TEME, Vol. L, Nº 2, April – June 2026, pp. 401–416

Original research paper
Received: January 27, 2026
Revised: January 28, 2026
Accepted: March 7, 2026

https://doi.org/10.22190/TEME260127026G
UDC 004.8:351.817
004+007]:004.056
343.53:004.7

ARTIFICIAL INTELLIGENCE AND CYBERSECURITY: SOCIAL AND TECHNOLOGICAL CHALLENGES

Vanja Glišin[1*], Srđan Milašinović[2], Aleksandar Aleksić[3]

[1]Institute for Political Studies, Belgrade, Serbia
[3]University of Criminal Investigation and Police Studies, Belgrade, Serbia
[3]Faculty of Diplomacy and Security, Belgrade, Serbia

ORCID iDs: Vanja Glišin https://orcid.org/0000-0002-7747-6205
Srđan Milašinović https://orcid.org/0000-0002-8251-339X
Aleksandar Aleksić https://orcid.org/0009-0009-2669-9781

Abstract

The scientific paper analyses the role and capabilities of artificial intelligence (AI) in improving cybersecurity, with a special focus on the social, educational and technological challenges of preventive application. The first chapter is dedicated to the conceptual and theoretical definition of AI, cyberspace and cybersecurity, as well as the logical connection of the aforementioned phenomena and processes. The ability of AI to process large amounts of data, recognise behavioural patterns, identify threats and offer responses in real time suggests that it is an important instrument for improving cybersecurity. The second chapter is dedicated to the role of AI in preventing cyber threats with the aim of presenting the advantages of using new technologies in preventing, detecting, protecting, responding and maintaining the security of systems and their components. In this context, in addition to analysing the technological aspects of the use of AI, attention is paid to the social and educational challenges that are indispensable for the efficient functioning and protection of systems in cyberspace. Using descriptive and explanatory methods of analysis and synthesis methods, the aim is to present that prevention in cyberspace is a complex technological, but also social process, which requires an integrated approach.

Key words: artificial intelligence, cyberspace, cybersecurity, prevention, society.

ВЕШТАЧКА ИНТЕЛИГЕНЦИЈА И САЈБЕР БЕЗБЕДНОСТ: ДРУШТВЕНИ И ТЕХНОЛОШКИ ИЗАЗОВИ

Апстракт

У научном раду се анализира улога и капацитети вештачке интелигенције (ВИ) у унапређењу сајбер безбедности, са посебним освртом на друштвене, обра-

* Corresponding author: Vanja Glišin, Institute for Political Studies, Dobrinjska Street 11, 11000 Belgrade, Serbia, vanja.glisin@ips.ac.rs

зовне и технолошке изазове превентивне примене. Прво поглавље посвећено је појмовном и теоријском одређењу ВИ, сајбер простора и сајбер безбедности, као и логичном повезивању наведених појава и процеса. Способност ВИ да обрађује велику количину података, препознаје обрасце понашања, идентификује претње и нуди одговоре у реалном времену, наводи на претпоставку да је реч о значајном инструменту за унапређење сајбер безбедности. Друго поглавље посвећено је улози ВИ у превенцији сајбер претњи са циљем да се презентују предности употребе нових технологија у превенцији, детекцији, заштити, реаговању и одржавању безбедности система и његових компоненти. У том контексту, осим анализе технолошких аспеката употребе ВИ, пажња се посвећује друштвеним и образовним изазовима који су неизоставни за ефикасно функционисање и заштиту система у сајбер простору. Коришћењем дескриптивне и експланаторне методе анализе и методе синтезе, циљ је да се презентује да је превенција у сајбер простору сложен технолошки, али и друштвени процес, који захтева интегрисани приступ.

Кључне речи: вештачка интелигенција, сајбер простор, сајбер безбедност, превенција, друштво.

INTRODUCTION

The development of modern technologies and artificial intelligence is reflected in the accelerated digitalisation of state and social systems, which inevitably leads to an increasing dependence of social, economic, institutional and other functions on cyberspace. Artificial intelligence (AI) is emerging as one of the key drivers of the digital transformation of security, especially in the field of cybersecurity. The ability of AI to process large amounts of data, detect, analyse and assess the behaviour of individuals and potential consequences, perform predictive analytics and offer responses in different circumstances and scenarios, allows the approach to security risks to change from a reactive to a proactive model, which would contribute to preventive actions. This means that AI, due to its speed, complexity and adaptability, can be used as a preventive instrument in the context of modern security challenges and threats. When analysing the modern security environment, it is necessary to take into account cyberspace, which, in addition to its function in the context of the development of information and communication systems, is becoming the central place for new security challenges and risks, including cybercrime, data misuse, hybrid threats, attacks on state systems, and more (Despotović & Glišin, 2025, Glišin, 2025; Glišin & Petrović, 2025; Glišin & Milašinović, 2025; Glišin, Milašinović & Lečić, 2025; Petrović & Glišin, 2025; Milojević, Milašinović & Milojković, 2025).

Theoretical consideration of the above topic is justified, since for practical consideration of the function and role of AI in the context of cybersecurity, a conceptual and theoretical understanding of the phenomena and processes we are analysing is first necessary. In this regard, the application of AI in cybersecurity is not only a technological issue, but is also conditioned by a number of social, educational and economic factors,

which are reflected in the efficiency and usability of AI-based systems. Social factors concern citizens' trust in institutions, human rights, privacy protection, perception of mass surveillance and its acceptability. In this regard, the educational aspect is very important, since the effective and responsible application of AI in cybersecurity depends on the level of education. Initially, this refers to experts and decision-makers at the state level, and then to the general public and users of new technologies. Lack of digital and security literacy and insufficient understanding of the ethical and legal implications of AI can lead to negative consequences and misuse of technology (Matković, Novakov & Glišin, 2023; Glišin, Milašinović & Lečić, 2025; Jevtović, Marić & Milašinović, 2025). The economic factor concerns the financial capabilities of the state and society to afford modern AI-based systems, i.e. to provide adequate education and protection.

The aim of the paper is to analyse the role of AI in cybersecurity, with a special focus on the social and technological challenges of preventive use in order to maintain security in cyberspace. It is based on the assumption that prevention in cyberspace cannot be understood only as a technical-technological process, but as a complex socio-technological mechanism based on institutional, educational and normative aspects. We will use the descriptive and explanatory methods of analysis and the synthesis method. The descriptive method of analysis is used with the aim of describing the results of theoretical research on the basic concepts, phenomena and processes related to the topic. Based on this, we use explanatory analysis to explain the connection between AI and cybersecurity, i.e. the potential advantages and disadvantages of prevention in cyberspace. Through a qualitative analysis of relevant scientific literature, we will offer answers to the subject and problem of the research.

DEFINING ARTIFICIAL INTELLIGENCE

The term Artificial Intelligence was defined by the American mathematician and computer scientist John McCarthy in 1955, which marked the formal beginning of research in this field (McCarthy, 2007; Glišin, 2025). McCarthy defines AI as an engineering discipline and defines it as: "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). This definition positions AI as a separate research field, but also as an area closely related to the understanding of human cognitive abilities.

From the very beginning, the development of AI has implied a multidisciplinary approach, which has led to the existence of different definitions and conceptual interpretations depending on the scientific field from which the concept is approached – computer science, philosophy, law, po-

litical science, sociology, military science, etc. (Russell & Norvig, 2010; Scherer, 2016; Glišin, 2025; Glišin & Petrović, 2025; Glišin & Milašinović, 2025; Glišin, Milašinović & Lečić, 2025, Petrović & Glišin, 2025; Beriša & Pavić, 2024). It is precisely the diversity of approaches that leads to conceptual ambiguity, since AI is essentially an attempt to simulate processes that are not fully explained even in a human context, primarily intelligence, thinking and decision-making (Sheikh, Prins & Schrijvers, 2023, p. 16; Despotović & Glišin, 2024). In this sense, some authors define AI as systems that 'think like humans,' which further complicates the theoretical debate (Haugeland, 1985).

More modern normative and policy approaches seek to offer more functional and operationalised definitions. Thus, the High-Level Expert Group on Artificial Intelligence (AI HLEG) defines AI as "systems that exhibit intelligent behaviour by analysing their environment and taking certain actions, with varying degrees of autonomy, in order to achieve predefined goals" (High-Level Expert Group on Artificial Intelligence, 2019). This understanding emphasises the ability of systems to operate in a dynamic environment, which is particularly important in the domain of cybersecurity and decision-making in difficult circumstances in real time (Glišin, 2025; Glišin & Petrović, 2025). McKendrick (2019) offers additional clarification of the concept in his definition of AI:

> 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, p. 6)

In contemporary theoretical frameworks, the central place in defining AI is occupied by the capabilities of learning and adaptation, which allow systems to improve their performance without direct and continuous human intervention (Shalev-Shwartz & Ben-David, 2014; Despotović & Glišin, 2024). The concept of 'machine learning' refers to such a type of system that, through data analysis and experience, builds its own decision-making and prediction models, without explicitly programming each individual step (Smola & Vishwanathan, 2008; Subotić, 2023). In this way, AI is increasingly viewed less as a static tool, and more as a dynamic system, whose basic functions are "learning, planning, reasoning, decision-making and problem-solving" (Prlja, Gasmi & Korać, 2021, p. 66; Glišin & Petrović, 2025).

Given the different levels of complexity and application, we find numerous classifications of AI in the literature. The most common distinc-

tion is made between the so-called weak AI, which is focused on solving specific and predefined tasks, and strong AI, which involves systems with general cognitive abilities comparable to humans (Prlja, Gasmi & Korać, 2021, p. 63; Glišin, 2025). In addition, AI can also be classified according to the areas of application into: 1) specialised AI - used only in certain areas; 2) general AI - as a general intelligence ability that can be applied in various areas (economy, business, education, medicine, military industry, etc.) and 3) super-intelligence (Kaplan & Haenlein, 2018; Wang & Siau, 2019). More complex typologies of AI, such as the one proposed by Alex Bekker: "1) interactive AI (example: personal assistants like Siri, Cortana, and Alexa); 2) functional AI (robots); 3) analytical AI (data analysis, machine learning); 4) textual AI (text recognition, speech-to-text conversion); and 5) visual AI (augmented reality technology)", indicate the functional diversity of AI (Bekker, 2019).

AI defined and conceptualised in this way represents the technological basis for its increasingly widespread application in the field of cybersecurity. The ability of AI systems to learn from large amounts of data, recognise patterns, predict behaviour, and make decisions in real time makes them particularly suitable for identifying and preventing cyber threats in a complex and dynamic digital environment (Glišin, 2025; Glišin & Petrović, 2025). Unlike traditional security mechanisms, which are mostly based on predefined rules, AI enables a proactive approach to cybersecurity, focused on early detection of risks and prevention or mitigation of potential consequences. Therefore, understanding the concept and characteristics of AI is a necessary theoretical prerequisite for analysing its role in modern models of prevention and protection in cyberspace.

CYBERSPACE AND CYBERSECURITY – A CONCEPTUAL UNDERSTANDING

Throughout history, new knowledge and technological developments have led to the discovery of new dimensions of space. From the initial possibilities of the existence and use of land as space, technological developments have enabled the exploration and use of water, air, and space. The last in the series is cyberspace, which represents a specific domain of the modern digital environment created by the development and global networking of information and communication systems. The term was coined by William Gibson in his work "Burning Chrome" from 1982, and elaborated in more detail in the novel "Neuromancer" from 1984 (Gibson, 2017). Cyberspace is "the environment created by the confluence of cooperative networks of computers, information systems, and telecommunication infrastructures commonly referred to as the Internet and the World Wide Web" (Sharp, 1999). However, in the literature we encounter a broader understanding of cyberspace, according to which it is not limited to the Inter-

net, but includes interdependent networks of computer and communication systems, the data generated, stored and processed in them, as well as the interactions of users and organisations in that environment (NSPD-54, 2008; Kuehl, 2009; Despotović & Glišin, 2024). Cyberspace was created thanks to the development of technology, that is, it is an artificially constructed, but real space of action, which has its own technical characteristics, as well as the dynamics of functioning. In this way, cyberspace differs from previous dimensions of space, which existed before the development of technology, but were further explored and used as such thanks to that development. According to Joseph Nye, "the geography of cyberspace is much more fluid than other environments. Mountains and oceans are difficult to move, but parts of cyberspace can be turned on or off with the flick of a switch" (Naj, 2012, p. 154).

In the last few decades, we have encountered different interpretations and definitions of cyberspace, which is, among other things, due to the rapid development of technology and sophisticated systems, as well as the acquisition of new knowledge about that space. According to the definition of the National Strategy to Secure Cyberspace from 2003, cyberspace is "the nervous system - the control system of the country... composed of hundreds of thousands of interconnected computers, servers, routers, switches, and fiber optic cables that allow our critical infrastructures to work" (White House, 2003). Cyberspace is a set of multiple activities, complex and layered networks, intersecting physical and virtual forms, uniquely coordinated and managed (Despotović & Glišin, 2024, p. 90). It is characterised by transnationality, speed of information flow, system complexity, accessibility, borderlessness and functionality, which, on the one hand, offers numerous opportunities for constructive action, but, on the other hand, can be a space suitable for various forms of abuse. Dan Kuehl (2009) offered the following definition:

> cyberspace is a global domain within the information environment whose distinctive and unique character is framed by the use of electronics and the electromagnetic spectrum to create, store, modify, exchange, and exploit information via interdependent and interconnected networks using information-communication technologies.
>
> (Kuehl, 2009, p. 28)

The scope for abuse is enormous, since the one who controls and manages cyberspace creates the so-called cyber power, which is "the ability to achieve a desired outcome through the use of electronically interconnected information resources of the cyber domain" (Naj, 2012, p. 153). Therefore, we are talking about actors who have the ability to control and influence multiple networks of information platforms, such as the Internet, satellites and mobile systems, as far as all those who have access to these platforms are concerned (Despotović & Glišin, 2024, p. 90). Therefore, the

analysis of actions in the digital environment requires the recognition and analysis of actors, their interests, patterns of behaviour, ways of communication, etc., which is a very demanding and extensive job, which is why we believe that the adequate application of AI is an important tool for all of the above. In the book "Geopolitics and Cryptopolitics," we point out three groups of actors who "collide in cyberspace and fight for cyber power – 1. governments; 2. organisations with highly structured networks; and 3. individuals and lightly structured networks" (Despotović & Glišin, 2024, p. 90-91). The fight in cyberspace is asymmetrical, everyone can be a target of attack and everyone can achieve certain effects with their attacks depending on their capacity and resources. For this reason, the USA has repeatedly emphasised that "the Internet is a state strategic space and that any threat to Internet security will be considered aggression" (Živen, 2019, p. 44).

In contemporary international relations, cyberspace is recognised as a new strategic environment, along with land, sea, air, and space, especially in the context of the development of AI. Therefore, cybersecurity is indispensable, as it is a set of technical, organisational and strategic measures aimed at protecting digital infrastructure, data, processes and entities in cyberspace. In the broadest sense, cybersecurity aims to preserve three fundamental principles known as the CIA triad: 1. confidentiality; integrity and availability (Samonas & Coss, 2014). The realisation of these principles is a basic prerequisite for the reliable functioning of digital systems, but also for the preservation of the state, institutions and users in cyberspace. The concept of cybersecurity "involves identifying potential vulnerabilities and instituting effective strategies to minimise the effect of possible threats" (Schiliro, 2022, p. 2).

Cybersecurity is increasingly viewed through the prism of AI as a power multiplier in cyberspace. This is very important to note, since AI, on the one hand, contributes to increasing the resilience of information systems through the automation of security responses to threats, while, on the other hand, it allows weaker actors to cause disproportionate damage in cyberspace using tools enhanced by AI technology (Glišin, 2025; Glišin & Milašinović, 2025; Glišin & Petrović, 2025).

There is no universal definition of cybersecurity, so we will highlight the following, which, in addition to the above, refers to the essence of system protection:

> Cybersecurity is a collection of best practices, including infrastructural changescritical in ensuring the integrity of a computer system. The concept entailsprotecting critical installations and sensitive information from potential digitalattacks. The measures are developed to combat threats from malicious attackerstargeting the networks and applications originating from without or within theorganization.
>
> (Abomhara & Koien, 2015)

The key functions of cybersecurity are: 1. Prevent; 2. Detect and 3. Respond. Prevention is a key element and involves taking all measures and investing in infrastructure, professionals, software and hardware to achieve goals (Schiliro, 2022). Establishing adequate system protection prevents potential unwanted access to the network and causing damage. Detection involves observing and identifying potential threats, in order to enable an adequate response and protection of the system (Kaur, Gabrijelčić & Klobučar, 2023). Therefore, constant monitoring of the system is carried out with the aim of timely detection of potential risks and threats. Response is an indispensable element that involves taking actions to prevent and/or mitigate the effects of malicious activities directed against the system in cyberspace (Schiliro, 2022). Timely detection enables effective response, which is why we believe that AI technology is very important for improving cybersecurity. Other sources mention the functions 1. Identify, 2. Protect, 3. Detect, 4. Respond and 5. Recover (Kaur, Gabrijelčić & Klobučar, 2023; NIST, 2025). This is a more complex structure that we have presented in Figure 1. The recovery function involves identifying damaged parts of the system due to attacks in cyberspace and taking activities to restore it, in order to normalise work and ensure cybersecurity. The protection function includes various measures and activities to increase the level of system security (Kaur, Gabrijelčić & Klobučar, 2023; NIST, 2025). In order for this to work, in addition to developing technology and improving the system in a technological sense, it is important to improve educational and professional staff, as well as raise social awareness about cyberspace and cybersecurity.

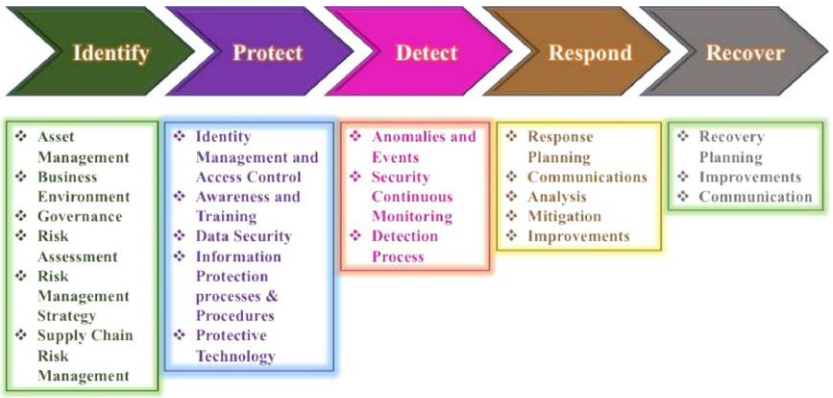

Figure 1. Functions of Cybersecurity
Source: Kaur, Gabrijelčić & Klobučar, 2023.

THE ROLE OF AI IN CYBER THREAT PREVENTION AND PRESERVATION OF CYBER SECURITY

The modern security environment in cyberspace is characterised by high dynamism, continuous growth of challenges and threats, asymmetric and sophisticated cyber attacks, which require a proactive approach and improvement of traditional protection models. In the last two decades, there has been increasing talk of network-centric wars, which, thanks to the rapid transfer of information, decentralisation of management, indoctrination of conflict participants and the creation of new channels for decision-making, achieve notable results. The idea of network-centric war is to connect everything on the Internet and make it interdependent, in order to further impose a desirable model of behaviour on everyone (with the appearance of freedom) (Despotović & Glišin, 2024, pp. 84-86). Therefore, "the implementation of the network represents the deprivation of states, peoples, armies and governments of the world, independence, sovereignty and subjectivity – their transformation into strictly managed and programmed mechanisms" (Dugin & Savin, 2018, p. 78). Control in the coordinates of network fields is established through the proper use of the so-called network code, which serves to "encode, decode and recode essential content, structures and information" (Despotović & Glišin, 2024, p. 89). It is important to note that network wars are not exclusively a struggle for cyberspace and virtual power, but also relate to establishing control over real spaces. The goal of network war is to establish control over everyone, including allies. The interconnected elements of a single network, among which there is a continuous exchange of information, are: communication systems, combat units, information support, public opinion formation, diplomatic steps, social processes, intelligence and counterintelligence, ethnopsychology, religious and collective psychology, economic support, academic science, technical innovations, and others (Dugin & Savin, 2018, p. 73). Unwanted events on the network can be grouped into four categories: (1) attacks – for the purpose of collecting data or for the purpose of imitation (e.g., IP address spoofing); (2) blocking – active interference and disruption of the operation of the operating system; (3) disclosure – disclosure of information, and (4) server intrusions – unauthorised access (Stevanović & Đurđević, 2019, pp. 53-54). The above points to the complexity and multidimensionality of security challenges in cyberspace threats, which is why it is necessary to continuously harmonise and improve defensive systems and train, first of all, professional personnel, and then the entire society, to prevent and/or reduce potential damage caused by cyber attacks through responsible behaviour.

Research confirms that conventional cybersecurity prevention technologies are ineffective in combating modern threats in cyberspace because they use fixed algorithms and physical devices such as sensors and detectors (Wiafe *et al*., 2020). Hence the need to incorporate AI into cyberspace

protection systems in order to adequately monitor and respond to the transformation of security challenges and threats. The National Institute of Standards and Technologies (NIST) encourages the use of more proactive and adaptive approaches by moving to real-time assessments, continuous monitoring and analytics (Kaur, Gabrijelčić & Klobučar, 2023; NIST, 2025). As we have already mentioned, AI enables the analysis and processing of large amounts of data in real time, recognises complex behavioural patterns and identifies potential threats, offering solutions and responses, which is why it is increasingly used for military and security purposes (Glišin, 2025; Glišin & Milašinović, 2025; Glišin & Petrović, 2025; Petrović & Glišin, 2025). "Development of intelligent systems that adapt to evolving threats and ensuring user safety holds the key to the future of cybersecurity" (Thakkar & Lohiya, 2020), some authors state. The importance of AI in the cybersecurity market is evidenced by the estimate that its value will grow to USD 102.78 billion by 2032 (Edris, 2025).

Given the rapid transformation of cyber threats, it is necessary to establish preventive defence capabilities in order to minimise the risk of excessive damage and losses. Therefore, the goal is to enable systems to predict, recognise and neutralise threats before damage occurs. Predictive analytics is one of the most important elements in protecting systems in cyberspace, thanks to machine learning algorithms that create possible scenarios, i.e. predict events (Chappell, 2020). By integrating and continuously monitoring events and data with analytics and learning from previous experiences, AI tools enable preventive action and a higher level of security in cyberspace (Hozouri, Mirzaei & Effatparvar, 2025). In this regard, the Intrusion Detection System (IDS) was created, "designed to detect intrusions and potential network or system abuse, send prompt alarms to managers, and document and log events for future analysis" (Hozouri, Mirzaei & Effatparvar, 2025, p. 3). Research in the field of IDS shows that machine and deep learning models can significantly improve the early detection and prevention of Zero-Day attacks, i.e. attacks that are not based on a previously known pattern (Buczak & Guven, 2016). IDS provides an additional layer of system defence and, with further development, will be able to respond to increasingly complex threats in cyberspace. Therefore, AI contributes to early warning and prediction of potential challenges and threats, which is an inevitable preventive role and a way to minimise risk (Glišin, 2025). However, given that AI has the ability to act and adapt in real time, we believe that automating responses to threats in real time and preparing adequate defences is a form of prevention (Schiliro, 2022). Timely optimisation of defences at a time when unwanted activities in cyberspace have already occurred adequately protects the system from potential damage.

The use of AI in cyber threat prevention is not only about preventing specific attacks, but also about early detection of vulnerable parts of the system, predicting different scenarios, and continuously adapting the sys-

tem to real and dynamic circumstances (Zhang *et al*., 2022). By applying machine learning, deep learning, and predictive analytics techniques, AI systems can identify shortcomings and weaknesses in the system, which allows for timely response and reduction of potential damage (Martínez Torres, Iglesias Comesaña, & García-Nieto, 2019; Wiafe *et al*., 2020).

We have identified two challenges that may arise in the context of the functioning and acceptance of AI systems: technological and social challenges. Given that AI systems analyse and process large data sets, it is possible that incomplete, biased, or outdated data can lead to false positives or false negatives, ultimately preventing adequate responses to security threats (Sangwan, Badr, & Srinivasan, 2023; Buczak, & Guven, 2016). In the context of cybersecurity, such failures can lead to erroneous interventions or the omission of real threats, thereby reducing the preventive role and eroding trust in AI systems (Sangwan, Badr, & Srinivasan, 2023; Buczak & Guven, 2016). In addition, the problem can arise when AI systems are increasingly automated and the human factor is marginalised, especially in the context of ethical judgment or political-security responsibility (Bostrom & Yudkowsky, 2011). Therefore, international standards and recommendations emphasise the Human in the Loop model, in order to ensure a balance between automation and human control (Crootof, Kaminski & Price II, 2023).

Societal challenges occupy a significant place when talking about the functioning of AI systems in cybersecurity. The application of advanced surveillance systems and behavioural analysis raises a number of questions regarding the protection of privacy, human rights and the limits of legitimate surveillance (Zubof, 2021). If AI is perceived as an instrument of mass surveillance, problems may arise in the form of resistance and loss of trust in institutions responsible for cybersecurity (Glišin, Milašinović & Lečić, 2025). Therefore, it is necessary to establish clear normative and institutional frameworks in order to avoid the collapse of the preventive and defensive function of AI systems in cyberspace (Novičić, 2021). In addition, the educational aspect is indispensable, since the effective application of AI systems in cyberspace depends on the level of expertise of personnel and decision-makers. Also, the human factor is the most vulnerable link in cybersecurity, which is why it is necessary to develop a general level of digital and security literacy (Furnell & Clarke, 2012; Despotović & Glišin, 2024, pp. 95-101). Therefore, the preventive role of AI in cybersecurity depends not only on technological development, but also on the ability of society to integrate technology into institutional, normative and value frameworks. We have seen that a systematic approach to modern security challenges is necessary in order to adequately respond to the challenges and threats in cyberspace.

CONCLUSION

Current international trends and the modern security environment have necessitated the transformation and adaptation of approaches to cyberspace and cybersecurity. In this context, artificial intelligence stands out as one of the key factors enabling the transition from a reactive to a proactive model of prevention and protection of systems in cyberspace. Cyberspace is a new strategic environment characterised by asymmetry, dynamism and global reach, which complicates the identification of threats and risks. Therefore, as we have previously pointed out, the capabilities of AI to collect and process large amounts of data, recognise complex behavioural patterns, perform predictive analytics and enable automated response in real time represent significant potential for improving cybersecurity.

Theoretical consideration of this topic is important, since by analysing the concepts and theoretical frameworks of AI, cyberspace and cybersecurity, we have shown that these are mutually conditioned and dynamic phenomena, the understanding of which is a necessary prerequisite for effective application in practice. By this, we mean primarily an educational character, since we believe that individuals with basic knowledge of the new security environment will have a more responsible approach to cyberspace and networks. In this regard, prevention in cyberspace, in addition to the technological aspect, inevitably has institutional and social aspects, which imply adequate institutional and normative frameworks, the development of educational capacities and social trust in AI-based systems. Lack of digital and security literacy, as well as insufficient understanding of the ethical and legal implications of the application of AI, can reduce the effects of technological solutions and open up space for their misuse.

Finally, we believe that adequate application of AI systems, within an integrated approach that connects technological, institutional and social factors, can contribute to improving security in cyberspace. Therefore, the synchronised development of AI technology, institutional capacities and social awareness can establish sustainable and effective systems of prevention and protection of states and societies in the modern security environment.

REFERENCES

Abomhara, M., & Koien, G. (2015). Cyber Security and the Internet of Things: Vulnerabilities, Threats, Intruders and Attacks. *Journal of Cyber Security*, 4, 65–88. DOI: 10.13052/jcsm2245-1439.414.

Bekker, A. (2019, May 13). 5 Types of AI to Propel Your Business. *Science Soft*. Available online at: https://www.scnsoft.com/blog/artificial-intelligence-types.

Beriša, H., & Pavić, A. (2024). Veštačka inteligencija i razvoj strategije nacionalne bezbednosti – izazovi i perspektive [Artificial Intelligence and National Security Strategy Development: Challenges and Perspectives]. *Nacionalni interes*, 48(2), 53-75.

Bostrom, N., & Yudkowsky, E. (2011). The ethics of artificial intelligence. In: Ramsey, W. & Frankish, K. (eds.): *Cambridge Handbook of Artificial Intelligence* (316-334). Cambridge University Press.

Buczak, A., & Guven, E. (2016). A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. *IEEE Communications Surveys & Tutorials*, 18 (2), 1153-1176. DOI: 10.1109/COMST.2015.2494502.

Chappell, J. H. (2020). *Artificial Intelligence: from Predictive to Prescriptive and Beyond*. Aveva.

Crootof, R., Kaminski, M., & Price II, N. (2023). Humans in the Loop. *Vanderbilt Law Review*, 76(2), 429-510.

Despotović, Lj., & Glišin, V. (2024). *Geopolitika i kriptopolitika* [Geopolitics and Cryptopolitics]. Beograd: Catena Mundi.

Despotović, Lj., & Glišin, V. (2025). *Savremeni međunarodni odnosi i geopolitika* [Contemporary International Relations and Geopolitics]. Beograd: Institut za političke studije.

Dugin, A., & Savin, L. (2018). *Mrežni ratovi* [Cyber Wars]. Beograd: Avala Pres.

Edris, E. K. K. (2025). Utilisation of Artificial Intelligence and Cybersecurity Capabilities: A Symbiotic Relationship for Enhanced Security and Applicability. *Electronics*, 14(10), 2057. DOI: https://doi.org/10.3390/electronics14102057.

Jevtović, A., Marić, K., & Milašinović, M. (2025). Propaganda and censorship on social networks in the age of political communication and digital diplomacy. *Diplomatija i bezbednost*, 8(1), 259-275.

Furnell, S., & Clarke, N. (2012). Power to the people? The evolving recognition of human aspects of security. *Computers & Security*, 31(8), 983-988. DOI: https://doi.org/10.1016/j.cose.2012.08.004.

Gibson, W. (2017). *Neuromancer*. London: Gollancz.

Glišin, V. (2025). Artificial intelligence in the security strategies of the USA and China at the beginning of the 21st century: The struggle for global dominance. *Srpska politička misao*, 94(6), 67–90. DOI: 10.5937/spm94-59859.

Glišin, V., & Milašinović, S. (2025). Artificial Intelligence and Contemporary Conflicts – The Example of the Israeli-Palestinian Conflict. *Zbornik Matice srpske za društvene nauke*, 76(3), 359-375. DOI: 10.2298/ZMSDN2595359G.

Glišin, V., & Petrović, J. (2025). The Use of Unmanned Aerial Vehicles for Military Purposes — A Global Perspective on the Development of Technology. *The Review of International Affairs*, 76, 399-420. DOI: 10.18485/iipe_ria.2025.76.1195.2.

Glišin, V., Milašinović, S., & Lečić, B. (2025). Artificial Intelligence as a Population Surveillance Tool: the Case of China. *Zbornik radova Filozofskog fakulteta u Prištini*, 55(4), 375-391. DOI: 10.5937/zrffp55-62227.

Haugeland, J. (1985). *Artificial Intelligence: The Very Idea*. Cambridge: MIT Press.

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

Hozouri, A., Mirzaei, A. & Effatparvar, M. (2025). A comprehensive survey on intrusion detection systems with advances in machine learning, deep learning and emerging cybersecurity challenges. *Discover Artificial Intelligence*, 5(314). DOI: https://doi.org/10.1007/s44163-025-00578-1.

Kaplan, A., & Haenlein, M. (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.

Kaur, R., Gabrijelčić, D., & Klobučar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion*, 97, 101804. https://doi.org/10.1016/j.inffus.2023.101804.

Kuehl, D. (2009). From Cyberspace to Cyberpower: Defining the Problem. In: Kramer, F., Starr, S., & Wentz, L. (eds.): *Cyberpower and National Security* (24-42). Lincoln: University of Nebraska Press.

Martínez Torres, J., Iglesias Comesana, C., & García-Nieto, P. J. (2019). Machine learning techniques applied to cybersecurity. *International Journal of Machine Learning and Cybernetics*, 10(10), 2823-2836.

Matković, A., Novakov, I., & Glišin, V. (2023). Prevalence and manifestations of digital harassment among young adults in Serbia. *Studia Securitatis*, 17(1), 88-103.

Milojević, S., Milašinović, S., & Milojković, B. (2025). Police and science: towards the development of academic discipline in the service of public safety. *Diplomatija i bezbednost*, 8(1), 7-40.

McCarthy, J. (2007). *What Is Artificial Intelligence?*. Stanford University. Available online at: http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html.

McKendrick, K. (2019). *Artificial Intelligence Prediction and Counterterrorism*. Chatham House.

Naj, Dž. (2012). *Budućnost moći* [The Future of Power]. Beograd: Arhipelag.

National Security Presidential Directives [NSPD]. (2008). *NSPD-54 - Cybersecurity Policy*. Washington: White House.

NIST. (2025). The CSF 1.1 Five Functions. *NIST*. Available online at: https://www.nist.gov/cyberframework/getting-started/online-learning/five-functions.

Novičić, Ž. (2021). Nova strategija sajber bezbednosti EU za digitalnu deceniju. In: Stanković, N., Dabić, D., & Bandov, G. (eds.): *Razvojni pravci Evropske unije nakon pandemije KOVID 19*. Beograd: IMPP. https://doi.org/10.18485/iipe_postkovid.2021.ch6.

Petrović, J., & Glišin, V. (2025). Use of Small Unmanned Aircraft Vehicles (SUAV) in Police Work. International Journal of Contemporary Security Studies, 1(2), 255-268. DOI: https://doi.org/10.18485/.

Prlja, D., Gasmi, G., i Korać, V. (2021). *Veštačka inteligencija u pravnom sistemu EU* [Artificial intelligence in the EU legal system]. Beograd: Institut za uporedno pravo.

Russell, S. & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach*. Prentice-Hall.

Samonas, S., & Coss, D. (2014). The CIA strikes back: redefining, confidentiality, integrity and availability in security. *Journal of Information System Security*, 10(3), 21-45.

Sangwan, R., Badr Y., & Srinivasan, S. (2023). Cybersecurity for AI Systems: A Survey. *Journal of Cybersecurity and Privacy*, 3(2), 166-190. DOI: 10.3390/jcp3020010.

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

Schiliro, F. (2023). Towards a Contemporary Definition of Cybersecurity. *ArXiv*. DOI: https://doi.org/10.48550/arXiv.2302.02274.

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding Machine Learning: From Theory to Algorithms*. Cambridge: Cambridge University Press.

Sharp, W. (1999). *Cyberspace and the Use of Force*. Virginia: Aegis Research Corporation.

Smola, A., & Vishwanathan, S. V. N. (2008). *Introduction to Machine Learning*. Cambridge: Cambridge University Press.

Stevanović, M., & Đurđević, D. (2019). *Izazovi nacionalne bezbednosti danas* [National Security Challenges Today]. Beograd: BIA/Službeni glasnik.

Subotić, O. (2023). *Digitalni rubikon* [Digital Rubicon]. Beograd: Catena Mundi.

Thakkar, A., & Lohiya, R. (2020). A Review of the Advancement in Intrusion Detection Datasets. *Procedia Computer Science*, 167, 636-645. DOI: https://doi.org/10.1016/j.procs.2020.03.330.

Wang, W., & Siau, K. (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

White House. (2003). *The National Strategy to Secure Cyberspace*. Washington, DC: The White House.

Wiafe, I., Koranteng, F., Obeng, E., Assyne, N., Wiafe, A., & Gulliver, S. (2020). Artificial Intelligence for Cybersecurity: A Systematic Mapping of Literature. *IEEE Access*, 8, 146598-146612.

Zhang, Z., Ning, H., Shi, F., Farha, F., Xu, Y., Xu, J., & Choo, K. K. R. (2022). Artificial intelligence in cyber security: research advances, challenges, and opportunities. *Artificial Intelligence Review*, 55(2), 1029–1053. DOI: https://doi.org/10.1007/s10462-021-09976-0.

Zubof, Š. (2021). *Doba nadzornog kapitalizma* [The Age of Surveillance Capitalism]. Beograd: Clio.

Živen, E. (2019). *Zapadni zid nije pao* [The Western Wall did not fall]. Novi Sad: Nova Evropa.

ВЕШТАЧКА ИНТЕЛИГЕНЦИЈА И САЈБЕР БЕЗБЕДНОСТ: ДРУШТВЕНИ И ТЕХНОЛОШКИ ИЗАЗОВИ

Вања Глишин[1], Срђан Милашиновић[2], Александар Алексић[3]

[1]Институт за политичке студије, Београд
[2]Криминалистичко-полицијски универзитет, Београд
[3]Факултет за дипломатију и безбедност, Београд

Резиме

Аутори у раду полазе од чињенице да се вештачка интелигенција намеће као један од кључних покретача дигиталне трансформације безбедности, поготову у области сајбер безбедности, те да је њена употреба у системима заштите неопходна како би се на адекватан начин одговорило на савремене изазове и претње. ВИ системи имају способност да обрађују велике количине података, детектују, анализирају и процењују понашање појединаца и потенцијалне последице тог понашања, врше предиктивну аналитику и нуде одговоре у различитим околностима и сценаријима. Захваљујући томе, могуће је променити приступ безбедносним ризицима са реактивног на проактивни модел, што би допринело превентивним деловањима. Дакле, ВИ због своје брзине, сложености и прилагодљивости може да се користи као превентивни инструмент у контексту савремених безбедносних изазова и претњи. Пажња се неизоставно посвећује сајбер простору, који осим функције у контексту развоја информационо-комуникационих система, постаје централно место нових безбедносних изазова и ризика. Према томе, аутори у раду анализирају улогу и капацитете ВИ у унапређењу сајбер безбедности, са посебним освртом на друштвене и технолошке изазове превентивне примене. Прво поглавље посвећено је појмовном и теоријском одређењу ВИ, сајбер простора и сајбер безбедности, као и логичком повезивању наведених појава и процеса. Друго поглавље посвећено је улози ВИ у превенцији сајбер претњи са циљем да се презентују предности употребе нових технологија у превенцији, детек-

ции, заштити, реаговању и одржавању безбедности система и његових компоненти. Коришћењем дескриптивне и експланаторне методе анализе и методе синтезе, циљ је да се презентује да је превенција у сајбер простору сложен технолошки, али и друштвени процес, који захтева интегрисани приступ.