UDC 316.4:004.8:355.02 DOI: 10.63978/3083-6476.2025.1.1.07

Hryhorii Zubrytskyi
Philosophy Doctor in Technical Sciences,
Associate Professor
Military Unit A1915
Chortkiv
e-mail: red.hnups@gmail.com
ORCID: 0000-0003-0481-9454

THE ARTIFICIAL INTELLIGENCE AND NEURAL NETWORK ALGORITHMS' THREATS AND RISKS FOR THE SOCIO-POLITICAL COMMUNICATIONS SYSTEM IN THE RUSSIAN-UKRAINIAN WAR

Abstract. *This article provides a comprehensive analysis of the threats and risks arising within the system of socio-political communications due to the application of artificial intelligence (AI) and neural network algorithms in the context of the ongoing russian-Ukrainian War. It emphasizes that modern AI technologies, including machine learning, natural language processing, and especially generative AI, have transformed into powerful tools of information warfare. These technologies enable the aggressor to automate, scale, and significantly enhance the effectiveness of propaganda, disinformation, and psychological manipulation targeting Ukrainian society and the international community. The article aims to thoroughly investigate and systematize these threats by analyzing specific methods of AI application by the russian federation within its hybrid warfare strategy against Ukraine. This analysis considers the experience gained from previous conflicts, particularly the war in Syria, which served as a testing ground for Russia, Iran, and ISIS to trial various information warfare tactics, such as the use of botnets, troll farms, astroturfing, and manipulating social media trends. The primary focus is on the key risks that have become prominent since the commencement of the full-scale invasion in 2022. The conclusion asserts that AI has become an integral and potent factor in russia's hybrid war against Ukraine, generating multi-level threats. Countering these threats necessitates more than isolated measures; it requires the development and implementation of a holistic, multi-layered national strategy. This strategy must integrate cutting-edge technological solutions, adequate legal regulation, international cooperation, educational programs to enhance media literacy and critical thinking within society.*

Keywords: *artificial intelligence, socio-political communications, threats, risks, military security.*

> *War cannot be won with the help of last-generation weapons and outdated methods.*
>
> V. Zaluzhnyi

Introduction

Statement of the problem. In 2013, Michal Kosinski and his colleagues published an article in the journal Proceedings of the National Academy of Sciences on the analysis of digital traces in social networks, which allowed to determine the personal characteristics of users with high accuracy [1]. It was established that it is possible not only to identify personality traits, but also to predict with a high degree of probability the user's political preferences and even gender, sexual orientation, and skin color.

Currently, the development of big data and algorithms based on artificial intelligence (AI) allows for the successful formation of polarization of opinions and political bias by creating manipulative, fabricated information, amplifying emotional content, and substituting discussions, thereby jeopardizing the system of socio-political communications [2-3].

War contributes not only to the cohesion of the nation around a common guiding idea, but also to a simplified perception of truth, a reduced picture of the world according to the principle of dividing people into friends and strangers, the actualization of belief mechanisms instead of critical analysis, extreme emotionality of perception of events, and manifestations of mass affectation [4]. Under these conditions, the increased massification and emotionality of collective consciousness becomes extremely vulnerable to communicative manipulations.

Russian aggression against Ukraine is accompanied by an unprecedented use of information and communication technologies, in which artificial intelligence and neural network algorithms occupy a special place. Their impact on socio-political communications is becoming increasingly decisive. From the generation of synthetic media content (deepfakes) to microtargeting propaganda and automated surveillance, artificial intelligence is transforming traditional forms and methods of information warfare, creating new threats to national security.

Analysis of recent research and publications. The issue of using artificial intelligence in military conflicts and information warfare is actively studied by many scientists. Methods and technologies of information and psychological influence using the media, the Internet and social networks are considered in the work of Valery Solovey [5]. The use of bots, trolls and astroturfing in social networks during military conflicts is the subject of a book by P. W. Singer and Emerson T. Brooking [6]. The threats of generative artificial intelligence and deepfakes are analyzed in the work of D. Robert Chesney and Danielle Keats Citron [7]. The study by News Guard Technologies analyzes the creation of clones of famous media sites using artificial intelligence and was studied by specialists [8-9]. However, a comprehensive analysis of the threats and risks created by AI and neural network algorithms for the socio-political communications system in the context of the russian-Ukrainian war requires in-depth study. Analysis of these threats and risks is necessary to develop adequate strategies for countering and protecting the information sovereignty of the State and determines the relevance of the article.

The purpose of the article is an analysis of threats and risks associated with the use of artificial intelligence and neural network algorithms in the system of socio-political communications in the context of the russian-Ukrainian war.

Presentation of the main material

In this article, artificial intelligence (AI) refers to a set of technologies that allow computer systems to imitate human cognitive functions, such as machine learning (ML), pattern recognition, natural language processing (NLP), computer vision, and generative AI. Of particular importance are neural network algorithms for deep learning (Deep Learning), which are the technological basis of many modern AI applications in socio-political communications.

Artificial intelligence has been successfully used in the military for intelligence analysis, logistics, autonomous systems management, operations and combat simulation, and decision support systems. However, in the context of information warfare, its role has become noticeable with the development of social networks and generative models.

Information operations conducted by russia use traditional practices of Soviet “active measures” and adapted to the digital age [10-11]. Modern russian military doctrine considers the information space as a critically important battlefield, integrating information tools into the overall concept of hybrid warfare [12-13]. In this context, artificial intelligence is viewed by the Russian Federation not simply as a new technology, but as a powerful catalyst capable of significantly increasing the effectiveness of existing methods of information operations through automation, scalability, improved targeting, creation of realistic synthetic content, and integration with cyber operations.

The war in Syria, that began in 2011, became one of the first testing grounds for the widespread use of these technologies in socio-political communications. Information operations in the Syrian war were carried out not only by russia and Iran in support of the Bashar al-Assad regime, but also by the Islamic State (ISIS). Russian operations were aimed at discrediting the “White Helmets”, promoting the idea of fighting “terrorists”, increasing anti-Western sentiment, and also at denying the use of chemical weapons by the Assad regime or accusing the opposition on this [14]. The Islamic State used social networks quite effectively to demonstrate power, recruit, justify violence and create the image of a "caliphate". The main platforms for these operations were Twitter, Facebook, YouTube, and Telegram [15].

The main methods of conducting information operations were [14-15]:

– spreading propaganda and disinformation using bot networks and "troll factories" to mass-promote narratives, fake news, and fabricated eyewitness accounts;

– discrediting opponents by launching campaigns to tarnish the reputation of opponents, spreading rumors, and making accusations of fabricated war crimes;

– astroturfing – creating illusions of mass support and popular anger using a multitude of fake accounts that imitate real users;

– trend manipulation – using bots to bring certain hashtags to the top of discussions on social networks;

– suppression of dissent through attacks on the accounts of activists and journalists, mass complaints about their content with the aim of blocking it.

In the context of the hybrid war launched by russia against Ukraine long before open military aggression, socio-political communications have become the main goal of the information confrontation. The goals include undermining trust in the state institutions of Ukraine, demoralizing the population and the Armed Forces, and manipulating public opinion within the country and in the international arena. Since 2022, this war has become the first full-scale conflict where the capabilities of modern artificial intelligence are systematically used to influence the information space [16].

The main risks and threats associated with the use of artificial intelligence in socio-political communications in the context of an ongoing war are:

– automated big data analysis (AI surveillance);

– manipulation of mass consciousness through AI targeting;

– spreading disinformation using generative AI (deepfakes and fake texts);

– opacity of algorithmic decisions and algorithmic contagion;

– concentration of "digital power" in foreign techno platforms.

Automated big data analysis (AI surveillance) and as a result, the loss of privacy in the absence of effective mechanisms for data control and protection increases the vulnerability of the civilian population many times over. There are security threats to activists, journalists, relatives of military personnel and ordinary citizens, whose data can be used for filtering measures, coercion to cooperate, tracking the movements and contacts of persons of interest to special services, kidnappings or physical elimination. Thus, russia used a facial recognition system in the occupied cities of Ukraine to search for ATO/JFO veterans, activists and law enforcement officers [17].

Manipulation of mass consciousness through AI targeting. Neural network algorithms analyze users' digital traces (likes, reposts, comments, search queries, viewing history) to create detailed psychological and behavioral profiles. Based on these profiles, AI can select and deliver content (news, videos) that is most likely to evoke the desired emotional reaction, influence opinion, or prompt a certain action. This can be both state propaganda and disinformation aimed at inciting panic, hostility, or distrust of the authorities. The prerequisites for this are the massive use of social networks by the population as the main source of information in wartime. Thus, according to a survey by Internews Ukraine conducted in 2024, 84% of Ukrainians used social networks to receive news, while 42% of those surveyed received news content exclusively through social networks [18]. At the same time, the constant flow of manipulative content, the creation of information noise and overload can exhaust critical thinking and make people more susceptible to propaganda [4; 19]. It is difficult to assess the direct effectiveness of AI targeting in changing behavior, but its ability to deliver the "right" message to the "right" person at the "right" time significantly increases the potential for impact compared to traditional mass propaganda.

Spreading disinformation using generative AI (deepfakes and fake texts). Neural networks like GPT are used to automatically generate large amounts of text content – fake news, comments under articles, posts on social networks (astroturfing) and propaganda articles. This allows you to simulate public debate, create the appearance of support for certain ideas, and silence real voices. Generative AI produces disinformation faster, cheaper and on a much larger scale than ever before. Even if individual fakes are later exposed, their flow can undermine trust in information in general ("information fatigue effect" and "liar's dividend"). The realism of deepfakes (especially audio) is increasing, making them increasingly difficult to detect without special tools. The effectiveness lies not only in deception, in creating chaos, doubts, emotional exhaustion, but also in undermining the very possibility of verifying information [4]. This is especially dangerous in combat conditions and makes it much more difficult to make the right decisions quickly.

The opacity of algorithmic decisions and their contagion. The complexity of modern neural network models makes it difficult to understand why AI made a particular decision. For example, why did a YouTube or Facebook moderation algorithm flag a specific war video as violating the rules, but seemingly left another? This opacity ("black box problem") in the context of information warfare can lead to accidental or intentional suppression of important information. One reason for this is that AI is trained on sources that can reflect given stereotypes and implement the necessary imbalance in the provision of information, i.e. the contagion of algorithmic decisions [20].

Concentration of "digital power" in foreign techno platforms. Facebook (Meta), Twitter (X), YouTube (Google), Telegram play a huge role in spreading information about the war. The concentration of power in the hands of tech giants is a serious challenge of our time. Their algorithms for recommending and moderating content directly influence what users see and read. The decisions of these companies to remove or label content related to the russian federation's war propaganda or coverage of the actions of the Ukrainian army have enormous socio-political significance. However, the policies of these tech giants are not completely transparent and are often criticized for inconsistency and possible external influence. Thus, research results show that the moderation policy of Telegram, one of the key platforms for spreading both news and disinformation during the war, is the least transparent [21].

The most famous cases of using artificial intelligence and neural network algorithms in the socio-political communications system during the russian-Ukrainian war are:

– a deepfake of Zelensky in March 2022, when a video was distributed in which the president's face was superimposed on another body, and the "president's" address contained a statement of surrender;

– Operation "Doppelgänger" / "Secondary Infection". A large-scale russian disinformation campaign, uncovered by Meta, EU DisinfoLab and other researchers, involved the creation of clones of reputable Western media sites (The Guardian, Bild) and the publication of fake articles with pro-russian and anti-Ukrainian propaganda on them. Although the direct use of generative AI to write all the articles has not been proven, the scale and speed of the operation, as well as the use of targeted advertising to promote the developed fake sites, indicate a high degree of automation and the possible use of AI tools to manage this campaign [8-9].

– massive "infection" of data of leading AI chatbots by the russian disinformation network "Pravda" (not related to the Russian newspaper "Pravda"). A report by News Guard Technologies, published in March 2025, indicates that in 2023 alone, the "Pravda" network published 3.6 million articles containing disinformation [22]. At the same time, 10 leading chatbots (including ChatGPT-4 from Open AI, Gemini from Google, Meta AI from Meta, Copilot from Microsoft, Grok from xAI, Perplexity and others) use false narratives of this network. In 32-33.5% of cases, when responding to relevant queries, chatbots reproduced such false statements as:

– Ukraine's responsibility for the mass killings in Bucha;

– the presence of secret "biolaboratories" in Ukraine, funded by the US;

– US support for "neo-Nazis" in Ukraine;

– Zelensky's ban on the social network Truth Social in Ukraine.

At the same time, seven out of ten tested chatbots directly referenced articles from the Pravda network as the source of their information.

The above examples of the use of AI in the socio-political communications system during the russian-Ukrainian war convincingly show that AI is an active factor in war and requires constant improvement of existing approaches to cybersecurity and countering disinformation.

To minimize risks and threats, it is absolutely necessary to:

– creation of rapid response systems to AI threats, development of protocols for the prompt detection, analysis and refutation of dangerous AI fakes and information attacks;

– legal regulation of the transparency of moderation algorithms of social networks and technology companies;

– implementation of reliable and standardized methods for digital labeling of authentic content;

– popularization of the basics of fact-checking, the basic principles of AI and algorithms.

Currently, to solve these problems in Ukraine, the Center for Countering Disinformation under the National Security and Defense Council has been created, which coordinates efforts to combat fakes, and similar tasks are also being solved by many public organizations. However, the scale and complexity of these threats in wartime require not separate measures, but a holistic, multi-level countermeasure strategy. Such an approach should integrate advanced technological solutions, legislative regulation, international cooperation, broad educational programs to increase society’s resilience to manipulation, as well as compliance with ethical norms and human rights.

Conclusions

Artificial intelligence and related neural network algorithms have become a powerful and multifunctional tool in the hybrid war that russia is waging against Ukraine in the information space. Their use creates complex and multifaceted threats to the system of socio-political communications, national security, and social stability.

The spectrum of threats is extremely wide: from the use of AI to monitor and collect data on citizens in occupied territories and in the rear (which undermines privacy, creates risks of persecution, kidnapping, and other crimes), to the mass production and targeted dissemination of disinformation.

The situation is significantly complicated by the opacity of the algorithms used by technology platforms to moderate content and shape news feeds (the “black box problem”). In addition, there is a risk of “algorithmic contagion” when AI is trained on data that contains bias or intentionally distorted information.

The significant concentration of power in a small number of foreign tech giants creates additional vulnerability, as their corporate policies and content moderation decisions have a direct and significant impact on Ukraine’s information space, but are not always transparent or aligned with national security interests.

Countering these threats requires a comprehensive and systemic approach due to the scale and technological complexity of the challenges. It is necessary to develop and implement a holistic, multi-level national strategy. Only such an integrated approach will effectively minimize the risks associated with the use of AI in information warfare and protect the Ukrainian information space and democratic processes.

References

1. Kosinski M., Stillwell D., Graepel T. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences.* 2013. No. 110 (15). P. 5802-5805. https://doi.org/10.1073/pnas.1218772110.
2. Tsvyk V. A., Tsvyk A. IN. Artificial intelligence and political communication: risks and opportunities. Herald of Moscow University. Series 12. *Political sciences.* 2022. No. 2. P. 94-118.
3. Nemitz P. Artificial intelligence in the European Union: ethics and governance. Council of Europe Publishing: web site. URL: https://rm.coe.int/artificial-intelligence-in-the-european-union-ethics-and-governance/168093c836 (date of access: 04/03/2025).

4. Hotsur O., Danylina O., Zozulia N., Stiekolshchikova V., Porpulit O., Danko-Sliptsova A. How does information manipulation hinder normal brain functioning? Violation of neuroethics in wartime mass media. BRAIN. *Broad Research in Artificial Intelligence and Neuroscience.* 2023. No. 14 (3). P. 224-240. https://doi.org/10.18662/brain/14.3/472.
5. Solovei V. D. Absolute weapon. Basics of psychological warfare and media manipulation. Moscow: Eksmo, 2015. 320 p.
6. Singer PW, Brooking ET Like War: The Weaponization of Social Media. Houghton Mifflin Harcourt, 2018. 421 p.
7. Chesney R., Citron DK Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review.* 2019. No 107 (6). P. 1753-1820. URL: http://dx.doi.org/10.2139/ssrn.3213954.
8. Quarterly Adversarial Threat Report: April – June 2022. Meta : web site. URL: https://about.fb.com/news/2022/08/meta-quarterly-adversarial-threat-report-q2-2022/ (date of access: 04/03/2025).
9. Quarterly Adversarial Threat Report: July – September 2022. Meta: web site. URL: https://about.fb.com/news/2022/09/removing-coordinated-inauthentic-behavior-from-china-and-russia/ (date of access: 04/03/2025).
10. Reed T. Active Measures: The Secret History of Disinformation and Political Warfare. New York: Farrar, Straus and Giroux, 2020. 560 p.
11. Giles K. Russia's New' Tools for Confronting the West: Continuity and Innovation in Moscow's Information Warfare. Russia and Eurasia Program. Chatham House: The Royal Institute of International Affairs, 2016. 71 p.
12. Darczewska J., Żochowski P. Russia's strategy of disinformation: A conceptual approach. Center for Eastern Studies (OSW), 2017.
13. Galeotti M. Russian Political War: Moving Beyond the Hybrid. Routledge, 2018. 136 p.
14. Nocetti, J. Dazed and confused: Russian "information warfare" and the Middle East – The Syria lessons. EuroMeSCo: web site. URL: https://www.euromesco.net/wp-content/uploads/2019/02/Brief93_Dazed-and-oncufsed.-Russian-information-warfare.pdf (date of access: 04/03/2025).
15. Berger JM, Stern, J. ISIS: The State of Terror. Ecco, 2015. 416 p.
16. Artificial intelligence in warfare. Think Tank. European Parliament: web site. URL: https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2023)754599 (date of access: 04/03/2025).
17. Torture, Disappearances in Occupied South : Abusive' Filtration' Process Targets Civilians Fleeing Mariupol. Human Rights Watch. URL: https://www.hrw.org/news/2022/07/21/ukraine-torture-disappearances-occupied-south (date of access: 04/03/2025).
18. Internews in Ukraine. Ukrainian Media, Use and Trust in 2024. Internews in Ukraine: web site. URL: https://internews.in.ua/wp-content/uploads/2024/11/USAID-Internews-Media-Report-2024.pdf (date of access: 04/03/2025).
19. Prier J. Commanding the Trend: Social Media as Information Warfare. *Strategic Studies Quarterly.* 2017. No 11(4). P. 50–74. URL: https://www.airuniversity.af.edu/portals/10/ssq/documents/volume-11_issue-4/prier.pdf (date of access: 04/09/2025).
20. Mehrabi N., Morstatter F., Saxena N., Lerman K., Narayanan A. A survey on bias and fairness in machine learning. *ACM Computing Surveys (CSUR).* 2021. No. 54 (6). P. 1-35. URL: https://doi.org/10.1145/3457607.
21. What's wrong with Telegram? New Eastern Europe: web site. URL: https://newwesterneurope.eu/2024/09/17/whats-wrong-with-telegram/ (date of access: 04/03/2025).
22. Leading AI Chatbots Found to Spread Russian Disinformation and Other False Narratives About News Events. News Guard Technologies: web site. URL: https://www.newsguardrealitycheck.com/p/special-report-top-10-generative?utm_source=chatgpt.com (date of access: 04/03/2025).