e-ISSN **3041-2498**

Public Management and Policy

https://www.eu-scientists.com/index.php/pmap

National Security in the Age of Artificial Intelligence: New Forms of Disinformation and Risks for Political Systems

Andriy Krap [1] *

[1] *Lviv Institute of Higher Education "Interregional Academy of Personnel Management" (Ukraine). Associate Professor of the Department of Law, PhD in Political Sciences.*

*** *Corresponding Author,*** e-mail: a.krap1412@gmail.com

ARTICLE INFO

Research Article

DOI: 10.70651/3041-2498/2025.12.09

Received:
2 November 2025
Accepted:
3 December 2025
Published online:
16 December 2025

ABSTRACT

The purpose of the article is to outline the threat vectors that AI generates for national security and political systems, as well as to justify the need to develop new mechanisms of legal regulation and accountability in the context of technological dominance. To achieve the goal, several tasks were solved, including an analysis of the impact of the phenomenon of the double "black box" on the mechanisms of constitutional deterrence; a review of new forms of disinformation; an outline of the risks associated with the use of AI for military purposes and internal control; a justification of the need to prevent the transition to technocratic governance. In preparing the article, general scientific and special methods were used. A systems method was used to study AI as an integrated element of the security architecture. Also, based on the available empirical material, attention was focused on possible scenarios of AI abuse by state and non-state actors of national security and politics. A case study was used to examine the impact of AI tools on the course of elections and the experience of digital transformation. The study finds that AI is significantly increasing the speed and scale of cyberattacks on critical infrastructure. This can lead to physical disruption and destabilization. AI has also become a weapon in information and cognitive warfare, enabling the creation of photo-realistic deepfakes and hyper-personalized narratives, which has a critical impact on public trust. It is emphasized that a significant risk is the transition from democratic accountability to technocratic governance, where decisions are made by opaque algorithms. This may ultimately contribute to the depreciation and neglect of fundamental democratic rights of citizens and societies. It also emphasizes the inherent vulnerabilities of AI systems (data poisoning, adversarial attacks, bias), which are used to compromise political opponents and form distrust of existing political systems and institutions. The practical significance of the article lies in substantiating the need for urgent implementation of regulatory measures aimed at ensuring transparency and human control over critical AI decisions. The experience of Estonia and the WINWIN memorandum emphasize the importance of international cooperation to strengthen cyber resilience. Prospects for further research should be aimed at developing universal protocols for auditing algorithms and creating effective mechanisms for international regulation that will ensure a clear balance between the innovative potential of AI and the need to prevent existential risks to human rights and national security. When introducing AI into various spheres of life, it is necessary to take into account the dynamism and rapidity of technological development, as well as the possibility of its unfair use.

KEYWORDS

war, disinformation, national security, politics, artificial intelligence (AI).

e-ISSN **3041-2498**

Публічне управління і політика

https://www.eu-scientists.com/index.php/pmap

Національна безпека у добу штучного інтелекту: нові форми дезінформації та ризики для політичних систем

Андрій П. Крап [1] *

[1] *Львівський інститут ПрАТ «ВНЗ «Міжрегіональна Академія управління персоналом» (Україна). Доцент кафедри права, канд. політ. наук.*

*** *Автор-кореспондент,*** e-mail: a.krap1412@gmail.com

СТАТТЯ

Дослідницька

DOI: 10.70651/3041-2498/2025.12.09

Отримана: 02.11.2025 р.
Прийнята: 03.12.2025 р.
Опублікована: 16.12.2025 р.

АНОТАЦІЯ

Метою статті є окреслення векторів загроз, які за допомогою ШІ генерують для національної безпеки та політичних систем, а також обґрунтування необхідності розробки нових механізмів правового регулювання та підзвітності в умовах технологічного домінування. Для досягнення мети було вирішено низку завдань, включаючи аналіз впливу феномену подвійної «чорної скриньки» на механізми конституційних стримувань; огляд нових форм дезінформації; окреслення ризиків, пов'язаних із використанням ШІ у військових цілях та внутрішньому контролі; обґрунтування необхідності запобігання переходу до технократичного управління. При підготовці статті були задіяні загальнонаукові та спеціальні методи. Для вивчення ШІ як інтегрованого елементу архітектури безпеки застосовано системний метод. Також, спираючись на наявний емпіричний матеріал, було зосереджено увагу на можоивих сценаріях зловживання ШІ з боку державних та недержавних акторів національної безпеки та політики. За допомогою кейс-стаді розглянуто вплив інструментарію ШІ на перебіг виборів та досвід цифрової трансформації. З результатів дослідження випливає, що ШІ значно посилює швидкість і масштаб кібератак на критичну інфраструктуру. Це може призводити до фізичних збоїв та дестабілізації. Також ШІ вже став зброєю у сфері інформаційної та когнітивної війни, уможливлюючи створення фото-реалістичних діпфейків та гіпер-персоналізованих наративів, що критично впливає на суспільну довіру. Наголошено, що суттєвим ризиком є перехід від демократичної підзвітності до технократичного управління, де рішення приймаються непрозорими алгоритмами. Це зрештою може сприяти знеціненню та нехтуванню фундаментальних демократичними правами громадян та суспільств. Також наголошено на внутрішніх вразливостях ШІ-систем (отруєння даних, змагальні атаки, упередженість), які використовують для компрометації політичних опонентів та формуванню недовіри до існуючих політичних систем і інституцій. Практичне значення статті полягає в обґрунтуванні необхідності термінового впровадження регуляторних заходів, спрямованих на забезпечення прозорості та людського контролю над критичними рішеннями ШІ. Досвід Естонії та меморандум WINWIN підкреслюють важливість міжнародної співпраці для зміцнення кіберстійкості. Перспективи подальших досліджень мають бути спрямовані на розробку універсальних протоколів для аудиту алгоритмів та створення ефективних механізмів міжнародного регулювання, які забезпечать чіткий баланс між інноваційним потенціалом ШІ та необхідністю запобігання екзистенційним ризикам для прав людини і національної безпеки. При впровадженні ШІ у різні сфери життя необхідно враховувати динамічність і стрімкість розвитку технологій, а також можливості його недобросовісного використання.

КЛЮЧОВІ СЛОВА

війна, дезінформація, національна безпека, політика, штучний інтелект (ШІ).

1. Introduction

The modern national security architecture is undergoing fundamental changes under the influence of the technological dominance of artificial intelligence (AI). Despite the obvious advantages, the introduction of AI in highly sensitive areas of national security generates several serious risks. In particular, the computational opacity of decisions made by AI systems (the so-called "black box" effect) calls into question the possibility of establishing clear cause-and-effect relationships and an effective accountability mechanism, which is critical to prevent uncontrolled escalation. The risk is compounded by the fact that national security decisions are traditionally among the least transparent and accountable government processes. As Ashley Dix rightly points out [1], the superimposition of an AI "black box" on an already existing "black box" of the government's decision-making process creates a double "black box" problem. This means that the mechanisms of constitutional checks and balances designed to control the actions of the executive branch become practically ineffective. Agencies, including intelligence and defense agencies, are already deploying AI to identify threatening activity, cyber operations, and potential targets. Thus, decisions with critical consequences that can lead to armed conflict or escalation can be made algorithmically – without a mandatory affirmative human decision. As a result, even security decision-makers cannot fully understand exactly how the final recommendation was formed. This aspect of opacity is vital, as it creates preconditions for the abuse of AI not only in the context of military operations, but also in the field of information manipulation [1].

The rapid development of AI technologies is outpacing the adaptability of national legal systems, which creates serious gaps in their regulation, especially in the field of national security. The challenges are to guarantee human rights, hold autonomous systems accountable, and ensure transparency and accountability of AI. The situation is complicated by the lack of agreed international standards, which increases the risk of abuse [2, p. 122].

2. Literature Review

The study of the practical and theoretical application of AI is one of the most active areas of modern scientific thought, both in Ukraine and abroad. Scientists focus on analyzing the implementation of AI systems in various industries and on the methodological justification of this process. Among the problems under study, attention is focused on the clarification of the conceptual apparatus and the need to increase the professional competence of users and employees for the effective use of new technologies [2–5]. Also, a number of publications are devoted to the study of the potential of AI in various spheres of human life and society and the state, but at the same time try to outline, analyze and predict the challenges and risks associated with its use [1; 6–8]. A number of studies deal directly with AI and national security, both at the state and geopolitical levels [9; 10].

However, taking into account the dynamics of the development of AI and its implementation in various spheres of life, in particular, in the field of national security, as well as the possibility of using AI to inflict multi-vector damage and losses, the study of this technology will not lose its scientific relevance for a long time.

3. Problem Statement

The outlined issues require a comprehensive study that goes beyond the traditional threat analysis. Thus, the article is aimed at identifying the existing vectors of threats that AI creates for national security and political systems, as well as substantiating the need to develop new mechanisms of legal regulation and accountability in the context of technological dominance and the emergence of various state and non-state actors in the geopolitical field. To achieve this goal, it is necessary to visualize new forms of disinformation and manipulation arising from the use of AI tools and assess their impact on electoral processes in democratic and authoritarian political systems. It is also important to outline the risks associated with the use of AI for military purposes and in the field of internal state control.

4. Methods and Materials

The methodological basis of the article was system and comparative analysis for the study of AI as an integrated component of modern national security. With the help of a systematic approach in the analysis of AI, the relationships between technological transformations and political risks were considered. The comparative approach was applied to study the experience of integrating AI into political and legal systems of a state and a supranational nature. The use of AI to manipulate public opinion and politically discredit the conduct of hostilities or political elections, which allows us to make assumptions about the speed and effectiveness and dynamics of threats of unfair and malicious use of AI. The case study came in handy when studying practical examples of the use of AI in the field of national security. The synthesis method made it possible to formulate conclusions and outline the problems of further development. The materials for the article were the developments of scientists, as well as analytical documents and publications of AI practitioners.

5. Results and Discussion

The debate on the definition of artificial intelligence (AI) has a long history, dating back to the middle of the XX century. Then the pioneers of the concept, in particular J. McCarthy and M. L. Minsky, viewed AI as the ability of machines to perform tasks that require human intellectual skills, such as logical thinking and learning. Over the decades, interpretation has been transformed under the influence of technological advances. Modern concepts, in particular those adopted by universal and regional organizations, have moved away from a narrow interpretation (imitation of human thinking) to a more inclusive one. AI is now seen as a system that has a comprehensive goal and is able to act autonomously, interact with the environment, and make optimal decisions based on the collection, analysis, and interpretation of large amounts of data. At the same time, there is an ongoing scientific and ethical debate about the limits and possibilities of AI. More and more researchers insist that the definition of AI cannot be exclusively technical and requires mandatory consideration of its social, legal and ethical aspects. This is necessary to avoid errors in the forecasting and regulation of technologies that have a fundamental impact on all spheres of life [2, p. 123]

The rapid development of AI is fundamentally changing the rules of the game in the political arena. AI technology not only improves old political tools, but also opens up fundamentally new ways for both political strategies and, unfortunately, manipulation. Today, we can say that AI has become a powerful analytical center. Politicians can now study electoral sentiments with exceptional accuracy, make detailed predictions of voting results, and almost completely automate communication with their voters. Thanks to the ability of AI to instantly process colossal amounts of data, it becomes possible to detect those hidden patterns in the behavior of citizens that were previously inaccessible to human analysis. This, in turn, allows you to influence the political views of voters much more effectively than ever. Essentially, we are seeing a shift towards hyper-personalized political advertising and the creation of uniquely tailored narratives. It's no longer just informing; it is the formation of a new perception of political reality. One of the key tools for political manipulation using AI is content personalization, especially through social media. Algorithms today do an incredible job: they analyze every user action, their interests, search history, and preferences. Based on this detailed profile, the AI creates advertising and informational messages that will resonate with the voter's individual beliefs as much as possible. The effectiveness of such targeted communication is growing exponentially, as it goes beyond rational information. It creates emotional bonds and, critically, deepens existing divisions in society. This leads to an increased risk of social polarization [6, pp. 131-132].

As a result of these processes, AI makes it possible not only to fine-tune electoral preferences but also to generate completely new forms of political manipulation. These include automated mass content creation (e.g., comments, posts, bots), as well as news falsification (Deepfakes, fake testimonies, etc.), which poses a direct threat to the integrity of political systems. For example, let's take a look at the recommendations of the European Citizens' Action Service (ECAS), which were developed as part of the popAI project. It is worth noting that the popAI project is a two-year initiative funded by the Horizon program, which brought together 13 organizations from 8 European member states. The main mission of the project is to strengthen citizens' trust in the use of AI by law enforcement agencies. Key recommendations were developed to ensure the responsible and effective implementation of AI in the field of security, focusing on three main areas:

- Support for law enforcement agencies. These are guidelines that are designed to help ethically and efficiently use AI tools in the performance of operational tasks. At the same time, law enforcement officers must be competent in terms of legal regulations on data protection, confidentiality and compliance with the principles of non-discrimination. It is also necessary to take into account the constant updates of AI capabilities. Therefore, such employees should always be aware of these updates through continuous training. to personal data.
- Recommendations for legislators and politicians. The purpose of these recommendations is to take them into account in the formation of future legislation related to the scope of AI in the field of security. In particular, special attention at the legislative level should be focused on the storage and use of biometric data. Also, the rules for storing such data should be synchronized at the level of the European Union. having signs of discrimination (ethnic origin, gender, etc.). An equally significant problem is the creation of independent supervisory bodies. which should monitor the use of such information and develop predictive control algorithms. Such measures are necessary to avoid the discriminatory use of AI. In addition, at the legislative level, it should be mandatory to inform citizens about when their data is used. The introduction of mechanisms that allow you to withdraw consent and request the deletion of personal data is also on the agenda.
- Recommendations for AI technology developers. They are based on ethical principles throughout the development process of AI tools. First of all, it is about the principle of non-discriminatory algorithms, which should run like a red thread through the entire cycle from design to the final product. AI tools are aimed at supporting humans, not replacing them. AI algorithms should be built on the diversity and representativeness of the data set that reflects the exact state of the population. results [11].

It is also worth dwelling on the example of Estonia as a country that, having regained its independence in 1991 after the Soviet occupation, inherited limited and outdated technological systems [12]. However, the country has turned this "zero" into a strategic advantage for rapid digital transformation and European integration. Decisive steps were taken almost instantly. In 1994, the strategic plan of the "Principles of Information Policy of Estonia" was developed and ratified, which provided for the allocation of 1% of GDP for IT financing. The peak point was the nationwide program "Tiger Leap" (1996), which aimed to update the infrastructure to European standards and make computer skills training a school priority. Such an ambitious policy led to a truly impressive result: Estonia, despite the difficult historical and political context, took a leading position in the Digital Development Index, and 99% of its citizens are regular Internet users.

In 1996, in parallel with the nationwide digitalization program, private banking structures began to provide the first electronic banking services. This was done for customer access in rural communities, which was a catalyst for encouraging citizens to use the internet and electronic identification. At the beginning of the XXI century. Estonia has introduced electronic meetings of the Cabinet of Ministers, creating a specialized database to optimize government decisions. This significantly reduced bureaucracy, reducing the average duration of meetings from 5 hours to 30 minutes. In 2000, the electronic tax board was launched, which allowed 99% of Estonians to declare taxes online, maximizing tax revenues. In the same year, m-Parking was introduced for mobile parking payments, which is used today in 95% of cases. The following year, in 2001, the X-Road integration platform was created, which became the basis of the e-state, combining public and private information systems. Thanks to X-Road, 100% of public services have become available online around the clock. In 2002, a mandatory identity card with the function of electronic identification and digital signature was introduced, which today covers 99% of citizens and annually saves 2% of GDP. The final stage was the introduction of i-Voting since 2005, which allowed local and national elections to be held online, in particular from more than 110 countries around the world; Currently, a third of votes are cast through this system [13; 14].

The high level of digitalization exacerbated the problem of cybersecurity, which was confirmed by a high-profile external cyberattack in April 2007. In response to cyberattacks, in 2008, Estonian cryptographers developed the scalable KSI blockchain technology, the purpose of which was to mitigate the threat of manipulation of insider data in state registers. In the same year, a nationwide e-health system began to operate, which integrates data from all medical institutions. Since 2013, Estonia has implemented mapping of problems and solutions in the development of public e-services. This approach provided a better understanding of the needs of the public and helped to address current shortcomings, guaranteeing the sustainability and further development of the e-state [16].

This is just a cursory overview of the path that the small Baltic country has traveled in a relatively short time, but with a great understanding of the importance of technological autonomy at the state level. In particular, the government's AI strategy, which outlines the current and future use of AI in public and private services, has been approved in Estonia since 2019. Its continuation was the National AI Strategy of Estonia (Kratt Strategy). It is a multi-vector initiative aimed at improving the efficiency of the public sector, supporting private businesses in the adoption of AI, promoting research and education, and developing a regulatory environment that ensures the responsible use of technology. Estonia sees AI as a powerful and effective tool for improving public administration, strengthening economic competitiveness and strengthening digital infrastructure, making it a core factor in the implementation of the national digital agenda. Kratt's strategy is structured around five key goals: expanding the adoption of AI in the public sector to develop digital services that increase accessibility and efficiency of the user experience; supporting the introduction of AI in private enterprises in order to integrate technologies into their activities to increase competitiveness; developing AI research and education through strengthening academic programs, research funding, and skilled workforce training; ensuring the availability of high-quality data by promoting open data policy and ethical management; development of a legal and regulatory framework for AI that harmonizes national legislation with EU regulations [17].

Against the backdrop of world leaders in digital transformation, in particular the example of Estonia, Ukraine is actively using international cooperation to accelerate its own technological development. Trilateral Memorandum of Cooperation in the Field of Digitalization. The document was concluded between the Ministry of Digital Transformation of Ukraine, on the one hand, and the Ministry of Foreign Affairs, Commonwealth Affairs and Development (together with the Department of Science, Innovation and Technology) of the United Kingdom of Great Britain and Northern Ireland and the Ministry of Justice and Digital Affairs of the Republic of Estonia, on the other. Representatives of Ukraine, Estonia and the United Kingdom emphasized the strategic importance of partnership. According to the Minister of Digital Transformation of Ukraine, cooperation with Estonia as a mentor and the United Kingdom as one of the first partners in the field of innovation is an important step forward and once again confirms the common vision of prospects aimed at creating breakthrough solutions and developing AI in public services to increase their convenience and efficiency. The main areas of joint work include: development of digital governance technologies and public services; exchange of experience in the field of AI, open data and cybersecurity; support for startups and innovative companies; Joint educational programs in digital literacy. The importance of such cooperation is also obvious for national security, as it strengthens cyber resilience and allows countering hybrid threats. However, we cannot ignore the technological threats that require partners to work together to develop not only innovative but also ethically regulated solutions. The signing of the memorandum is part of the implementation of the national WINWIN Strategy, which provides for strengthening the country's technological potential through international partnerships and Ukraine's integration into the global innovation space [18].

However, despite the not yet fully realized opportunities that AI technology and tools provide to humanity, we already have several serious challenges and threats to national security, which are carried by the uncontrolled or unfair use of such capabilities. Analogies with the possibilities of nuclear energy, examples of the use of which are not for the benefit of humanity, but for its destruction and intimidation, we know from not-so-distant history. And nuclear intimidation by Russia of the civilized world also does not lose its relevance. As M. Branca noted, currently nine nuclear powers are not only preserving, but also actively modernizing their arsenals, ignoring calls for elimination. The United States and Russia, which own about 87% of the world's nuclear warheads, have slowed down the pace of arms reductions. China is also already openly demonstrating the growth of both rhetoric and capabilities of its nuclear potential. that Saudi Arabia enjoys Pakistan's de facto nuclear protection, and some analysts even suggest that medium-sized powers such as Canada and Germany, which have previously been in favor of disarmament, should consider obtaining their own nuclear weapons. The situation is complicated by the integration of AI and other emerging technologies that radically transform decision-making processes and military command and control structures. Algorithmization promises to reduce reaction times, but this critically increases the risk of errors and unpredictable escalation. Nuclear danger of the XXI century. not a relic of the Cold War, but a more complex, less predictable and potentially more catastrophic threat [19].

With the help of AI technologies, there are interferences in electoral processes in different countries of the world, as well as attempts to undermine the situation in the country by spreading various disinformation and influencing public moods and political views. One of the first large-format testing grounds for the testing and application of the latest technologies was the war in Syria, which began in 2011. and the Islamic State of Iraq and the Levant (ISIL). The Russians' operations were aimed at discrediting their opponents ("white helmets" or Syrian Civil Defense); promoting the narrative of fighting exclusively against "terrorists"; denying the use of chemical weapons by the Assad regime or blaming the opposition. In turn, ISIS effectively used social networks to demonstrate strength, recruit and create the image of a caliphate. Bot networks and troll factories, there was a massive spread of fake news, fabricated testimonies, and propaganda narratives. In addition, with the help of discrediting campaigns, targeted campaigns were launched to denigrate the reputation of opponents, spread rumors and false accusations. It was also used to create the illusion of mass support or popular anger through a large number of fake accounts imitating real users. Manipulating trends, bots were used to artificially display certain hashtags or topics in the top discussions of social networks. There was also a suppression of dissent by conducting attacks and mass complaints against the accounts of activists and journalists in order to block them. In general, Russia's modern policy in the context of military doctrine is based on the use of the information space as a battlefield. Information operations conducted by the Russian Federation are a direct adaptation of Soviet active actions to the digital age. but as a powerful catalyst. It can significantly increase the efficiency of existing methods of information warfare by automating and scaling operations; improving targeting of target audiences; creation of realistic synthetic content; integration with cyber operations [4].

So, as we can see in various examples of the use of AI in the field of national security, the risks associated with it are multi-vector and specific, which requires their systematic understanding. The complexity and adaptability of AI systems make their behavior less predictable. Threats come from several sources. First of all, we are talking about data bias. which has significant ethical and social implications. There is also a risk of competitive attacks, when attackers create specially modified inputs to bypass the model's ethical filters and generate harmful, disinformational, or malicious content (e.g., phishing emails). Data poisoning and backdoor attacks are also becoming widespread, which occur both at the stage of preliminary training and during additional training of the model. These methods allow attackers to implement stealthy malicious behavior that is only activated by certain triggers, which is especially dangerous in specialized security applications. In addition, there is a risk of model hijacking when attackers replicate its functionality by sending requests in bulk, actually obtaining the intellectual property and capabilities of the system. These attack vectors require the immediate implementation of transparency and cyber protection mechanisms at all stages of the AI lifecycle [20].

A significant risk is the use of AI to automate cyberattacks, which significantly increases their speed and scale. If AI controls critical infrastructure systems, the consequences of an attack can have a far-reaching, destabilizing effect. Over the past decades, cyber threats targeting critical infrastructure have evolved significantly. If earlier attacks were limited to simple failures or denial of service, then with the increase in the interconnectedness of networks and dependence on industrial control systems, their scale and destructive impact have increased significantly. Today, the goals of cyberattacks have shifted from simple data theft to causing large-scale, physical disruption and destabilization, which is a direct challenge to national security. Critical infrastructure (energy, transport, finance, healthcare) has become an attractive target for cybercriminals and state-sponsored entities. Such attacks not only lead to financial losses but can also paralyze vital services, as was the case during the attack on the Colonial Pipeline in the United States (2021), which caused fuel shortages and chain failures [21, p. 6].

With the help of AI technology, they create qualitatively new threats to national security that go far beyond cyberattacks on physical infrastructure. As noted in 2025 during the World Economic Forum's work in analyzing the adverse effects of AI technologies, there is a high risk of undermining geopolitical stability and public health. The most acute threat concerns political elections and democratic processes. Generative AI has dramatically increased the opportunities for automated disinformation and cognitive warfare. Thanks to AI, creating highly credible fake content and fake social profiles has become a matter of a few clicks. These campaigns can target specific demographics by using personalized language and narratives. Similar threats have already materialized, as evidenced by an incident during the 2024 Romanian presidential election, which was suspended due to interference through social platforms. This confirms that the risks are not only theoretical but also have a direct destabilizing effect on political systems. In the medium term, as AI improves, the intensity and

sophistication of such threats will inevitably increase. No less threatening are the financial and economic consequences of using AI, which can be a tool of economic destabilization. AI-generated fake news is being turned into a weapon to manipulate stock markets. For example, in 2023, an AI-generated image of an explosion near the Pentagon led to a drop in the value of stocks in US markets, demonstrating how false information exposure can cause real financial losses and panic [22].

Today, AI is already considered a transformational technology of recent decades, comparable in importance to the emergence of the Internet [3]. It has the potential to radically reformat economic and socio-political systems. However, as with any powerful innovation, AI's potential for abuse increases exponentially when it falls into the hands of governments. History has repeatedly confirmed: when states acquire new instruments of control, they rarely abandon them. The main risk is that governments could use AI technologies to give themselves special powers and exempt themselves from legal regulations that apply to citizens and private companies, and in an "emergency", even suspend basic individual freedoms. With public funding and access to sophisticated models, governments can create deepfakes that are indistinguishable from reality to fabricate evidence, discredit journalists and opponents. This allows the creation of totally fictitious threats to justify internal or international abuse of authority or even military conflicts. The prospect of introducing "social credit" systems following the example of China is particularly alarming. AI can combine financial information, social media, political preferences, and opinions expressed into a detailed scoring system. Such a system can automatically distribute rewards and punishments, limiting centralized access to vital services (banking, housing, healthcare), which becomes a powerful tool for coercion and suppression of dissent. Politicians can invoke AI, claiming that the "algorithm" is neutral, efficient, and devoid of human bias. However, all AI systems are trained on data that inevitably reflects the historical biases, political assumptions, and interests of those who created the system [8].

Thus, AI has radically transformed the information environment, creating qualitatively new threats to political systems, which is especially relevant in the context of mass election campaigns. Generative AI allows you to create photorealistic deepfakes and personalized narratives that are difficult to distinguish from real images and informational messages. This leads to a crisis of trust and a blurring of the perception of reality. The experience of the election campaigns of Argentina and Turkey in 2023 has shown that AI technologies are actively used to denigrate opponents and create political disinformation (for example, fabricated videos of support for terrorists or fictitious statements) [23]. Thus, the risk of introducing AI into public administration is also the shift of management responsibilities from democratic accountability to technocracy.

Thus, we can state that AI, despite significant benefits for productivity and innovation, is quite critical and risky for national security and political systems. As of 2025, AI is not so much creating new threats as dramatically increasing the speed, scale, and effectiveness of existing ones, allowing even less skilled actors to carry out previously unreachable attacks. The greatest influence of AI is concentrated in the digital sphere (cyberattacks, fraud, impersonation), as well as in the plane of political stability. The proliferation of synthetic media (deepfakes) directly threatens the erosion of democratic participation and public trust in state institutions. In addition, embedding AI into physical systems, including critical infrastructure, carries growing physical security risks. The global community's preparedness to mitigate these threats is uneven, and global regulation remains incomplete and likely does not keep pace with future technological developments, highlighting the need for the immediate implementation of effective countermeasures strategies.

6. Conclusions

The integration of AI into the functioning of state institutions and society is a multi-vector challenge for national security that goes far beyond the usual cyber threats and cyberattacks. First of all, we are talking about the problem of accountability and escalation, when the opacity of AI algorithms is superimposed on the limited accountability of the security sector. Such a situation can lead to the destruction of established and tested mechanisms of constitutional and democratic deterrence, which creates preconditions for the intensification of conflict situations and erroneous decisions with tactical and strategic consequences. We can conduct analogies with nuclear weapons and nuclear blackmail, which we have today as a significant factor in geopolitics.

The use of AI, in particular to create disinformation content, dramatically increases the scale and effectiveness of information and cognitive warfare, both nationally and internationally. This not only

affects the course and results of political processes, in particular, election campaigns, but also leads to the formation and strengthening of distrust in existing democratic institutions, social polarization and undermining of political systems, and contributes to the formation of authoritarian governance.

It is emphasized that the growing dependence on AI revealed its internal vulnerability. It's about bias and data poisoning, competitive attacks, and model hijacking. All this can be used to compromise political systems and manipulate the results of elections, referendums, etc.

It is emphasized that the experience of leading countries in digital transformation demonstrates that by minimizing the risks associated with the use of AI, there is a clear state strategy, attracting investments and their strategy, cyber resilience, data protection and international cooperation.

Further research should focus on developing a comprehensive regulatory framework through which human-centered control over AI decisions and decisions could be formed. At the same time, it is necessary to take into account the human factor of influencing AI, the use of technology not for the good, but to the detriment of humanity and individuals. Strengthening international partnerships for the exchange of experience and joint counteraction to threats to national security in the field of information and influence on political processes currently looks like a promising and far-sighted step.

References

1. Deeks, A. (August 14, 2024). The Double Black Box: AI Inside the National Security Ecosystem. *Just Secyritu*. https://www.justsecurity.org/98555/the-double-black-box-ai-inside-the-national-security-ecosystem/
2. Hrebeniuk, D. (2024). Vykorystannia shtuchnoho intelektu v systemi zabezpechennia natsionalnoi bezpeky: perspektyvy, vyklyky ta pravovi aspekty. [The use of artificial intelligence in the system of ensuring national security: prospects, challenges and legal aspects]. *Natsionalni interesy Ukrainy*, (5), 121–128. https://doi.org/10.52058/3041-1793-2024-5(5)-121-128 (in Ukrainian)
3. Zhorniak, A. (2025). Shtuchnyi intelekt yak innovatsiinyi instrument sotsialnykh zmin: vyklyky, perspektyvy, ryzyky ta sotsialni naslidky. [Artificial intelligence as an innovative tool for social change: challenges, prospects, risks and social consequences]. *Vidkryta nauka ta innovatsii,* (1), 57–75. http://doi.org/10.62405/osi.2025.01.05 (in Ukrainian)
4. Zubrytskyi, H. (2025). Zahrozy ta ryzyky shtuchnoho intelektu i neiromerezhevykh alhorytmiv dlia systemy sotsialnopolitychnykh komunikatsii u rosiisko-ukrainskii viini. [Threats and risks of artificial intelligence and neural network algorithms for the socio-political communications system in the Russian-Ukrainian war]. *Military Strategy and Technology*, (1), 69–75. https://doi.org/10.63978/3083-6476.2025.1.1.07 (in Ukrainian)
5. Sorokina, O. (2025) Vprovadzhennia shtuchnoho intelektu (ShI) v robotu orhanizatsii. [Introduction of artificial intelligence (AI) into the work of organizations]. *Humanities Studies, 22*(99), 132–138. https://doi.org/10.32782/hst-2025-22-99-14 (in Ukrainian)
6. Lazarovych, M., Honcharuk-Cholach, T., & Dokash, O. (2025). Shtuchnyi intelekt u politychnykh manipuliatsiiakh ta transformatsiia vyborchoi kultury. [Artificial Intelligence in Political Manipulations and the Transformation of Electoral Culture]. *Politykus*, (1), 131–135. https://doi.org/10.24195/2414-9616.2025-1.20 (in Ukrainian)
7. Ångström, R., Björn, M., Dahlander, L., Mähring, M., & Wallin, M. (2023). Getting AI Implementation Right: Insights from a Global Survey. *California Management Review, 66* (1), 5–22. https://doi.org/10.1177/00081256231190430
8. Roth, V. P. (2025, October 10). The many dangers of AI in state hands. *GIS.* https://www.gisreportsonline.com/r/dangers-ai-in-state-hands/
9. Mansur, M. A. (2025). National Security and Cyber Defense in the Rise of Artificial Super Intelligence. *European Scientific Journal, ESJ*, (39), 32–329. https://doi.org/10.19044/esj.2025.v21n10p116
10. Srivastava, K. (2023). Artificial intelligence and national security: Perspective of the global south. *International Journal of Law in Changing World*, (2), 77–87. https://doi.org/10.54934/ijlcw.v2i2.63
11. European Citizen Action Service (ECAS). (2024, August 19). *Recommendations towards a safe use of AI tools in the security domain*. https://ecas.org/publication/recommendations-towards-a-safe-use-of-ai-tools-in-the-security-domain
12. Raun, T. U. (2009). Estonia after 1991: Identity and Integration. *East European Politics and Societies and Cultures, 23*(4), 526–534. https://doi.org/10.1177/0888325409342113
13. Horowitz, B. & Tamkivi, S. (2014, February 6). Estonia: The Little Country That Cloud. Andreessen Horowitz. *Andreessen Horowitz*. https://a16z.com/estonia-the-little-country-that-cloud/

14. e-Estonia Briefing Centre. (n.d.). *e-Estonia: We have built a digital society & we can show you how* [Homepage]. https://e-estonia.com
15. Dilek, E., Talin, Ö. & Bensghir, T. K. (2023). An Examination of Estonia 2007 Cyber Attacks and the Effects on National Cyber Security Policies of Countries. *Bilgi Yönetimi, 6*(2), 332–347. https://doi.org/10.33721/by.1392577
16. Estonia leads the way with advanced e-services for citizens. (2016, January 12). *European Commission.* https://ec.europa.eu/regional_policy/en/projects/estonia/estonia-leads-the-way-with-advanced-e-services-for-citizens
17. Ministry of Economic Affairs and Communications. (2021). Estonia's national artificial intelligence strategy or Kratt Strategy for 2022–2023. *Digwatch*. https://dig.watch/resource/estonias-national-artificial-intelligence-strategy-kratt-strategy-for-2022-2023
18. Uriadovyi portal. (2025, November 6). Ukraina, Velyka Brytaniia ta Estoniia rozvyvatymut spivpratsiu u sferi tsyfrovykh tekhnolohii ta innovatsii [Ukraine, the United Kingdom and Estonia will develop cooperation in the field of digital technologies and innovations]. *Yedynyi veb-portal orhaniv vykonavchoi vlady Ukrainy*. https://www.kmu.gov.ua/news/ukraina-velyka-brytaniia-ta-estoniia-rozvyvatymut-spivpratsiu-u-sferi-tsyfrovykh-tekhnolohii-ta-innovatsii (in Ukrainian)
19. Branka, M. (2025, November 7). Opinion: The nuclear illusion: Mistaking luck for security. *Digital Journal*. https://www.digitaljournal.com/world/op-ed-the-nuclear-illusion-mistaking-luck-for-security/article
20. Narula, S., Ghasemigol, M., Carnerero-Cano, J., Minnich, A., Lupu, E., & Takabi, D. (2025). Exploring AI security: A systematic mapping study. *IEEE Access*. Advance online publication. https://www.researchgate.net/publication/391477701_Exploring_AI_Security_A_Systematic_Mapping_Study
21. Khan, M. (2025). Cybersecurity in Critical Infrastructure: Challenges and Future Directions. *Researchgate.* https://www.researchgate.net/profile/Muskan-Khan-5
22. Insua, D. R. (2025, September 29). AI poses risks to national security, elections and healthcare. Here's how to reduce them. *The Conversation. Academic rigour, journalistic flair.* https://doi.org/10.64628/AAK.vy39veg7e
23. Polishchuk, A. (2024, January 26). AI Poses Risks to Both Authoritarian and Democratic Politics. *The Russia File. A blog of the Kennan Institute*. https://www.wilsoncenter.org/blog-post/ai-poses-risks-both-authoritarian-and-democratic-politics