задачи". Дата обращения 15 июня 2025. https://www.uni-sco.ru/stat/2/stat_2.html

Слободан Поповић[*]

Институт за међународну политику и привреду, Београд

Драган Траиловић[**]

Институт за међународну политику и привреду, Београд

КАДА ГЛОБАЛНА ЦИВИЛИЗАЦИЈСКА ИНИЦИЈАТИВА СРЕТНЕ ПРАКСУ – ШАНГАЈСКА ОРГАНИЗАЦИЈА ЗА САРАДЊУ[***]

Резиме

Рад се бави Шангајском организацијом за сарадњу (ШОС) као једним од најрепрезентативнијих примера практичне манифестације кинеске Глобалне цивилизацијске иницијативе (ГЦИ), покренуте 2023. године. Предмет рада је испитивање начина на који принципи и вредности ГЦИ, као што су прави мултилатерализам, равноправност, поштовање разноликости и одбацивање конфронтације, налазе своје институционално и практично утемељење у раду ШОС-а, као и какви су њени одговори на регионалне и глобалне изазове. Циљ рада је да покаже да ШОС својим институционалним развојем, механизмима сарадње и дипломатском праксом у великој мери остварује замишљене оквире Глобалне цивилизацијске иницијативе (ГЦИ), чиме се гради „нови тип међународних односа" заснован на дијалогу, међусобном поверењу и солидарности. Рад се заснива на квалитативном и интерпретативном приступу, уз ослањање на анализу докумената ШОС-а, званичних говора кинеских лидера и релевантне научне литературе. Резултати анализе показују да је ШОС кроз Шангајски дух институционализовао међуцивилизацијски дијалог, да

[*] Имејл: slobodan.popovic@diplomacy.bg.ac.rs; ORCID: 0000-0002-8094-8445

[**] Имејл: dragan.trailovic@diplomacy.bg.ac.rs; ORCID: 0000-0001-9707-9867

[***] Овај рад представља резултате студије развијене у оквиру истраживачког пројекта под називом „Србија и изазови у међународним односима 2025. године", који је финансирало Министарство науке, технолошког развоја и иновација Републике Србије, а спровео Институт за међународну политику и привреду у Београду током 2025. године.

промовише разноврсност култура и система вредности, и настоји да успостави регион поуздане предвидивости у понашању њених чланица и заједничке безбедности. Истраживање истиче и развој нових форми сарадње и то у области образовања, туризма, спорта и културне размене, које додатно ојачавају људске и институционалне везе међу државама чланицама. Иако постоје изазови у пракси, посебно у односима појединих чланица, досадашњи рад ШОС-а сведочи о снажној способности да цивилизацијске разлике претвара у фактор стабилности, а не у линије подела. На тај начин, ШОС се позиционирала као важан оквир у којем различите државе и културе, од великих и регионалних сила попут Кине, Русије и Индије, до мањих држава Централне Азије, могу усаглашавати интересе и креирати заједничке политике. Закључак рада је да Шангајска организација за сарадњу представља један од најуспешнијих облика примене Глобалне цивилизацијске иницијативе у пракси, потврђујући значај кинеске дипломатске праксе и мисли у обликовању алтернативног, инклузивног и хармоничног модела глобалног поретка.

Кључне речи: Шангајска организација за сарадњу, Глобална цивилизацијска иницијатива, дипломатија, Кина, заједница, заједничка будућност, истински мултилатерализам

* This manuscript was submitted on July 23, 2025, and accepted by the Editorial Board for publishing on November 10, 2025.

REVIEWS

UDC 327.56::351.88(497.12)
DOI: 10.5937/spm94-60393
Review

Српска политичка мисао
(Serbian Political Thought)
No 6/2025.
Vol. 94
pp. 271-286

Željko Oset*

American-Slovenian Education Foundation, Ljubljana

PILOT RETURN MIGRATION POLICY INDEX: CASE STUDY OF SLOVENIA**

INTRODUCTION

In response to the growing global importance of return migration as a key element of brain circulation, this paper presents the development of a pilot Return Migration Policy Index tailored to the Slovenian context. Designed to evaluate the feasibility and transferability of international return migration measures, the index is grounded in a multidimensional framework encompassing five core dimensions: Economy, Legal Order, Societal Openness to Immigrants, Migration History and Cultural Perception, and Official Language Accessibility. Drawing on interdisciplinary research, international best practices, and national policy analysis, each dimension was operationalized through specific indicators and weighted scoring criteria. The index was pilot-tested on 43 return migration measures adopted by various countries, allowing for both cross-national comparison and contextual relevance

* E-mail: zeljko.oset@gmail.com; ORCID: 0000-0001-6688-4126

** This paper is the result of the research projects Mechanisms for Attracting Foreign and Reintegrating Domestic Experts to Strengthen Slovenia's Position in Global Digitalization Trends and Between Brain Drain and Brain Circulation: Study and Postdoctoral Training in Finland, both financially supported by the Slovenian Research and Innovation Agency.

to Slovenia. The results demonstrate the index's utility in identifying promising practices and institutional barriers, as well as its potential to inform more transparent, coherent, and evidence-based policymaking. By institutionalizing the index, Slovenia can strengthen its capacity to attract and reintegrate global talent, contributing to its long-term development and competitiveness in the global knowledge economy.

BRAIN CIRCULATION IN THE CONTEXT OF GLOBAL MOBILITY

In the context of increasing global mobility and transnational flows of people, many countries have begun to formulate and implement targeted strategies aimed at encouraging the return of their own nationals, particularly those who are highly educated and professionally trained abroad. This phenomenon is part of a broader trend known as brain circulation, which is increasingly replacing the one-directional narrative of brain drain with a more dynamic understanding of knowledge exchange, professional mobility, and return migration (Saxenian 2005, 35; Wickramasekara 2011, 4–5).

Across Europe, Asia, and Latin America, national governments are developing returnee programs to attract individuals who left their home country for education, training, or employment opportunities abroad. These programs often include financial incentives, tax breaks, streamlined administrative procedures, support for housing and family reintegration, and career development schemes (Organisation for Economic Co-operation and Development [OECD] 2013, 118). In countries with significant diasporas, efforts are also directed toward the return of the descendants of emigrants, particularly those who maintain a cultural or emotional attachment to the country of their ancestors (Levitt and de la Dehesa 2003, 588; Ragazzi 2014, 73).

The motivation behind such policies is both economic and symbolic. On the one hand, states seek to harness the human capital and international experience of returnees to address domestic skills shortages, foster innovation, and enhance global competitiveness (Kapur and McHale 2005, 14). On the other hand, the return of emigrants or their descendants serves as a powerful affirmation of national identity, continuity, and global belonging (Délano and Gamlen 2014, 46).

Return migration policies are particularly relevant in the context of educational and professional mobility. Many individuals who study

abroad initially do not intend to settle permanently in the host country. However, attractive career opportunities, research conditions, or personal circumstances often lead to long-term or permanent stays (Teferra 2005, 54). Recognizing this, countries of origin have begun designing proactive measures to maintain connections with their nationals abroad and to facilitate their eventual return–whether temporarily or permanently (International Organization for Migration [IOM] 2020, 7–12).

Some governments, such as those of Portugal, Spain, Ireland, Latvia, and Israel, have developed comprehensive national strategies to support the reintegration of returnees. These may include programs tailored to entrepreneurs, researchers, or young professionals, as well as initiatives aimed at revitalizing local communities by attracting returnees to less developed regions (Ministry of Foreign Affairs of the Republic of Latvia 2021). Other countries, such as Germany and Sweden, focus on creating favorable conditions for return through strong research infrastructure, social support systems, and inclusive labor market policies (Cassarino 2004, 268).

Moreover, the return of second- or third-generation emigrants–those born and raised abroad–is gaining traction as a specific policy focus. For such individuals, return is not always driven by economic necessity but often by a desire to reconnect with cultural roots, seek new opportunities in a more meaningful or values-driven environment, or benefit from programs aimed at diaspora engagement (King and Christou 2011, 454; Brinkerhoff 2016, 40). Countries such as Italy, Poland, and Armenia have introduced symbolic and material measures to welcome such returnees and facilitate their integration into local society (Düvell and Garapich 2011, 32).

Ultimately, the global competition for talent has pushed national governments to view their emigrants and their descendants not as lost citizens, but as mobile assets–individuals who can contribute to the development of their home country if appropriate pathways and incentives are provided (Gamlen 2014, 25). The challenge remains to ensure that return is not only possible but desirable, requiring continuous efforts to improve institutional openness, economic opportunities, and social inclusion in the country of return.

TYPES OF POLICY MEASURES ADOPTED BY STATES TO ENCOURAGE THE RETURN OF SKILLED MIGRANTS AND THEIR DESCENDANTS

In recent decades, a growing number of countries have introduced targeted policies to attract the return of highly skilled emigrants and the descendants of their diasporas. These measures respond to structural labor shortages, demographic shifts, global talent competition, and the increasing strategic importance of diasporas as agents of development, innovation, and transnational cooperation (Kapur and McHale 2005, 21; Gamlen 2014, 25).

The policy instruments implemented by national governments to encourage the return of skilled emigrants and their descendants can be broadly categorized into several thematic areas, each addressing distinct facets of the return migration process. First, financial and fiscal incentives represent a common approach, encompassing measures such as temporary tax exemptions (e.g., Italy's *Decreto Rientro dei Cervelli* and Portugal's non-habitual resident scheme), relocation grants, and entrepreneurship support through startup funding or preferential credit access (Le Coz 2021, 7–19). Second, research and academic mobility programs target returning scholars and scientists by offering dedicated grants, enhanced access to research infrastructure, and streamlined recognition of foreign qualifications (Teferra 2005, 54). Third, states have introduced measures under the rubric of administrative and legal facilitation, including centralized "one-stop shops" for returnee services, simplified visa and residency procedures, and expanded digital public services (Cassarino 2004, 268; IOM 2020, 7–12). Fourth, integration and reintegration support aims to ease the transition of returnees into local societies through mentorship initiatives, recognition of prior professional experience, and targeted family support policies (Brinkerhoff 2016, 40; King and Christou 2011, 454). Fifth, diaspora engagement and cultural reconnection initiatives strengthen identity-based ties through language courses, youth exchanges, and legal provisions such as the "right of return" or fast-track citizenship (Levitt and de la Dehesa 2003, 588; Ragazzi 2014, 73). Finally, symbolic and recognition-based measures seek to affirm the societal value of returnees by publicly acknowledging their contributions, organizing diaspora-focused events, and integrating returnee perspectives into policymaking processes (Délano and Gamlen 2014, 46; Gamlen 2019, 14). Together, these categories illustrate the

multidimensional nature of return migration governance and the diverse policy tools deployed to attract and retain mobile talent.

WHY DEVELOP A PILOT INDEX FOR EVALUATING RETURN MIGRATION MEASURES?

In recent years, Slovenia has increasingly recognized the strategic importance of facilitating the return of educated and professionally trained individuals residing abroad. While several policy initiatives already exist–such as targeted programs for researchers and entrepreneurs–there is currently no comprehensive tool for systematically evaluating the effectiveness, coherence, and long-term impact of these measures. This institutional gap prompted the project team, under the framework of the research project "Mechanisms for Attracting Foreign and Returning Domestic Experts to Strengthen Slovenia's Position in Global Digitalization Trends", to propose the development of a pilot index to assess return migration policy.

The decision to create such an index was driven by several key factors identified in the project's baseline assessment: 1. Lack of a centralized and consistent database on returning experts, which would enable longitudinal analysis of their contributions; 2. Insufficient institutional coordination across relevant ministries and implementing agencies (e.g., education, labor, diaspora affairs); 3. Fragmentation of policy instruments, which are often under-publicized or difficult to access for potential returnees. Absence of performance indicators to evaluate which programs work, for whom, and under what conditions. These shortcomings mirror challenges observed in other small states with significant diasporas and limited policy integration (Gamlen 2014, 25; Cassarino 2004, 268). Additionally, the international literature and best practices emphasize that return migration should not be understood merely as a one-time event, but rather as a complex, multidimensional process embedded in broader patterns of brain circulation and transnational engagement (Saxenian 2005, 35; Wickramasekara 2011, 4).

Countries such as Ireland, Estonia, and Portugal have moved in recent years toward evidence-based models of return policy design, introducing monitoring frameworks and performance metrics to assess the cost-effectiveness and societal benefits of their programs (Le Coz 2021, 14–16). In this context, Slovenia currently lacks an equivalent evaluative mechanism, which makes it difficult to assess the return on

public investment in such schemes or to adapt them based on real-time feedback.

The development of the pilot index is driven by a set of interrelated objectives aimed at enhancing the strategic management of return migration policies in Slovenia. Foremost among these is the creation of a structured and comparative framework that allows for the systematic evaluation of return-related programs over time. In doing so, the index establishes a foundation for regular reporting, enabling policymakers to monitor progress, identify implementation gaps, and formulate data-driven reforms. Crucially, the index is designed to integrate both quantitative metrics–such as program participation and employment rates–and qualitative dimensions, including user experience and accessibility, into a cohesive evaluative tool. By incorporating the perspectives of returnees themselves, the index ensures that policy assessment reflects not only institutional intentions but also lived experiences (Brinkerhoff 2016, 40; King and Christou 2011, 454). Furthermore, the tool promotes greater transparency and accountability by making information on return migration measures more accessible to domestic and international stakeholders. Importantly, the index is not envisioned as a purely statistical mechanism; rather, it serves as a strategic instrument for institutional learning, cross-sectoral coordination, and adaptive policymaking. By enabling Slovenia to benchmark its return migration efforts against relevant international practices–while remaining attuned to its national context–the index aspires to contribute to a more inclusive, sustainable, and evidence-based migration governance system.

DESIGNING THE PILOT INDEX: METHODOLOGICAL APPROACH AND KEY DIMENSIONS

The development of the pilot index for evaluating return migration measures in Slovenia was grounded in a recognition of the need for a systematic, multidimensional, and evidence-based assessment of public policies aimed at facilitating the return and reintegration of skilled nationals and members of the diaspora. Drawing from the findings of the national CRP research project and informed by international frameworks (e.g., OECD, IOM, EU), the project team adopted a multi-step methodology focused on conceptual clarity, stakeholder relevance, and practical usability.

METHODOLOGICAL APPROACH

The process began with a comprehensive review of international good practices and analytical tools, including return migration dashboards used in countries such as Portugal, Spain, Austria, Germany, Italy, Sweden, and Israel. These were complemented by policy evaluation methodologies used by the European Commission and diaspora engagement indicators proposed by IOM and the Migration Governance Framework (IOM 2020, 7–12).

At the national level, we conducted a mapping of Slovenian return-related programs, including the Dr. Aleš Debeljak Program (for returning researchers), incentives for diaspora entrepreneurship, and diaspora cultural engagement activities. Semi-structured interviews were organized with returnees, policymakers, and civil society actors to identify gaps, barriers, and opportunities within the existing system. We also conducted interviews with experts from Portugal, Finland, and Israel.

The guiding methodological principles were: 1. Relevance (alignment with Slovenia's strategic migration and development goals); 2. Feasibility (availability of data and responsible institutions); 3. Comparability (potential to benchmark across time or countries), 4. Inclusiveness (incorporation of returnee experience and expert feedback). Based on this process, five main dimensions (Economy, Legal Order, Societal Openness to Immigrants, Migration History and Cultural Perception, and Official Language Accessibility) were selected, each representing a critical area of state intervention.

DEFINING DIMENSIONS AND INDICATORS FOR THE PILOT INDEX

To enable a structured and nuanced evaluation of the feasibility and applicability of various return migration measures in Slovenia, we developed a pilot index composed of five key dimensions. Each dimension includes specific indicators, and where appropriate, sub-criteria. These dimensions reflect a combination of economic, legal, sociocultural, and historical factors relevant to the return and integration of skilled migrants and diaspora members.

The selection of the five dimensions–Economy, Legal Order, Societal Openness to Immigrants, Migration History and Cultural Perception, and Official Language Accessibility–was based on a thorough

interdisciplinary analysis of return migration literature, national policy gaps, and contextual priorities specific to Slovenia. Each dimension addresses a distinct yet interrelated domain that can significantly influence the success of return policies.

Economy was selected to capture the structural readiness of the national economic system to absorb returnees and offer meaningful, sustainable employment opportunities. High-skilled individuals are more likely to return if they can find or create jobs that match their expertise and aspirations. Economic incentives and the overall productivity of the labor market are key motivating factors.

Legal Order was included due to its central role in regulating the conditions of return, residence, work authorization, and entrepreneurship. A transparent, efficient, and inclusive legal framework lowers the administrative burden for returnees and increases policy credibility. The rule of law is also critical for protecting rights and building institutional trust.

Societal Openness to Immigrants was chosen to assess the sociocultural climate in which returnees and newcomers must navigate. Public attitudes toward immigrants, the accessibility of services, and support networks are essential for successful integration and long-term retention. This dimension helps capture intangible barriers that may not be reflected in legal provisions alone.

Migration History and Cultural Perception was added to address the legacy of past emigration, particularly forced or politically motivated displacement. For many second- or third-generation diaspora members, returning to Slovenia involves negotiating complex emotional and identity-based dynamics shaped by family histories and social narratives. This dimension recognizes the long-term influence of historical memory on policy outcomes.

Official Language Accessibility was incorporated to evaluate how language functions as both an enabler and barrier to reintegration. While English is widely used in global academia and business, the acquisition of Slovenian is often essential for social inclusion, civic participation, and deeper cultural integration. Balancing these linguistic needs is crucial in shaping inclusive policy. These dimensions were deliberately selected to ensure the index captures both structural (legal, economic) and human-centered (cultural, linguistic, historical) elements of return migration, making it a holistic tool for evaluating the feasibility of diverse policy measures.

ECONOMY

This dimension assesses the structural conditions of the national economy, focusing on labor market flexibility and the overall capacity of the economic environment to support high-impact return migration policies. The analysis was guided by several key questions: How does the socio-economic system in Slovenia compare with those in selected reference countries? What financial and structural advantages can Slovenia offer to potential returnees, particularly in light of its GDP trends and fiscal outlook? And crucially, is the Slovenian economy sufficiently robust to absorb skilled returnees and sustain high-value-added employment opportunities?

To operationalize this dimension, we identified three equally weighted indicators. The first is the growth of value added per employee, which captures productivity, innovation potential, and the economy's readiness to integrate high-skilled labor. The second is labor market flexibility, measured through job vacancy rates and the ability of employers to recruit qualified personnel. The third indicator is the standard of living, calculated based on a financial threshold equivalent to 1.5 times the national average salary. Collectively, these indicators not only reflect the structural attractiveness of Slovenia as a return destination but also highlight the potential economic gains that return migration can generate.

LEGAL ORDER

This dimension examines the extent to which the legal and institutional framework in Slovenia facilitates return migration and supports entrepreneurial activity among returnees. The analysis was informed by several core questions: How accessible and transparent is Slovenia's legal environment? To what degree can returnees efficiently regularize their residence and employment status? And what forms of institutional support are available through universities and public agencies for individuals seeking to establish businesses upon return?

To evaluate these aspects, five equally weighted indicators were defined. The first, clarity of the legal framework, assesses the availability and comprehensibility of relevant legal documents. The second, status regularization mechanisms, captures the speed and ease with which residency and work permits can be obtained. The third

indicator, academic and entrepreneurial conditions, evaluates the quality of the higher education sector and the availability of incentives for startups. The fourth, social benefits and entitlements, considers the scope and accessibility of targeted programs for returnees. Finally, debureaucratization efforts measure the degree to which policy reforms and digital public services reduce administrative burdens.

The legal and institutional environment is especially significant for returnees who may be unfamiliar with domestic regulations or administrative procedures. A clear, supportive, and efficient legal framework not only enhances policy credibility but also plays a decisive role in returnees' decisions to reintegrate and invest their skills in the home country.

SOCIETAL OPENNESS TO IMMIGRANTS

This dimension explores the overall societal openness in Slovenia toward foreigners and returnees, with particular attention to their ability to access essential information, build support networks, and integrate into local communities. The evaluation was guided by key questions such as: How accessible is Slovenia to international students and researchers? And what kinds of institutional or informal support structures are available to facilitate their integration?

To capture these dynamics, four indicators were established. The first, information accessibility, measures the availability of relevant resources in both Slovenian and English, which is critical for effective navigation of public services and institutions. The second, social support networks, assesses the presence of community organizations, social events, and mentorship programs that assist newcomers in building connections. The third indicator, general sense of safety, draws on public safety rankings and recorded instances of discrimination to evaluate the societal climate for returnees. The fourth, integration metrics, considers the migration balance of highly educated individuals as a proxy for the inclusiveness and attractiveness of the host environment. Societal receptiveness and clear integration pathways are vital for creating a welcoming atmosphere that not only encourages return but also supports the long-term reintegration of returnees into the social and professional fabric of the country.

MIGRATION HISTORY AND CULTURAL PERCEPTION

This qualitative dimension addresses the historical and cultural context shaping emigration and return in Slovenia. It reflects the enduring influence of past experiences–particularly those related to post-World War II political exile–on contemporary perceptions of returnees. Discussions within the project highlighted how public narratives, societal attitudes, and lingering stereotypes can affect both the willingness of emigrants to return and the ease of their reintegration. In this regard, return migration is not merely a logistical or economic process, but also a deeply symbolic and identity-driven experience.

To assess this dimension, two equally weighted indicators were defined. The first, cultural-historical background, evaluates how historical narratives of emigration influence present-day readiness for integration, including the presence of stigma or contested belonging. The second, support mechanisms, considers the existence and effectiveness of organizations that facilitate cultural reintegration and actively work to counter bias or exclusion.

This dimension is particularly significant for second- and third-generation descendants of emigrants, for whom the decision to return is often shaped not only by opportunity structures, but also by emotional ties, inherited memories, and a desire to reconnect with ancestral heritage. Understanding these cultural and historical dynamics is therefore essential to designing return policies that are both inclusive and responsive to the complex realities of transnational identity.

OFFICIAL LANGUAGE ACCESSIBILITY

The final dimension assesses the role of language in the return and reintegration process, recognizing it as both a potential barrier and a critical enabler of successful integration. The analysis was guided by two central questions: To what extent are public services and educational resources available in a widely spoken lingua franca, particularly English? And how accessible are Slovenian language learning opportunities for newcomers and returnees alike?

To operationalize this dimension, two indicators were established. The first, the availability of information in English, evaluates whether essential information and services are provided in a language that facilitates immediate functionality within the host society. The second,

promotion of Slovenian language learning, examines the extent to which affordable and publicly supported language courses are available, helping returnees and immigrants acquire the linguistic skills necessary for long-term social inclusion and civic participation.

Linguistic accessibility is especially important for returnees who may have grown up abroad without exposure to the Slovenian language or cultural context. Ensuring both the availability of English-language resources and robust support for Slovenian language acquisition is therefore vital for fostering an inclusive environment that accommodates diverse return pathways.

WEIGHTED EVALUATION APPROACH

The index was designed as a weighted scoring system, rather than an unweighted composite. Each dimension is assigned a weight based on its relative importance to the specific policy being assessed. The overall index score (ranging from 0 to 5) represents the sum of all weighted dimension scores: Final Index Score = Σ (Weight × Dimension Score).

For each return policy or measure evaluated, detailed justifications for the assigned weights and scores were provided. This approach allows for flexible adaptation and ensures that the index reflects the specific requirements and feasibility conditions of diverse return migration initiatives.

PILOT TESTING AND EVALUATION OF POLICY MEASURES

The creation of a multidimensional index for evaluating return and attraction policies is inherently exploratory and experimental. As such, our process began with a pilot assessment of selected policy measures using the weighted scoring method outlined above. Both quantitative and qualitative data were applied to generate meaningful and balanced evaluations.

The first phase involved a test scoring of return migration policies from selected countries. Indicators were scored on a five-point scale (1–5), and scores were multiplied by dimension-specific weights. The selection of weights was internally validated through structured discussions within the research team. Particular attention was given to dimensions 2 (Legal Order), 4 (Migration History and Cultural Perception), and 5 (Language

Accessibility), where it was challenging to balance the targeted nature of the measure in the original country against its transferability and applicability to the Slovenian context.

To address this, a consensus was reached: 50% of the final policy score would derive from the internal quality of the measure, while the remaining 50% would assess its relevance and feasibility for implementation in Slovenia. This adjustment proved crucial, especially for culturally specific or niche policies–such as reconciliation measures or initiatives attracting elite talent–where contextual factors dominate.

Subsequently, a total of 43 policy measures were distributed among members of the research team for individual evaluation. Each member applied the weighting scheme to score the measures independently. This was followed by a cross-validation phase: each member was assigned a peer's evaluated set of measures to review and critique. Comments, reflections, and discrepancies were compiled in a shared document and discussed in a team meeting to ensure consistency, reliability, and transparency across evaluations. The resulting dataset not only enabled comparative analysis of return migration policies but also served as a proof of concept for refining the pilot index in future iterations.

PROPOSED POLICY MEASURES AND IMPLEMENTATION CONSIDERATIONS FOR SLOVENIA

Based on the comparative analysis of international return migration policies and the application of our pilot evaluation index, we propose a series of policy measures that could significantly enhance Slovenia's ability to attract and reintegrate both foreign experts and returnees from its diaspora. These recommendations are grounded in best practices identified in countries such as Ireland, Portugal, Italy, Germany, and Israel, and are adapted to Slovenia's demographic, economic, and administrative context.

Based on the comparative evaluation of international return migration policies and the application of the pilot index, several concrete policy measures are proposed to enhance Slovenia's capacity to attract and reintegrate returnees and diaspora professionals. These recommendations draw on successful practices from countries such as Portugal, Ireland, Germany, Israel, and Italy, while being tailored to Slovenia's specific administrative, economic, and demographic context.

First, the creation of a Central Coordinating Office for Return Migration would provide a unified institutional platform for return-related services. Inspired by Portugal's "Welcome Office" and Israel's returnee support centers, such an office in Slovenia could offer centralized legal advice, housing and employment assistance, and integration support. Effective implementation would require inter-ministerial coordination–particularly among the Ministries of Foreign Affairs, Education, Labor, and Internal Affairs–as well as targeted investments in digital infrastructure and the inclusion of diaspora representatives in the design and governance of services.

Second, the introduction of return incentives and financial support schemes could significantly improve the attractiveness of return. These might include temporary tax relief, subsidies for returnee entrepreneurs, and housing grants, modeled on programs such as Italy's *Rientro dei Cervelli* and Ireland's *Safe Home* initiative. Key implementation steps would involve revising national tax legislation to accommodate return-specific provisions, coordinating with the national development fund and employment agency, and defining transparent eligibility criteria based on duration abroad, qualifications, or sectoral relevance.

Third, measures aimed at recognizing transnational experience and supporting dual careers are essential. Returnees frequently bring valuable international knowledge and skills, particularly in fields such as academia and healthcare. Policymakers should facilitate the automatic recognition of foreign qualifications, provide dual-career support for returning couples, and expand access to national and European funding programs targeting returnee researchers and entrepreneurs.

Fourth, Slovenia should formalize diaspora engagement through institutional representation. Following the examples of Ireland and Germany, this could involve establishing a Council for Slovenians Abroad with consultative status, integrating diaspora affairs into national strategic frameworks, and offering dual citizenship or expedited reintegration rights for the descendants of exiles and emigrants.

Fifth, it is necessary to develop tailored integration measures for second- and third-generation returnees, who often face distinct cultural and linguistic challenges. This could include culturally sensitive reintegration programs, Slovenian language courses both abroad and domestically, and partnerships with diaspora schools and cultural organizations to foster smoother transitions.

Sixth, Slovenia should invest in strategic communication and outreach campaigns to reshape public perceptions of return migration and highlight its societal benefits. This includes launching media campaigns that profile successful returnees, building a digital platform for skills matching between returnees and domestic employers, and mobilizing influencers and diaspora networks to strengthen trust and engagement.

Finally, the institutionalization of robust monitoring, evaluation, and feedback mechanisms is critical. The pilot index developed in this study could serve as a foundation for ongoing monitoring of return migration policies. To support this, the government should designate a statistical agency or policy observatory to maintain and refine the index, conduct regular data collection on return flows, and fund qualitative research that captures the lived experiences and evolving needs of returnees. Together, these measures offer a roadmap for developing a more coordinated, inclusive, and evidence-based approach to return migration in Slovenia.

REFERENCES

Brinkerhoff, Jennifer M. 2016. *Institutional Reform and Diaspora Entrepreneurs: The In-Between Advantage.* New York: Oxford University Press.

Cassarino, Jean-Pierre. 2004. "Theorising Return Migration: The Conceptual Approach to Return Migrants Revisited." *International Journal on Multicultural Societies* 6 (2): 253–279. DOI: 10.1590/S1980-85852013000200003

Délano Alonso, Alexandra, and Alan Gamlen. 2014. "Comparing and Theorizing State-Diaspora Relation." *Political Geography* 41: 43–53. DOI: 10.1016/j.polgeo.2014.05.005

Düvell, Franck, and Michael Garapich. 2011. *Polish Migration to the UK: Continuities and Discontinuities* (Working Paper No. 84). Oxford: Centre on Migration, Policy and Society; University of Oxford.

Gamlen, Alan. 2014. "Diaspora Institutions and Diaspora Governance." *International Migration Review* 48 (1): 180–217. DOI: 10.1111/imre.12136

Gamlen, Alan. 2019. *Human Geopolitics: States, Emigrants, and the Rise of Diaspora Institutions*. Oxford: Oxford University Press.

International Organization for Migration [IOM]. 2020. *Mobilizing diaspora for development: Lessons for Armenia*. Geneva: IOM.

Kapur, Devesh, and John McHale. 2005. *Give Us Your Best and Brightest: The Global Hunt for Talent and Its Impact on the Developing World*. Washington, D.C.: Center for Global Development.

King, Russell, and Anastasia Christou. 2011. "Of Counter-Diaspora and Reverse Transnationalism: Return Mobilities to and from the Ancestral Homeland." *Mobilities* 6 (4): 451–466. DOI: 10.1080/17450101.2011.603941

Le Coz, Camille. 2021. "EU Strategy on Voluntary Return and Reintegration: Crafting a Road Map to Better Cooperation with Migrants' Countries of Origin." *Migration Policy Institute*. May 2021. https://www.migrationpolicy.org/sites/default/files/publications/mpi_eu-strategy-voluntary-return-reintegration_final.pdf?utm_source=chatgpt.com

Levitt, Peggy, and Rafael de la Dehesa. 2003. "Transnational Migration and the Redefinition of the State: Variations and Explanations." *Ethnic and Racial Studies* 26 (4): 587–611. DOI: 10.1080/0141987032000087325

Ministry of Foreign Affairs of the Republic of Latvia. 2021. "Diaspora Policy." July 1, 2021. https://www.mfa.gov.lv/en/diaspora-policy?utm_source=chatgpt.com&utm_source=https%3A%2F%2Fchatgpt.com%2F

Organisation for Economic Co-operation and Development [OECD]. 2013. *Recruiting Immigrant Workers: Germany*. OECD Publishing.

Ragazzi, Francesco. 2014. "A Comparative Analysis of Diaspora Policies." *Political Geography* 41: 74–89. DOI: 10.1016/j.polgeo.2013.12.004

Saxenian, AnnaLee. 2005. "From Brain Drain to Brain Circulation: Transnational Communities and Regional Upgrading in India and China." *Studies in Comparative International Development* 40 (2): 35–61. DOI: 10.1007/BF02686293

Teferra, Damtew. 2005. "Brain Circulation: Unparalleled Opportunities, Underlying Challenges, and Outmoded Presumptions." *Journal of Studies in International Education* 9 (3): 229–250. DOI: 10.1177/1028315305277619

Wickramasekara, Piyasiri. 2011. *Circular Migration: A Triple Win or a Dead End*. Global Union Research Network Discussion Paper No. 15. Geneva: International Labour Office; Global Union Research Network.

* This review was submitted on July 24, 2025, and accepted by the Editorial Board for publishing on November 10, 2025.

AUTHOR GUIDELINES

The academic journal *Serbian Political Thought* publishes articles that result from the latest theoretical and empirical research in the field of political science. Authors should refer mainly to the results of scientific research published in academic journals, primarily in political science journals.

Manuscripts should be submitted in Serbian (Cyrillic script) with a mandatory English translation, or in English.

The journal is published six times a year. The deadlines for submitting the manuscripts are February 1st, April 1st, June 1st, August 1st, October 1st, and December 1st.

Two consecutive issues cannot contain articles written by the same author, whether single-authored or co-authored.

Papers are submitted to the Editorial Board by uploading them to the CEON platform using the following link: https://aseestant.ceon.rs/index.php/spm/login.

Authors are obliged to submit a signed and scanned declaration of authorship when submitting their works. The declaration form can be downloaded from the journal's website: https://www.ips.ac.rs/en/magazines/srpska-politicka-misao/authors_directions/

All submitted manuscripts are checked for plagiarism or auto-plagiarism. Various forms of chat boxes and other artificial intelligence software cannot be (co)authors of the papers under consideration. These tools can only be used for stylistic language editing, not for writing sections of the paper, and authors who use them are obliged to specify the purpose of using such tools at the point where they are used.

Authors are required to provide their ORCID numbers along with their (preferably) institutional email addresses, which they include in the manuscript text in a footnote alongside their names and surnames.

Research articles can have up to 40,000 characters with spaces, including footnotes. When counting the characters leave out the reference list. Exceptionally, a monographic study can be larger in scope in accordance with the provisions of *the Rulebook on procedure, method of evaluation, and quantitative presentation of scientific research results.*

Reviews can have up to 15,000 characters with spaces.

Book reviews can have up to 10,000 characters with spaces.

CITING AND REFERENCING

The journal *Serbian Political Thought* uses a partially modified Chicago style of citation (17th edition of the *Chicago Manual of Style*), which implies specifying bibliographic parentheses (brackets) according to the author-date system in the text, as well as a list of references with full bibliographic data after the text of the paper.

Data in bibliographic parentheses and the list of references should be written in Latin script.

Below are the rules and examples for citing the bibliographic information in the reference list and in the text. For each type of source, a citation rule is given first, followed by an example of citation in the reference list and bibliographic parenthesis.

The bibliographic parenthesis is usually set off at the end of the sentence, before the punctuation mark. It contains the author's surname, the year of publication, and page numbers pointing to a specifically contextual page or range of pages, as in the following example: (Mearsheimer 2001, 15–17).

Books

Books with one author

Surname, Name. Year of publication. *Title*. Place of publication: Publisher.

Mearsheimer, John J. 2001. *The Tragedy of Great Power Politics.* New York: W. W. Norton & Company.
(Mearsheimer 2001)

Books with two or three authors

Surname, Name, and Name Surname. Year of publication. *Title*. Place of publication: Publisher.

Brady, Henry E., and David Collier. 2010. *Rethinking Social Inquiry: Diverse Tools, Shared Standards.* Lanham: Rowman & Littlefield Publishers.
(Brady and Collier 2010, 211)

Pollitt, Christopher, Johnston Birchall, and Keith Putman. 1998. *Decentralising Public Service Management*. London: Macmillan Press.
(Pollitt, Birchal and Putman 1998)

Books with four or more authors

Surname, Name, Name and Surname, Name and Surname, and Name and Surname. Year of publication. *Title*. Place of publication: Publisher.

> Pollitt, Christopher, Colin Talbot, Janice Caulfield, and Amanda Smullen [Pollitt *et al.*]. 2005. *Agencies: How Governments do Things Through Semi-Autonomous Organizations*. New York: Palgrave Macmillan.
> (Pollitt *et al.* 2005)

Editor(s) or translator(s) in place of the author(s)

Surname, Name, Name and Surname, ed. Year of publication. *Title*. Place of publication: Publisher.

> Kaltwasser, Cristobal Rovira, Paul Taggart, Paulina Ochoa Espejo, and Pierre Ostigoy [Kaltwasser *et al.*], eds. 2017. *The Oxford Handbook of Populism.* New York: Oxford University Press.
> (Kaltwasser *et al.* 2017)

Chapter in an edited book

Surname, Name. Year of publication. "Title of the chapter." In *Title*, ed. Name Surname, pages range. Place of publication: Publisher.

> Lošonc, Alpar. 2019. "Discursive dependence of politics with the confrontation between republicanism and neoliberalism." In *Discourse and Politics*, eds. Dejana M. Vukasović and Petar Matić, 23-46. Belgrade: Institute for Political Studies.
> (Lošonc 2019)

Journal Articles

Regular issue

Surname, Name. Year of publication. "Title of the article." *Journal* Volume, if available (issue): page range. DOI.

> Ellwood, David W. 2018. "Will Brexit Make or Break Great Britain?" *Serbian Political Thought* 18 (2): 5-14. DOI: 10.22182/spt.18212018.1.
> (Ellwood 2018)

Special issue

Surname, Name. Year of publication. "Title of the article." In "Title of the special issue", ed. Name Surname, Special issue, *Journal*: page range. DOI.

Chin, Warren. 2019. "Technology, war and the state: past, present and future." In "Re-visioning war and the state in the twenty-first century." Special issue, *International Affairs* 95 (4): 765–783. DOI: 10.1093/ia/iiz106.
(Chin 2019)

Encyclopedias and dictionaries

When the author/editor is known

Surname, Name, Name Surname, ed. Year of publication. *Title*. Vol. Place of publication: Publisher.

Badie, Bertrand, Dirk Berg-Schlosser, and Leonardo Morlino, eds. 2011. *International Encyclopedia of Political Science*. Vol. 1. Los Angeles: Sage Publications.
(Badie, Berg-Schlosser and Morlino 2011)

When the author/editor is unknown

Title. Year of publication. Place of publication: Publisher.

Webster's Dictionary of English Usage. 1989. Springfield, Massachusetts: Merriam-Webster Inc.
(*Webster's Dictionary of English Usage* 1989)

PhD dissertation

Surname, Name. Year of publication. "Title of the dissertation." PhD diss. University.

Munger, Frank J. 1955. "Two-Party Politics in the State of Indiana." PhD diss. Harvard University.
(Munger 1955, 17–19)

Newspapers and magazines

Signed articles

Surname, Name. Year of publication. "Title of the article." *Newspaper/Magazine* Date: page range.

Clark, Phil. 2018. "Rwanda's Recovery: When Remembrance is Official Policy." *Foreign Affairs*, January/February 2018: 35–41.
(Clark 2018)

Unsigned articles

Title of the newspaper/magazine. Year of publication. "Title of the article." Date: page range.

> *New York Times*. 2002. "In Texas, Ad Heats Up Race for Governor." July 30, 2002.
> (*New York Times* 2002)

Corporate Author

Name of the corporate author [acronym if needed]. Year of publication. *Title of the publication*. Place of publication: Publisher.

> International Organization for Standardization [ISO]. 2019. *Moving from ISO 9001:2008 to ISO 9001:2015.* Geneva: International Organization for Standardization.
> (International Organization for Standardization [ISO] 2019) – *The first in-text citation*
> (ISO 2019) – *Second and all subsequent citations*

Special cases of referencing

Citing editions other than the first

Surname, Name. Year of publication. *Title*, edition number. Place of publication: Publisher.

> Bull, Hedley. 2012. *The Anarchical Society: A Study of Order in World Politics*, 4th edition. New York: Columbia University Press.
> (Bull 2012)

Multiple sources of the same author

1) *Multiple sources by the same author* should be arranged chronologically by year of publication in ascending order.

> Mearsheimer, John J. 2001. *The Tragedy of Great Power Politics.* New York: W. W. Norton & Company.
>
> Mearsheimer, John J. 2010. "The Gathering Storm: China's Challenge to US Power in Asia." *The Chinese Journal of International Politics* 3 (4): 381–396. DOI: 10.1093/cjip/poq016.

2) *Multiple sources by the same author from the same year* should be alphabetized by title, with lowercase letters attached to the year. Those letters should be used in parenthetical citations as well.

Walt, Stephen M. 2018a. *The Hell of Good Intentions: America's Foreign Policy Elite and the Decline of U.S. Primacy*. New York: Farrar, Straus and Giroux.
(Walt 2018a)

Walt, Stephen M. 2018b. "Rising Powers and the Risk of War: A Realist View of Sino-American Relations." In *Will China's Rise be Peaceful: Security, Stability and Legitimacy*, ed. Asle Toje. 13–32. New York: Oxford University Press.
(Walt 2018b)

3) *Single-authored sources precede multiauthored sources beginning with the same surname* or written by the same person.

Pollitt, Christopher. 2001. "Clarifying convergence. Striking similarities and durable differences in public management reform." *Public Management Review* 3 (4): 471–492. DOI: 10.1080/14616670110071847.

Pollit, Christopher, Johnston Birchall, and Keith Putman. 1998. *Decentralising Public Service Management*. London: Macmillan Press.

4) *Multiauthored sources with the same name and surname* as the first author should continue to be alphabetized by the second author's surname.

Pollitt Christopher, Johnston Birchall, and Keith Putman. 1998. *Decentralising Public Service Management*. London: Macmillan Press.

Pollitt Christopher, Colin Talbot, Janice Caulfield, and Amanda Smullen. 2005. *Agencies: How Governments do Things Through Semi-Autonomous Organizations*. New York: Palgrave Macmillan.

Special cases of parenthetical citation

Exceptions to the rule of placing the parenthetical citation at the end of a sentence

1) If the *author is mentioned in the text*, even if used in a possessive form, the year must follow in parenthesis, and page numbers should be put in the brackets at the end of the sentence.

For the assessment, see Kaltwasser *et al*. (2017) ... (112).
According to Ellwood (2018) … (7).

2) When *quoting directly*, if the name of the author precedes the quotation, the year and page numbers must follow in parenthesis.

Mearsheimer (2001, 28) claims that: "...".

3) When *using the same source multiple times in one paragraph*, the parenthetical citation should be placed either after the last reference (or at the end of the paragraph, preceding the final period) if the same page (or page range) is cited more than once, or at the first reference, whereas the subsequent citations should only include page numbers.
Do not use *ibid* or *op. cit.* with repeated citations.

Using brief phrases such as "see", "compare" etc.

Those phrases should be enclosed within the parenthesis.

(see: Ellwood 2018)

Using secondary source

When using a secondary source, the original source should be cited in parenthesis, followed by "quoted/cited in" and the secondary source. The reference list should only include the secondary source.

"Its authority was greatly expanded by the constitutional revision of 1988, and the Court of Arbitration can now be regarded as a 'genuine constitutional court'" (De Winter and Dumont 2009, 109 cited in: Lijphart 2012, 39–40).

Lijphart, Arend. 2012. *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*, 2nd edition. New Haven & London: Yale University Press.

Multiple sources within the same parentheses

1) When *multiple sources* are cited, they should be separated by semicolons.

(Mearsheimer 2001, 34; Ellwood 2018, 7)

2) When *multiple sources by the same author*, but published in different years are cited, the name of the author is cited only the first time. The different years are separated by commas or by semicolons where page numbers are cited.

(Mearsheimer 2001, 2010) or (Mearsheimer 2001, 15–17; 2010, 390)

3) When *different authors share the same surname*, include the first initial in the parenthesis.

(M. Chiti 2004, 40), (E. Chiti 2004, 223)
Chiti, Edoardo. 2004. "Administrative Proceedings Involving European Agencies." *Law and Contemporary Problems* 68 (1): 219–236.

Chiti, Mario. 2004. "Forms of European Administrative Action." *Law and Contemporary Problems* 68 (1): 37–57.

Legal and Public Documents

Sections, articles, or paragraphs can be cited in the parentheses. They should be appropriately abbreviated.

Constitutions and laws

The title of the legislative act [acronym if needed], "Official Gazette of the state" and the number of the official gazette, or the webpage and the date of last access.

> The Constitution of the Republic of Serbia, "Official Gazette of the Republic of Serbia", No. 98/06.
> (The Constitution of the Republic of Serbia, Art. 33)

> The Law on Foreign Affairs [LFA], "Official Gazette of the Republic of Serbia", No. 116/2007, 126/2007, and 41/2009.
> (LFA 2009, Art. 17)

> Succession Act [SA], "Official Gazette of the Republic of Croatia", No. 48/03, 163/03, 35/05, 127/13, and 33/15 and 14/19.
> (SA 2019, Art. 3)

> An Act to make provision for and in connection with offences relating to offensive weapons [Offensive Weapons Act], 16th May 2019, www.legislation.gov.uk/ukpga/2019/17/pdfs/ukpga_20190017_en.pdf, last accessed 20 December 2019.
> (Offensive Weapons Act 2019)

Government decisions and decisions of the institutions

The name of the government body or institution [acronym or abbreviation], the title and number of the decision, the date of the decision passing, or the webpage and the date of the last access.

> Protector of Citizens of the Republic of Serbia [Protector of Citizens], Opinion No. 19–3635/11, 11 January 2012, https://www.ombudsman.org.rs/attachments/064_2104_Opinion%20HJC.pdf, last accessed 20 December 2019.
> (Protector of Citizens, 19–3635/11)

> U.S. Department of the Treasury [USDT], Treasury Directive No. 13–02, July 20, 1988, https://www.treasury.gov/about/role-of-treasury/orders-directives/Pages/td13-02.aspx, last assessed 20 December 2019.
> (USDT, 13–02)

Legislative acts of the European Union

The title of the legislative act, the number of the official gazette, the publication date, and the number of the page in the same format as on the *EUR-lex* website: https://eur-lex.europa.eu/homepage.html.

> Regulation (EU) No 182/2011 of the European Parliament and of the Council of 16 February 2011 laying down the rules and general principles concerning mechanisms for control by Member States of the Commission's exercise of implementing powers, OJ L 55, 28.2.2011, p. 13–18.
> (Regulation 182/2011, Art. 3)

Treaties

European Union founding treaties

Title of the treaty or title of the consolidated version of the treaty [acronym], information on the treaty retrieved from the official gazette in the same format as on the *EUR-lex* website: https://eur-lex.europa.eu/homepage.html.

> Treaty on European Union [TEU], OJ C 191, 29.7.1992, p. 1–112.
> (TEU 1992, Art. J.1)
>
> Consolidated version of the Treaty on European Union [TEU], OJ C 115, 9.5.2008, p. 13–45.
>
> (TEU 2008, Art. 11)
>
> Consolidated version of the Treaty on the Functioning of the European Union [TFEU], OJ C 202, 7.6.2016, p. 1–388.
>
> (TFEU 2016, Art. 144)

Other treaties

Title of the treaty [acronym or abbreviation], date of conclusion, UNTS volume number, and registration number on the *United Nations Treaty Collection* website: https://treaties.un.org.

> Marrakesh Agreement Establishing the World Trade Organization [Marrakesh Agreement], 15 April 1994, UNTS 1867, I-31874.
> (Marrakesh Agreement 1994)
>
> International Covenant on Civil and Political Rights [ICCPR], 19 December 1966, UNTS 999, I-14668.
> (ICCPR 1966)
>
> Treaty of Peace between the State of Israel and the Hashemite Kingdom of Jordan [Israel Jordan Peace Treaty], 26 October 1994, UNTS 2042, I-35325.
> (Israel Jordan Peace Treaty 1994)

Decisions of international organizations

The name of the international organization and its body [acronym], the decision number, the title of the decision, and the date of the decision passing.

United Nations Security Council [UNSC], S/RES/1244 (1999), Resolution 1244 (1999) Adopted by the Security Council at its 4011th meeting, on 10 June 1999.
(UNSC, S/RES/1244)

Parliamentary Assembly of the Council of Europe [PACE], Doc. 14326, Observation of the presidential election in Serbia (2 April 2017), 29 May 2017.
(PACE, Doc. 14326, para. 12)

Case law

Case law of the courts in the Republic of Serbia

The type of the act and the name of the court [acronym of the court], the case number with the date of the decision passing, the name and number of the official gazette where the decision is published – if available.

Decision of the Constitutional Court of the Republic of Serbia [CCRS], IUa-2/2009 of 13 June 2012, "Official gazette of the Republic of Serbia", No. 68/2012.
(Decision of CCRS, IUa-2/2009)

Decision of the Appellate Court in Novi Sad [ACNS], Rzr–1/16 of 27 April 2016.
(Decision of ACNS, Rzr–1/16)

Case law of the International Court of Justice

The name of the court [acronym], *the case title*, type of the decision with the date of the decision passing, the name and number of I.C.J. Reports issue where the decision is published, page number.

International Court of Justice [ICJ], *Application of the Interim Accord of 13 September 1995 (the Former Yugoslav Republic of Macedonia* v. *Greece)*, Judgment of 5 December 2011, I.C.J. Reports 2011, p. 644.
(ICJ Judgment 2011)
International Court of Justice [ICJ], *Accordance with the International Law of the Unilateral Declaration of Independence in Respect of Kosovo*, Advisory Opinion of 22 July 2010, I.C.J. Reports, p. 403.
(ICJ Advisory Opinion 2010)

Case law of the Court of Justice of the European Union

The case title, the case number, the type of the case with the date of the decision passing, ECLI.

> *United Kingdom of Great Britain and Northern Ireland v. European Parliament and Council of the European Union*, Case C-270/12, Judgment of the Court (Grand Chamber) of 22 January 2014, ECLI:EU:C:2014:18.
> (*United Kingdom of Great Britain and Northern Ireland v. European Parliament and Council of the European Union*, C-270/12) or
> (CJEU, C-270/12)

> *United Kingdom of Great Britain and Northern Ireland v. European Parliament and Council of the European Union*, Case C-270/12, Opinion of Advocate General Jääskinen delivered on 12 September 2013, ECLI:EU:C:2013:562.
> (Opinion of AG Jääskinen, C-270/12)

Case law of the European Court of Human Rights

The case title, number of the application, type of the case with the date of the judgment passing, ECLI.

> *Pronina v. Ukraine*, No. 63566/00, Judgment of the Court (Second Section) on Merits and Just Satisfaction of 18 July 2006, ECLI:CE:ECHR:2006:0718JUD006356600.
> (*Pronina v. Ukraine* 63566/00, par. 20) or
> (ECHR, 63566/00, par. 20)

Case law of other international courts and tribunals

The name of the court [acronym], the case number, *the case title*, the type of the decision with the date passing.

> International Tribunal for the Prosecution of Persons Responsible for Serious Violations of International Humanitarian Law Committed in the Territory of the Former Yugoslavia since 1991 [ICTY], Case No. IT-94-1-A-AR77, *Prosecutor v. Dusko Tadic*. Appeal Judgement on Allegations of Contempt Against Prior Counsel, Milan Vujin. Judgment of 27 February 2001.
> (*Prosecutor v. Dusko Tadic*, IT-94-1-A-AR77) or
> (ICTY, IT-94-1-A-AR77)

Archive sources

Name of the repository [acronym], title or number of the fond [acronym], box number, folder number – if available, reference code, "title of the document" – or, if it is not available, provide a short description by

answering the questions who? whom? what?, place and date – or n.d. if no date is provided.

Arhiv Srbije [AS], MID, K-T, f. 2, r93/1894, "Izveštaj Ministarstva inostranih dela o postavljanju konzula", Beograd, 19. april 1888.
(AS, MID, K-T, f. 2)
(AS, MID, f. 2) – *When the folder number is known only*

Dalhousie University Archives [DUA], Philip Girard fonds [PG], B-11, f. 3, MS-2-757.2006-024, "List of written judgements by Laskin," n.d.
(DUA, PG, B-11, f. 3)

Web sources

Surname, Name, or name of the corporate author [acronym]. Year of publication or n.d. – if the year of publication cannot be determined. "The name of the web page." *The name of the website*. Date of creation, modification, or the last access to the web page, if the date cannot be determined from the source. URL.

Bilefsky, Dan, and Ian Austen. 2019. "Trudeau Re-election Reveals Intensified Divisions in Canada." *The New York Times*. https://www.nytimes.com/2019/10/22/world/canada/trudeau-re-elected.html.
(Bilefsky and Austen 2019)

Institute for Political Studies [IPS]. n.d. "The 5th International Economic Forum on Reform, Transition and Growth." *Institute for Political Studies*. Last accessed 7 December 2019. http://www.ips.ac.rs/en/news/the-5th-international-economic-forum-on-reform-transition-and-growth/.
(Institute for Political Studies [IPS] n.d.) – *First in-text citation*
(IPS n.d.) – *Second and every subsequent citation*

Associated Press [AP]. 2019. "AP to present VoteCast results at AAPOR pooling conference." May 14, 2019. https://www.ap.org/press-releases/2019/ap-to-present-votecast-results-at-aapor-polling-conference.
(AP 2019)

TEXT FORMATTING

General guidelines for writing the manuscript

The manuscript should be written in Word, in the following manner:

- Paper size: A4;
- Margins: Normal 2.54 cm;
- Use Times New Roman font (plain letters) to write the text, unless specified otherwise;
- Line spacing: 1.5;
- Footnote line spacing: 1;
- Title font size: 14 pt;
- Subtitles font size: 12 pt;
- Text font size: 12 pt;
- Footnote font size: 10 pt;
- Tables, charts and figures font size: 10 pt;
- Use Paragraph/Special/First line at 1.27 cm;
- Text alignment: Justify;
- Font color: Automatic;
- Page numbering: Arabian numerals in lower right corner;
- Do not break the words manually by inserting hyphens to continue the word in the next line;
- Save the manuscript in the .doc format.

Research article manuscript preparation

The manuscript should be prepared in the following manner:

Name and surname of the first author*

* Footnote: E-mail address: The institutional e-mail address is strongly recommended. ORCID:

Affiliation

Name and surname of the second author**

** Footnote: E-mail address: The institutional e-mail address is strongly recommended. ORCID:

Affiliation

TITLE OF THE PAPER***

*** Footnote: if necessary, specify one of the following (or similar) data: 1) the name and number of the project; 2) the proceeding where the manuscript was presented under the same or similar title; 3) statements of gratitude.

Abstract

Abstract, within 100–250 words range, contains the subject, aim, theoretical and methodological approach, results and conclusions of the paper.

Keywords: Below the abstract, five to ten **key words** should be written. Key words should be written in roman font and separated by commas.

The manuscript can have maximally three levels of subtitles. **Subtitles** should not be numbered. They should be used in the following manner:

FIRST LEVEL SUBTITLE

Second level subtitle

Third level subtitle

Tables, charts, and figures should be inserted in the following manner:

- Above the table/chart/figure, center the name of the Table, Chart or Figure, an Arabic numeral, and the title in Times New Roman font;
- Below the table/chart/figure, the source should be cited in the following manner: 1) if the table/chart/figure is taken from another source, write down *Source:* and include the parenthetical citation information of the source; or 2) if the table/chart/figure is not taken from another source, write down *Source: Author*.

Use in-text references according to *Citing and referencing*.

Use the footnotes solely to provide remarks or broader explanations.

REFERENCES

References should be listed after the text of the paper, before the Resume in the following manner:

- the first line of each reference should be left indented, and the remaining lines should be placed as hanging by 1.27 cm using the option Paragraph/Special/Hanging;
- all the references should be listed together, without separating legal acts of archives;
- the references should not be numbered;
- list only the references used in the text.

After the reference list, write the name and surname of the author, the tile of the paper and resume in Serbian in the following manner:

Име и презиме првог аутора*

*Фуснота: Имејл-адреса аутора: Препоручује се навођење институционалне имејл-адресе аутора. ORCID:

Установа запослења

Име и презиме другог аутора**

** Фуснота: Имејл-адреса аутора: Препоручује се навођење институционалне имејл-адресе аутора. ORCID:

Установа запослења

НАСЛОВ РАДА***

*** Фуснота: по потреби, навести један од следећих (или сличних) података: 1) назив и број пројекта у оквиру кога је чланак написан; 2) да је рад претходно изложен на научном скупу у виду усменог саопштења под истим или сличним називом 3) исказ захвалности.

Резиме

Resume (Резиме) up to 1/10 length of the paper contains the results and conclusions of the paper which are presented in greater scope than in the abstract.

Keywords (Кључне речи): Keywords should be written in Times New Roman font and separated by commas.

Review preparation

A review should be prepared in the same manner as the research article, but leaving out the abstract, keywords, resume, or book cover.

Book review preparation

When writing book reviews, split the text into **two columns**. Book reviews should be prepared in the following manner:

Name and surname of the author*	
* In the footnote: E-mail address: The institutional e-mail address is strongly recommended. ORCID:	Below the title **place the image of the front cover;**
Affiliation	Below the image of the front cover list the book details according to the following rule:
TITLE OF THE BOOK REVIEW***	Name and surname of the author. Year of publication. *Title of the book*. Place of publication: Publisher, total number of pages.
*** Footnote: if necessary, specify one of the following (or similar) data: 1) the name and number of the project; 2) the proceeding where the manuscript was presented under the same or similar title; 3) statements of gratitude.	**The text** of the book review should be prepared following the guidelines of the research article preparation.

REVIEWERS' GUIDELINES

The role of reviewers is to contribute to maintaining the high quality of our journal. All submitted manuscripts undergo a double-blind peer review, ensuring anonymity in both directions.

Requests for peer review are submitted through the SCIndeks Assistant system. The deadline for accepting or declining the review request is seven days from receipt, while the review itself must be completed within two weeks of receiving the request. The content of the review is confidential and must not be disclosed to individuals outside the journal's Editorial board. If, at any point, a reviewer becomes aware of any conflict of interest related to the manuscript under review, they are required to inform the Editorial board as soon as possible.

When reviewing a manuscript, the reviewer is required to complete the attached review form:

Title of the manuscript:

Relevance, social, and scientific significance of the topic under consideration:

To what extent has the author clearly outlined the theoretical and methodological approach in the manuscript?

Is the manuscript based on contemporary and relevant literature, particularly in terms of the author's use of the latest research published in scientific journals and conference proceedings (especially in political science journals and proceedings)?

Scientific and social contribution of the manuscript. General comments on the quality of the manuscript:

Suggestions for the author on how to improve the quality of the manuscript, if necessary:

Please select one of the recommendations for categorizing the manuscript:

1. Original research article
2. Review article
3. Scientific critique, polemic, or commentary

Please select one of the recommendations regarding the publication of this manuscript:

1. Publish without revision
2. Publish with minor revisions

3. After revision, submit for a new round of review
4. Reject

Additional comments for the editor regarding ethical concerns (e.g., plagiarism, fraud) or other aspects of the article that may assist in making a final decision on its status.

Date of review:

Reviewer's name, surname, and academic title:

The use of chatbots and other artificial intelligence software is strictly prohibited in the preparation of reviews.

SERBIAN POLITICAL THOUGHT 2025 – LIST OF REVIEWERS

SPT 1/25

1. **Veran Stančetić**, Associate Professor, Faculty of Political Sciences, University of Belgrade;
2. **Vanja Glišin**, Research Associate, Institute for Political Studies, Belgrade;
3. **Nikola Mladenović**, Research Associate, Institute for Political Studies, Belgrade;
4. **Ana Čović**, Principal Research Fellow, Institute for Political Studies, Belgrade;
5. **Igor Janev**, Principal Research Fellow, Institute for Political Studies, Belgrade;
6. **Dušan Dostanić**, Research Associate, Institute for Political Studies, Belgrade;
7. **Rastko Lompar**, Research Associate, Institute for Balkan Studies, Serbian Academy of Sciences and Arts, Belgrade;
8. **Tijana Perić Diligenski**, Senior Research Associate, Institute for Political Studies, Belgrade;
9. **Jovan Janjić**, Research Associate, Institute for Political Studies, Belgrade;
10. **Petar Matić**, Research Associate, Institute for Political Studies, Belgrade;
11. **Milena Pešić**, Research Associate, Institute for Political Studies, Belgrade;
12. **Milovan Subotić**, Senior Research Associate, Strategic Research Institute, University of Defence, Belgrade;
13. **Srđan Starčević**, Associate Professor, University of Defense, Belgrade;

SPT 2/25

1. **Zoran Milošević**, Principal Research Fellow, Institute for Political Studies, Belgrade;
2. **Marija Đorić**, Senior Research Associate, Institute for Political Studies, Belgrade;
3. **Mario Kalik**, Research Associate, Institute for Political Studies, Belgrade;

4. **Ana Čović**, Principal Research Fellow, Institute for Political Studies, Belgrade;
5. **Bogdana M. Stjepanović**, Research Associate, Institute for Comparative Law, Belgrade;
6. **Dragan Trailović**, Research Associate, Institute of International Politics and Economics, Belgrade;
7. **Dragan Stanar**, Research Associate, Strategic Research Institute, University of Defense, Belgrade;
8. **Rajko Petrović**, Research Associate, Institute of European Studies, Belgrade;
9. **Dušan Ilić**, Research Associate, Institute of European Studies, Belgrade;
10. **Vladimir Ajzenhamer**, Assistant Professor, Faculty of Security Studies, University of Belgrade;
11. **Petar Matić**, Research Associate, Institute for Political Studies, Belgrade;
12. **Momčilo Subotić**, retired Principal Research Fellow, Institute for Political Studies, Belgrade;
13. **Jelena Todorović Lazić**, Senior Research Associate, Institute for Political Studies, Belgrade;
14. **Marta Nešković**, Research Associate, Institute for Political Studies, Belgrade;

SPT 3/2025

1. **Boris Latinović**, Assistant Professor, Faculty of Economics and Finance, University "UNION – Nikola Tesla", Belgrade;
2. **Violeta Talović**, Full Professor, Faculty of Diplomacy and Security, University "UNION – Nikola Tesla", Belgrade;
3. **Dušan Proroković**, Senior Research Associate, Institute of International Politics and Economics, Belgrade;
4. **Jelena Todorović Lazić**, Senior Research Associate, Institute for Political Studies, Belgrade;
5. **Bogdana M. Stjepanović**, Research Associate, Institute for Political Studies, Belgrade;
6. **Jelena Ž. Kostić**, Senior Research Associate, Institute for Comparative Law, Belgrade;
7. **Vladan Stanković**, Senior Research Associate, Institute for Political Studies, Belgrade;

8. **Marija Đorić**, Principal Research Fellow, Institute for Political Studies, Belgrade;
9. **Sanja Lazarević Radak**, Principal Research Fellow, Institute for Political Studies, Belgrade;
10. **Ivan Matić**, Research Associate, Institute for Political Studies, Belgrade;
11. **Mladen Lišanin**, Research Associate, Institute of International Politics and Economics, Belgrade;
12. **Saša Mijalković**, Full Professor, University of Criminal Investigation and Police Studies, Belgrade;
13. **Marko Pejković**, Research Associate, Institute for Political Studies, Belgrade;
14. **Srđan Starčević**, Associate Professor, University of Defense, Belgrade;
15. **Milovan Subotić**, Senior Research Associate, Strategic Research Institute, University of Defence, Belgrade;
16. **Aleksandra Bulatović**, Senior Research Associate, Institute for Philosophy and Social Theory, University of Belgrade;

SPT 4/2025

1. **Aleksandar Lukić**, Senior Research Associate, Institute for Recent History of Serbia, Belgrade;
2. **Slobodan Selinić**, Principal Research Fellow, Institute for Recent History of Serbia, Belgrade;
3. **Mijodrag Radojević**, Research Associate, Institute for Political Studies, Belgrade;
4. **Bogdana M. Stjepanović**, Research Associate, Institute for Political Studies, Belgrade;
5. **Duško Bodroža**, Research Associate, Institute of Economic Sciences, Belgrade;
6. **Marija Đorić**, Principal Research Fellow, Institute for Political Studies, Belgrade;
7. **Milovan Subotić**, Senior Research Associate, Strategic Research Institute, University of Defence, Belgrade;
8. **Sanja Jelisevac Trošić**, Senior Research Associate, Institute of International Politics and Economics, Belgrade;
9. **Mitko Arnaudov**, Research Associate, Institute of International Politics and Economics, Belgrade;

10. **Miloš Šolaja**, Full Professor, Faculty of Political Sciences, University of Banja Luka;
11. **Dragan Stanar**, Research Associate, Strategic Research Institute, University of Defense, Belgrade;
12. **Miloš Milenković**, Full Professor, Faculty of Philosophy, University of Belgrade;
13. **Neven Obradović**, Assistant Professor, Faculty of Philosophy, University of Niš;
14. **Siniša Atlagić**, Full Professor, Faculty of Political Sciences, University of Belgrade;
15. **Aleksandra Bulatović**, Senior Research Associate, Institute for Philosophy and Social Theory, University of Belgrade;
16. **Vesna Stanković Pejnović**, Senior Research Associate, Institute for Political Studies, Belgrade;
17. **Srđan Starčević**, Associate Professor, University of Defence, Belgrade;
18. **Ivana Luknar**, Research Associate, Institute for Political Studies, Belgrade;
19. **Svetlana Stanarević**, Associate Professor, Faculty of Security Studies, University of Belgrade;

SPT 5/2025

1. **Marija Đorić**, Principal Research Fellow, Institute for Political Studies, Belgrade;
2. **Milovan Subotić**, Senior Research Associate, Strategic Research Institute, University of Defence, Belgrade;
3. **Petar Matić**, Research Associate, Institute for Political Studies, Belgrade;
4. **Nikola Mladenović**, Research Associate, Institute for Political Studies, Belgrade;
5. **Vladan Stanković**, Senior Research Associate, Institute for Political Studies, Belgrade;
6. **Marko Pejković**, Research Associate, Institute for Political Studies, Belgrade;
7. **Kirill Shevchenko**, Professor, Branch of Russian Social University in Minsk, Minsk;
8. **Irina Antanasijević**, Full Professor, Faculty of Philology, University of Belgrade;

9. **Dragan Trailović**, Research Associate, Institute of International Politics and Economics, Belgrade;
10. **Milena Pešić**, Research Associate, Institute for Political Studies, Belgrade;
11. **Filip Čukljević**, Research Associate, Institute of Philosophy, Faculty of Philosophy, University of Belgrade;
12. **Aleksandra Vučković**, Research Associate, Institute of Philosophy, Faculty of Philosophy, University of Belgrade;
13. **Boris Bursać**, Research Associate, Institute for Serbian Culture, Priština-Leposavić;
14. **Dragan Tančić**, Senior Research Associate, Institute for Serbian Culture, Priština-Leposavić;

SPT 6/2025

1. **Dušan Proroković**, Senior Research Associate, Institute of International Politics and Economics, Belgrade;
2. **Dragan Stanar**, Research Associate, Strategic Research Institute, University of Defense, Belgrade;
3. **Stevan Rapaić**, Senior Research Associate, Institute for Political Studies, Belgrade;
4. **Marina L. Galas**, Professor, Chief Researcher, Department of Political Science, Financial University under the Government of the Russian Federation, Moscow;
5. **Nikola Avrejski**, Full Professor, Institute for Historical Studies, Bulgarian Academy of Sciences, Sofia;
6. **Dragana Lazić**, Associate Professor, Faculty of Business and Law, University "MB", Belgrade;
7. **Mihajlo Vučić**, Senior Research Associate, Institute of International Politics and Economics, Belgrade;
8. **Petar Matić**, Research Associate, Institute for Political Studies, Belgrade;
9. **Nenad Putnik**, Full Professor, Faculty of Security Studies, University of Belgrade;
10. **Ayjaz Wani**, Fellow, Observer Research Foundation, New Delhi;
11. **Ljubiša Despotović**, Principal Research Fellow, Institute for Political Studies, Belgrade;
12. **Aleksandar Gajić**, Principal Research Fellow, Institute for European Studies, Belgrade;

13. **Milan Žarković**, retired Full Professor, University of Criminal Investigation and Police Studies, Belgrade;
14. **Božidar Banović**, Full Professor, Faculty of Security Studies, University of Belgrade;
15. **Rifat Olgun**, Associate Professor, Vocational School of Serik G-S. Süral, Akdeniz University, Antalya;
16. **Mert Çakir**, Assistant Professor, Faculty of Architecture, Süleyman Demirel University, Isparta;
17. **Natalija Perišić**, Full Professor, Faculty of Political Sciences, University of Belgrade;
18. **Nataša Mrvić Petrović**, Principal Research Fellow, Institute for Comparative Law, Belgrade;
19. **Srđan Milašinović**, Full Professor, University of Criminal Investigation and Police Studies, Belgrade;
20. **Milan Čabarkapa**, Assistant Professor, Faculty of Engineering, University of Kragujevac;
21. **Miloš Agatonović**, Senior Lecturer, Academy of Applied Preschool Teaching and Health Studies, Kruševac.

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

32

СРПСКА политичка мисао = Serbian Political Thought / главни и одговорни уредник Миша Стојадиновић. - Vol. 1, бр. 1/4 (1994)- . - Београд : Институт за политичке студије, 1994- (Београд : Донат граф). - 23 cm

Двомесечно. - Текст на срп. и енгл. језику. - Преузео је: Serbian Political Thought = ISSN 1450-5460. - Друго издање на другом медијуму: Српска политичка мисао (CD-ROM изд.) =

ISSN 1452-3108

ISSN 0354-5989 = Српска политичка мисао

COBISS.SR-ID 102782215

1994

СРПСКА
ПОЛИТИЧКА МИСАО

ISSN 0354-5989 UDC 32 year XXXII vol. 94 No. 6/2025.