Evgeny N. Pashentsev[1]
Institute of Contemporary International Studies,
Diplomatic Academy of the Ministry of Foreign Affairs
of the Russian Federation
Moscow (Russia)

Ivan S. Blekanov[2]
Saint-Petersburg State University,
Programming Technology Department
Saint-Petersburg (Russia)

Anastasia O. Chernobrivchenko[3]
International Centre for Social and Political Studies
Moscow (Russia)

342.72/.73(470)
004.8:004.056.54(470)
007:004.056.5(470)
343.53:004.738.5(470)
Review scientific paper
Submitted 21/02/2023
Accepted 10/04/2023
doi: 10.5937/socpreg57-42986

PERSONAL DATA PROTECTION IN RUSSIA AND THE RISKS OF MALICIOUS USE OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES: NEW CHALLENGES TO PSYCHOLOGICAL SECURITY

Abstract: The paper focuses on the relationship between personal data protection and technologies of artificial intelligence (AI) in the context of threats to the psychological security of society on the case study of the Russian Federation. The research identifies existing and prospective risks of malicious use of AI involving personal data to affect the psyche of people. The paper examines the possibilities of a comprehensive response to new threats of psychological security. The research methodology is based on the systemic approach, dialectical method and comparative analysis of the national and international components of the research problem.

Keywords: artificial intelligence, malicious use, personal data protection, psychological impact, digital privacy

1. INTRODUCTION

There is no doubt that the importance of artificial intelligence (AI) technologies in human development continues to grow. Without AI, it is impossible to achieve high productivity measure and ensure primacy in the global competitive market (Lepskij, Rajkov, 2022, p. 5). In recent years, AI has been repeatedly proclaimed as the most actively developing information and communication technology in the world (Glazkov et al., 2023, p. 10). In many countries, governments as well as private business and research centres pay

[1] icspsc@mail.ru
[2] i.blekanov@spbu.ru
[3] chernobrivchenko.ana@yandex.ru

significant attention to AI development and its implementation. Russia is no exception: an appropriate legal framework aimed at the development of AI has been created, the research done by private companies is encouraged by the authorities, and new technologies are being introduced into the daily lives of citizens.

At the same time, in order to learn how to act "similarly to human thinking mechanisms" (Kuteynikov, Izhaev, Zenin, Lebedev, 2019, p. 77), AI processes a huge amount of completely different types of data. The functioning and the use of AI technologies are inextricably linked to personal data, which includes information about a particular person or, formally, a data subject or his behavioural activities. Consequently, malicious use of AI (MUAI) can affect user security and information about them in various ways. It is important to study not only physical damage from MUAI but also its negative implications for psychological security of a person and society to determine the real harm caused to people and find ways to neutralize it, including by creating new legislative norms.

The very concept of AI as a system able to mimic human thought processes, even within a limited functionality, using large volumes of data about people, their intellectual and physiological characteristics, implies serious risks to society. If AI technologies are used with malice, there is not only the threat of damage to nature, technosphere and physical harm to humans, but also the danger of a serious negative effect on their psychological and mental state. Examining psychological security threats in the case of MUAI using personal data, it is important to underline that AI development and data security degree varies greatly from state to state. Moreover, neither well-developed data protection legislation (without taking into account the capabilities of AI), nor the lack of research in the field of AI in a state can protect its citizens from threats of MUAI using personal data. Any information about a person and his digital footprint on the internet can be used by intruders.

The article aims to identify the interconnection between personal data and AI technologies in the context of psychological consequences of its malicious use in the Russian Federation.

The researchers set the following objectives:

1) Determine the essence and levels of threats to psychological security through MUAI;
2) Detect the interrelation between personal data and AI in the context of its malicious use;
3) Examine the specific manifestations and psychological implications of MUAI related to personal data in Russia.

The research methodology is based on the systemic approach, dialectical method and comparative analysis of the national and international components of the research problem.

2. MALICIOUS USE OF ARTIFICIAL INTELLIGENCE AND PSYCHOLOGICAL SECURITY

The notion of psychological security can be found in many studies (Roshhin and Sosnin, 1995; Grachev, 1998; Afolabi and Balogun, 2017). Renowned US psychologist Abraham Maslow believed that once basic physiological needs are met, the need for security moves to the forefront. In more specific terms, it is the need for protection, stability, confidence in the future and good health, etc. Apart from personal security, a person also feels the

need for public security: people prefer certainty to uncertainty, and want to be confident that their environment is safe and free from threats (Maslow et al., 1945). Certain groups of researchers have proposed a distinction between psychological and cognitive security based on the separation of psychological and cognitive operation targets: "psychological operations are significantly different from cognitive operations aimed at the destruction of ... a holistic worldview" (Kefeli and Yusupov, 2017, p. 196). When a psychological operation results in the defeat of the enemy's will, a cognitive one leads to the defeat of individual consciousness. In the context of confronting MUAI, it is important to consider the cumulative impact of its destructive effect on the will and consciousness of individuals and society; the consequences of such an effect hinder social progress.

Threats from MUAI have become increasingly relevant with the growth of geopolitical rivalries, the activity of various state and non-state antisocial actors and the development and increasing accessibility of various AI technologies. It leads to nothing less than attempts by various interest groups to use AI to influence public consciousness for their own purposes.

Recent years have revealed great potential for MUAI in the psychological sphere. Despite a significant and rapidly growing number of academic publications on the technical aspects of MUAI, its general socio-economic and political implications, and the first attempts to classify MUAI (Brundage et al., 2018; Caldwell et al., 2020), there are relatively few publications on specific MUAI issues in the context of psychological security, and almost no one has yet begun a comprehensive examination of MUAI capabilities in terms of psychological security threats. Despite its importance, separate analysis of malicious psychological impact from deepfakes, bots, predictive analytics and so on, neither takes into account the synergy of their effects, nor provides a comprehensive view of the risks to the psychological security of individuals and to the entire international security system. The lack of comprehensive analysis is explained by the novelty of the issue. However, the risks of MUAI to psychological security exist nowadays and their impact on national and international development will continue to grow in the near future. A comprehensive study of this topic is necessary due to the massive cross-border spread of AI technologies, their relative accessibility and the possibility of targeted psychological impact on people of different gender, age, profession and nationality, etc. It does not matter that there has not yet been a negative experience of MUAI comparable to Hiroshima; it is crucial to make sure that worst-case scenarios never happen.

It is necessary to understand the opportunities for the psychological impact of various AI technologies that can be used by antisocial actors. This purpose requires research on various issues such as social engineering (Ozkaya, 2018), cyberpsychology (Aiken, 2017), the role of manipulative technologies in society (Grudin, 2006; Alvarez et al., 2009; Jacobs and Shapiro, 2002; King and Roth, 2006; Higdon et al., 2019; Woolley and Howard, 2018), psychological warfare (Brusnitsyn, 2001; Armistead, 2010; Paul, 2008; Welch, 2011; Bazarkina et al., 2020) and the role of media in political warfare (Hammond, 2008; Singer and Brooking, 2019; Simons, 2016). In the course of the study, the authors of the article worked with mass media publications on the MUAI topic that allowed us to trace how the specifics of the positioning of such issues in the media intentionally or unintentionally shape certain public biases. The entrenchment of prejudice in people's minds can affect the effectiveness of both MUAI and measures to counter it.

One of the authors of the article in his previous research studies identified three levels of threats to psychological security through MUAI (Pashentsev, 2020; Bazarkina, Pashentsev, 2020; Pashentsev, 2021). The first level is associated with the deliberate formation of a distorted (from excessively negative to excessively positive) attitude towards AI technologies. It may result in taking erroneous decisions and even worsening of socio-political situation in a country. At the second level psychological impact is directly connected with MUAI, but is not the main goal of a malicious act. On the contrary, at the third level AI is deliberately used to negatively influence the individual, group or public consciousness, in the long run – up to the establishment of effective and long-term control over them. The classification described allows any AI technology to be considered for real or potential threats of negative mental or psychological effects, even if there have been no previous cases of malicious use. Such an anticipatory analysis has substantial advantages when it comes to the risks of anti-social AI application up to large-scale destruction of public consciousness with subsequent processes of political, economic, military and cultural destabilization. The results of anticipatory analysis make it possible to develop preventive measures adapted to a particular situation.

The research devoted to the general issues of MUAI and psychological security provides a certain methodological basis for further case analysis of such threats as MUAI using personal data.

3. PERSONAL DATA PROTECTION AND PSYCHOLOGICAL RISKS OF MALICIOUS USE OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES: THE RUSSIAN EXPERIENCE

Before describing Russian experience in personal data protection and psychological risks of MUAI, it is worth mentioning that although the country is not among the leaders in the development of AI (The Global AI Index, n.d.), Russian society is quite receptive to the new ICTs emerging around the world. Thus, all types of psychological risks described earlier in the article are more or less relevant for the Russian Federation. Moreover, in Russia, AI technologies are already widely used in banking, road and transportation sector, medicine, services, recruitment, heavy industry and even in humanities research (Digital Petr, 2021). In view of ICT active development, taking into account a range of documents on strategic development of Russia (The Presidential Decree "About the Strategy of Scientific and Technical Development in Russia", 2016; Program "Digital economy of Russia", 2017; Executive Order on the development of AI in Russia, 2019) the authors decided to identify certain areas that are most susceptible to the psychological consequences of MUAI in Russia.

In Russia large banking companies are among the leaders in the development and implementation of AI technologies. In particular, the Russian government has recently signed an agreement on development of AI with one of the country's largest banks called "Sberbank" (The Russian Government, 2023). More than that, Russian digital banking is recognized as one of the most dynamic in the world (Wodzicki, Majewski, MacRae, 2020, p.8). Given these circumstances in Russia the malicious use of banking chat-bots may become a rather dangerous type of MUAI aimed at obtaining users' personal data. The

issue of malicious and even terrorist use of bots, created not for communication, has long been discussed in academia and professional circles. For instance, such bots are found to be used to manipulate public opinion and cause reputational damage including during election campaigns (Bazarkina, Pashentsev, 2019, p. 155), attract new members to criminal organizations and coordinate their activities (Mihalevich, 2022). Meanwhile, intruders use popular in Russia chat-bots for other purposes: logical vulnerabilities allow them to be used to steal bank customer data (Ilyina, 2021). Obviously, chat-bots can simply be hacked in order to obtain information directly from users. It is worth mentioning that this technology is also used in the Russian unified online system for providing public services to citizens called "Gosuslugi". Despite the fact that data leaks are impossible through chat-bot of "Gosuslugi", at the highest peak of the COVID-19 pandemic it was still subjected to a cyberattack: criminals used it to misinform people about the existence of the Coronavirus and threatened vaccinated citizens with death (Ushkov, Balashova, 2021). This example vividly illustrates that chat-bots are a vulnerable technology and its use by intruders can both cause psychological harm to an individual and affect psychological security of an entire country.

According to the 2021 research (Statista, 2021), more than 10% of Russian citizens regularly use smart voice assistants in their everyday life. For comparison, in the USA that is vying with China for leadership in the field of AI, this indicator reached 30% the same year (Edison Research, 2022). Thereby, the authors suggest that there is a real threat of voice assistant's malicious use in Russia. Hacking voice assistants may result in the same circumstances that are typical for cyberattacks on chat-bots. In addition, hijacking a smart home system or even simply connecting to a smart speaker through this technology would allow attackers to violate people's privacy and affect their psychological state by intercepting control of devices in their homes.

Considering the Russian experience in protecting personal data and psychological risks of MUAI, it is expedient to pay attention to deepfake technology. The fact that the first ever deepfake series was filmed in Russia in 2022 (PMZHejson, 2022) demonstrates the level of deepfakes technology development in the country. In parallel, this technology is used in business to manipulate people. There has been a reported case of an interview with a virtual interlocutor created with the use of deepfake technology (Adamov, 2022). Among the threats of malicious use of deepfakes (MUD) can be "Dead souls" frauds where a criminal steals the data of a deceased person in order to use their personality to make a profit. "Stolen identity" can be used to gain access to online services and accounts or to apply for credit cards, loans, etc. Besides, the created (synthesized) identity of a non-existent person can be used to conduct a large financial transaction or obtain credit (Panda Mediacenter, 2021).

In September 2021, fraudsters made a deepfake ad using the image of Tinkoff Bank founder Oleg Tinkov. In the video the fake billionaire encourages people to invest and receive bonuses by clicking on a link below. The fake ad was published on a fake Tinkoff Bonus Facebook page. Its profile picture resembled the logo of the bank. According to Fakecheck, when users clicked on a link below the video, they were redirected to a landing page with the bank logo where people were supposed to answer a few questions about investing and fill in a form with their name, email and phone number (Dulneva, Milukova, 2021). Obviously,

scams like this can easily provoke stress and panic of deceived people, especially in a critical situation. As deepfake technologies continue to improve and more effective schemes of manipulative influence emerge, their psychological impact will only grow.

Intelligent biometric identification system is one of the most actively developing area of AI in Russia. The main psychological risks connected with malicious use of such technologies are the theft of biometric data and their use for criminal purposes, in particular to create deepfakes or harass and blackmail people. Apparently, it violates their right to privacy and personal security. Nowadays facial recognition technologies are commonly used in video monitoring system in several cities of Russia. In 2018 police used them to enforce law and order at the World Cup in Russia (NtechLab, n.d.). Besides, Russia has one of the most developed "Safe City" systems in the world and a unified network of city cameras is a significant part of it (Koleganov, Kuvshinov, Pigina, Fedotov & Shedrov, 2021). Therefore, it is only a matter of time before the government begins to expand the geographic application of AI technologies for personal identification to ensure public safety, but it may also entail the emergence of new psychological risks of MUAI.

Contactless payment system also functions with the help of AI technologies that process users' biometric data. For instance, all lines of the Moscow metro operate a facial recognition system for fare payments. Although, there is no reason to consider this technology unreliable, it seems that there is still a lack of public confidence in this kind of technology. It is evidenced by the fact that in 1.5 years of full-scale operation of the project less than 2% of Moscow citizens have joined it (Deptrans Moskvy, 2022). In addition, the negative attitude towards the implementation of AI into the transport payment system is deliberately supported by a number of opponents of technological progress and human rights activists. They believe that such technologies are designed to establish government and enterprises control over Russian citizens. The described situation is a vivid example of the first-level threat, according to the classification given at the beginning of the article. In this case, unduly negative attitude to AI technologies leads to a delay in technological development of the country and a growing distrust in a government.

However, such a trend is gradually losing its relevance in Russia. The Russian experience shows that the proper state policy helps to reduce the influence of the first-level threat. According to the recent poll of Russian Public Opinion Research Centre (VCIOM) on Russians' attitude to AI, the level of people's trust in AI technologies as well as the degree of awareness of them, is actively growing, and citizens are becoming more objective in assessing the capabilities of AI and the risks associated with its use (VCIOM, 2022). Nevertheless, the risk of growing social tensions due to new AI technology introduction is still quite high and, in some spheres, people are not willing to entrust AI even with auxiliary functions. It appears that there is a lack of adequate assessment of AI capabilities and threats associated with it, not only in Russia but around the world. It means that governments should continue educational work and create conditions for the safest possible use of AI technologies.

Assessing the growth of threats of the malicious use of AI technologies in Russia is impossible without taking into account external risks in this area. The US Big Tech sector proved itself as a powerful tool for confronting Russia in cyberspace. Brad Smith, president and vice chair of *Microsoft*, writes in no uncertain terms about the role of his company in Ukraine.

> "Ukraine's government has successfully sustained its civil and military operations by acting quickly to disburse its digital infrastructure into the public cloud, where it has been hosted in data centres across Europe. This has involved urgent and extraordinary steps from across the tech sector, including by Microsoft. While the tech sector's work has been vital, it's also important to think about the longer-lasting lessons that come from these efforts" (Microsoft, 2022).

General Paul Nakasone, director of the National Security Agency, confirmed in his interview to Sky News as of June, 2022 that the United States had conducted offensive hacking operations in support of Ukraine: "We've conducted a series of operations across the full spectrum; offensive, defensive, [and] information operations" (Martin 2022). Such operations are impossible without the involvement of the Big Tech sphere. Thus, high-tech agenda setting in the United States, that today is unthinkable without the full use of AI technologies, has turned out to be openly subordinated to military and political interests and the needs of psychological warfare.

The use of Western-produced AI technologies in the ongoing military conflict in Ukraine is significant. US facial recognition start-up "Clearview AI" has provided technical support to Ukraine. Clearview AI's tools can identify faces in videos, comparing them to a company's database of 20 billion images from public networks and identifying potential spies and killed people. AI tools also play an important role in Ukraine's propaganda war and in processing critical information about the conflict. A program from the US company "Primer" can perform speech recognition, transcription and translation. It intercepts and analyzes Russian data, including conversations between Russian soldiers in Ukraine. A Swiss encrypted chat service called "Threema" allows Ukrainian users to send this data to the military without revealing their identities (Global Times, 2022).

To sum up, the complex of various internal and external development factors determines the existing and prospective risks of MUAI involving personal data to affect the psyche of people. It is fair to say that the use of personal data for AI development violates a number of fundamental data protection principles (Datatilsynet, 2018) enshrined both in General Data Protection Regulation (GDPR), which has become the benchmark, and in laws of some other states, including Russia (Federal Law "On Personal Data", 2006). The authors of the article offer several solutions to the outlined earlier problem. On the one hand, it is necessary to ensure compliance with the principles by limiting the scope of AI technologies application and at the same time significantly improving AI algorithms, for example, to eliminate biased decision-making. On the other hand, it is also possible to create a "compromise legislation" in which approaches to personal data and privacy will be developed taking into account the increasing role of AI technologies. In practice, there are cases when governments introduce light-touch regulation on data protection in order to create conditions for more effective implementation and development of AI. For instance, Moscow city has had such an "experimental regime" since 2020. It allows, *inter alia*, personal data to be used for the development of AI, provided that it is depersonalized (Lukackij, 2020). Despite the fact that this approach is sometimes criticized, it gives the Russian government an opportunity to ensure the development of useful and safe for society AI and, at the same time, identify sensitive areas that require additional legislative regulation.

4. CONCLUSION

Russia meets with internal and external threats of MUAI in the sphere of psychological security. Moreover, the latter are clearly increasing with the growth of international tensions, active hybrid warfare against Russia waged by the United States and its allies. Obviously, with the increasing development of AI in various states the probability of using practically any type of AI for unlawful purposes is becoming higher. Thus, it seems advisable to establish regional and international cooperation in order to jointly develop measures to counteract MUAI which is using personal data that threatens the security of all countries. Besides, interstate cooperation is also necessary to determine the interrelation between personal data and AI and to establish interdisciplinary standards. What is more, it is crucial not only to determine the cases in which the use of personal data for AI would be considered a breach, but also to work out measures for their protection, up to the restriction of the use and further development of AI under certain conditions.

The analysis has shown that in a large number of cases of MUAI personal data is acting both as an object and as a means of an unlawful act. In most situations people can suffer not only material and physical damage, but also psychological one. Consequently, it is extremely important to take into account the psychological security of individuals and society in the development of legislation and public policy in the field of AI and personal data protection. Governments should also involve psychologists and sociologists in this process. And last but not least, it is extremely important to deal with psychological impact of MUAI straight away without any hesitation as this problem to become particularly acute.

Евгениј Н. Пашенцев[1]
Институт за савремена међународна истраживања,
Дипломатска академија Министарства спољних послова Руске Федерације
Москва (Русија)

Иван С. Блеканов[2]
Државни универзитет у Санкт Петерсбургу, Катедра за технологију програмирања
Санкт Петербург (Русија)

Анастасија О. Чернобривченко[3]
Међународни центар за друштвена и политичка истраживања
Москва (Русија)

ЗАШТИТА ПОДАТАКА О ЛИЧНОСТИ У РУСИЈИ И РИЗИЦИ ЗЛОУПОТРЕБЕ ТЕХНОЛОГИЈА ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ: НОВИ ИЗАЗОВИ ЗА ПСИХОЛОШКУ БЕЗБЕДНОСТ

(Translation *In Extenso*)

Сажетак: Овај рад усредсређен је на однос између заштите података о личности и технологија вештачке интелигенције (енгл. AI) у контексту претњи по психолошку безбедност друштва у студији случаја Руске Федерације. Истраживање утврђује постојеће и будуће ризике злоупотребе вештачке интелигенције у вези са подацима о личности и утицај на људску психу. У раду се испитују могућности свеобухватног одговора на нове претње по психолошку безбедност. Истраживачка методологија заснована је на систематском приступу, дијалектичкој методи и упоредној анализи националних и међународних компоненти истраживачког проблема.

Кључне речи: вештачка интелигенција, злоупотреба, заштита података о личности, психолошки утицај, дигитална приватност

1. УВОД

Нема сумње да је значај технологија вештачке интелигенције све већи у људском развоју. Без вештачке интелигенције немогуће је остварити високу производну меру и осигурати првенство на глобалном конкурентном тржишту (Lepskij, Rajkov, 2022, p. 5). У последњих неколико година, за вештачку интелигенцију се стално истиче да је информациона и комуникациона технологија која се најактивније развија на

[1] icspsc@mail.ru
[2] i.blekanov@spbu.ru
[3] chernobrivchenko.ana@yandex.ru

свету (Glazkov et al., 2023, p. 10). У многим земљама владе, као и приватне компаније и истраживачки центри, придају огромну пажњу развоју и примени вештачке интелигенције. Русија није изузетак: усвојен је одговарајући законски оквир у погледу развоја вештачке интелигенције, приватне компаније спроводе истраживања уз подршку власти, а у свакодневни живот грађана уводе се нове технологије.

У исто време, да би се сазнало како треба поступати „слично човековим механизмима размишљања" (Kuteynikov, Izhaev, Zenin, Lebedev, 2019, p. 77), вештачка интелигенција обрађује огромну количину потпуно различитих врста података. Функционисање и употреба технологија вештачке интелигенције нераскидиво су повезани са подацима о личности, што подразумева информације о одређеној особи или, формално говорећи, о власнику података или његовом понашању и активностима. Самим тим, злоупотреба вештачке интелигенције може на разне начине утицати на безбедност корисника и на информације о њима. Важно је проучавати не само физичку штету услед овакве злоупотребе већ и њене негативне импликације по психолошку безбедност појединца и друштва како би се утврдила стварна штета нанета људима и пронашли начини за њену неутрализацију, укључујући и осмишљавање нових законодавних норми.

Сам концепт вештачке интелигенције као система способног да подражава процесе човековог размишљања, чак и у оквиру ограничене функционалности, користећи велике количине података о људима, њиховим интелектуалним и физиолошким особинама, указује на озбиљне ризике по друштво. Ако се технологије вештачке интелигенције користе злонамерно, постоји не само опасност од наношења штете природи, техносфери и физичке штете људима, већ и опасност озбиљног негативног утицаја на њихово психолошко и ментално стање. Приликом испитивања претњи по психолошку безбедност у случају злоупотребе технологија вештачке интелигенције, важно је нагласити да се степен развоја вештачке интелигенције и безбедности података у знатној мери разликује између држава. Осим тога, ни развијено законодавство о заштити података (не узимајући у обзир способности вештачке интелигенције) нити недостатак истраживања у области вештачке интелигенције у некој држави не могу заштитити грађане од опасности од злоупотребе ове интелигенције у коришћењу података о личности. Уљези могу употребити све информације о некој особи и њеном дигиталном отиску на интернету.

Рад има за циљ да препозна међусобну повезаност података о личности са технологијама вештачке интелигенције у контексту психолошких последица њене злоупотребе у Руској Федерацији.

Истраживачи су поставили следеће циљеве:

1) одредити суштину и нивое претњи по психолошку безбедност од злоупотребе вештачке интелигенције;
2) утврдити међусобну повезаност података о личности са вештачком интелигенцијом у контексту њене злоупотребе;
3) испитати посебне манифестације и психолошке импликације злоупотребе вештачке интелигенције у вези са подацима о личности у Русији.

Истраживачка методологија заснована је на систематском приступу, дијалектичкој методи и упоредној анализи националних и међународних компоненти истраживачког проблема.

2. ЗЛОУПОТРЕБА ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ И ПСИХОЛОШКА БЕЗБЕДНОСТ

Појам психолошке безбедности може се наћи у бројним студијама (Roshhin & Sosnin, 1995; Grachev, 1998; Afolabi & Balogun, 2017). Признати амерички психолог Ејбрахам Маслов веровао је да, када се испуне основне физиолошке потребе, у први план избија потреба за безбедношћу. Прецизније говорећи, то је потреба за заштитом, стабилношћу и веровањем у будућност и добро здравље итд. Осим личне безбедности, особа такође има потребу за јавном безбедношћу: појединци више воле извесност од неизвесности, те желе да буду уверени у то да је њихово окружење безбедно и без претњи (Maslow et al., 1945). Поједине групе истраживача предложиле су прављење разлике између психолошке и когнитивне безбедности на основу раздвајања психолошких и когнитивних оперативних циљева: „психолошке операције знатно се разликују од когнитивних које су усмерене на уништавање... холистичког погледа на свет" (Kefeli & Yusupov, 2017, p. 196). Када нека психолошка операција доведе до пораза непријатељеве воље, когнитивна операција доводи до пораза свести појединца. У контексту спречавања злоупотребе вештачке интелигенције, важно је сагледати збирни утицај разорног ефекта на вољу и свест појединаца и друштва; последице таквог ефекта спречавају друштвени напредак.

Претње од злоупотребе вештачке интелигенције постале су све важније због повећаних геополитичких супарништава, активности разних државних и недржавних антидруштвених актера и повећане доступности разних технологија вештачке интелигенције. То доводи чак и до покушаја разних интересних група да користе вештачку интелигенцију да у сопствене сврхе утичу на јавну свест.

У последњих неколико година откривен је велики потенцијал за злоупотребу вештачке интелигенције у сфери психологије. Упркос значајном и све већем броју академских публикација о техничким аспектима вештачке интелигенције, њеним општим друштвено-економским и политичким импликацијама, као и првим покушајима класификације злоупотребе вештачке интелигенције (Brundage et al., 2018; Caldwell et al., 2020), и даље је релативно мали број публикација о посебним питањима злоупотребе вештачке интелигенције у контексту психолошке безбедности, док нико још није почео свеобухватно испитивање злоупотребе вештачке интелигенције у смислу претњи по психолошку безбедност. И поред свог значаја, засебна анализа штетног психолошког утицаја дипфејка (енгл. *deepfakes*), ботова, предиктивне аналитике итд. не узима у обзир синергију њихових ефеката нити пружа свеобухватан поглед на ризике по психолошку безбедност појединаца и целокупног међународног безбедносног система. Недостатак свеобухватне анализе објашњава се тиме да је ово питање новијег датума. Међутим, данас постоје ризици од злоупотребе вештачке интелигенције по психолошку безбедност, који ће све више утицати на национални и међународни развој у ближој будућности. Свеобухватна студија ове теме неопходна је због масовног прекограничног ширења технологија вештачке интелигенције, њихове релативне доступности и могућности циљаног психолошког утицаја на људе различитог пола, старости, занимања и националности итд. Није важно то што се од Хирошиме наовамо није поновило негативно искуство у вези са злоупотребом вештачке интелигенције; пресудно је постарати се да се најгори сценарио никада не догоди.

Неопходно је схватити могућности психолошког утицаја разних технологија вештачке интелигенције који могу да врше антидруштвени актери. Ова сврха захтева истраживање разних питања као што су социјални инжењеринг (Ozkaya, 2018), сајберпсихологија (Aiken, 2017), улога манипулативних технологија у друштву (Grudin, 2006; Alvarez et al., 2009; Jacobs & Shapiro, 2002; King & Roth, 2006; Higdon et al., 2019; Woolley & Howard, 2018), психолошки рат (Brusnitsyn, 2001; Armistead, 2010; Paul, 2008; Welch, 2011; Bazarkina et al., 2020) и улога медија у политичком рату (Hammond, 2008; Singer & Brooking, 2019; Simons, 2016). Током истраживања, аутори овог рада прегледали су публикације масовних медија на тему злоупотребе вештачке интелигенције, који су нам омогућили да пратимо како специфично позиционирање ових питања у медијима намерно или ненамерно обликује извесне предрасуде у јавности. Укорењене предрасуде у људском уму могу да утичу на делотворност како злоупотребе вештачке интелигенције, тако и мера за њено спречавање.

Један од аутора овог рада је у претходним истраживањима препознао три нивоа претњи од злоупотребе вештачке интелигенције по психолошку безбедност (Pashentsev, 2020; Bazarkina, Pashentsev, 2020; Pashentsev, 2021). Први ниво повезан је са намерним стварањем искривљеног става (од крајње негативног до крајње позитивног) према технологијама вештачке интелигенције. То може довести до доношења погрешних одлука, па чак и до погоршања друштвено-политичке ситуације у некој земљи. На другом нивоу, психолошки утицај је непосредно повезан са злоупотребом вештачке интелигенције, али није главни циљ злонамерне радње. Напротив, на трећем нивоу, вештачка интелигенција се намерно користи за негативно утицање на свест појединца, групе или јавности у дужем временском периоду – до успостављања делотворне и дугорочне контроле над њима. Описана класификација омогућава да се свака технологија вештачке интелигенције сматра стварном или потенцијалном претњом негативних менталних или психолошких ефеката, чак и ако претходно није било случајева злоупотребе. Таква антиципаторна анализа има знатне предности када је реч о ризицима од антидруштвене примене вештачке интелигенције, све до великог уништавања јавне свести даљим процесима политичке, економске, војне и културолошке дестабилизације. Резултати антиципаторне анализе омогућавају развој превентивних мера прилагођених одређеној ситуацији.

Истраживања посвећена општим питањима злоупотребе вештачке интелигенције и психолошке безбедности пружају одређену методолошку основу за даљу анализу случаја ових претњи у вези са коришћењем података о личности.

3. ЗАШТИТА ПОДАТАКА О ЛИЧНОСТИ И ПСИХОЛОШКЕ ОПАСНОСТИ ОД ЗЛОУПОТРЕБЕ ТЕХНОЛОГИЈА ВЕШТАЧКЕ ИНТЕЛИГЕНЦИЈЕ: ИСКУСТВО РУСИЈЕ

Пре него што опишемо руско искуство са заштитом података о личности и психолошким ризицима од злоупотребе вештачке интелигенције, требало би поменути да, иако Русија није међу лидерима у развоју вештачке интелигенције (The Global AI Index, n.d.), руско друштво је прилично отворено за нове ИКТ које се јављају широм

света. Због тога су све врсте психолошких ризика претходно описане у овом раду више или мање релевантне за Руску Федерацију. Осим тога, у Русији се технологије вештачке интелигенције већ увелико примењују у банкарству, сектору друмског саобраћаја и превоза, медицини, услужном сектору, регрутовању, тешкој индустрији, па чак и у истраживањима друштвених наука (Digital Petr, 2021). Када је реч о активном развоју ИКТ, узимајући у обзир низ докумената о стратешком развоју Русије (председнички декрет „О стратегији научног и техничког развоја у Русији", 2016; програм „Дигитална економија Русије", 2017; Извршна наредба о развоју вештачке интелигенције у Русији, 2019), аутори су одлучили да препознају одређене области које су најосетљивије на психолошке последице злоупотребе вештачке интелигенције у Русији.

У Русији су велике банкарске компаније међу лидерима у развоју и примени технологија вештачке интелигенције. Пре свега, руска влада је недавно потписала споразум о развоју вештачке интелигенције са Сбербанком као једном од највећих банака у земљи (The Russian Government, 2023). Осим тога, руско дигитално банкарство препознато је као једно од најдинамичнијих на свету (Wodzicki, Majewski, MacRae, 2020, p. 8). С обзиром на околности у Русији, злоупотреба банкарских chat-bots може постати прилично опасна врста злоупотребе вештачке интелигенције чији је циљ добијање корисничких података о личности. Проблем злонамерне или чак терористичке употребе ботова који нису створени за комуникацију, већ дуго је предмет академске заједнице и стручних кругова. На пример, утврђено је да се ови ботови могу искористити за манипулисање јавним мњењем и за наношење штете репутацији, укључујући и током изборних кампања (Bazarkina, Pashentsev, 2019, p. 155), за привлачење нових чланова у криминалистичке организације и за координацију њихових активности (Mihalevich, 2022). У међувремену, уљези у друге сврхе користе популарне чет-ботове у Русији: логичке осетљиве тачке омогућавају њихово коришћење за крађу података о клијентима банке (Ilyina, 2021). Чет-ботови се очигледно могу хаковати да би се добиле информације директно од корисника. Треба поменути да се ова технологија користи и у руском обједињеном онлајн систему за пружање јавних услуга грађанима, познатом као „Госуслуги". Иако кроз чет-бот овог система није могуће цурење података, у јеку пандемије COVID-19 он је ипак био предмет сајбернапада: криминалци су га користили за погрешно информисање људи о коронавирусу и о томе да вакцинисаним грађанима прети смрт (Ushkov, Balashova, 2021). Овај пример сликовито илуструје то да су чет-ботови осетљива технологија, те да њена употреба од стране уљеза може нанети психолошку штету појединцима и утицати на психолошку безбедност читаве државе.

Према истраживању из 2021. године (Statista, 2021), више од 10% грађана Русије редовно користи тзв. паметне гласовне асистенте у свакодневном животу. Поређења ради, у САД, које се такмиче са Кином за преимућство у области вештачке интелигенције, овај показатељ достигао је 30% исте године (Edison Research, 2022). Стога аутори указују на постојање стварне опасности од злоупотребе гласовног асистента у Русији. Хаковање гласовног асистента може довести до истих околности које су типичне за сајбернападе на чет-ботове. Надаље, напад на систем паметне куће, па чак и повезивање на паметни спикер помоћу ове технологије омогућило би нападачима да повреде приватност људи и утичу на њихово психолошко стање уплитањем у контролу справа у њиховим домовима.

Имајући у виду искуство Русије у заштити података о личности и са психолошким ризицима злоупотребе вештачке интелигенције, неопходно је обратити пажњу на дипфејк технологију. Чињеница да је прва дипфејк серија снимљена управо у Русији 2022. године (PMZHejson, 2022) показује ниво развијености дипфејк технологије у овој земљи. Истовремено, ова технологија се користи у пословању да би се манипулисало људима. Забележен је случај интервјуа са виртуелним саговорником направљен помоћу дипфејк технологије (Adamov, 2022). Једна од опасности злоупотребе дипфејка могу бити преваре познате као „мртве душе", где преступник краде податке покојника како би његову личност употребио за стицање добити. Крађа идентитета може се искористити за приступ онлајн услугама и рачунима или за пријављивање за кредитне картице, зајмове итд. Осим тога, створени (синтетизовани) идентитет непостојеће особе може се употребити за обављање великих финансијских трансакција или за добијање кредита (Panda Mediacenter, 2021).

У септембру 2021. године, преваранти су направили дипфејк рекламу користећи слику Олега Тинкова, оснивача банке *Тинкоф*. На снимку се види како лажни милијардер подстиче људе да инвестирају и добију бонусе тако што ће кликнути на наведени линк. Лажна реклама објављена је на лажном профилу банке *Тинкоф* на Facebook-у. Профилна слика личила је на лого банке. Приликом провере на овој друштвеној мрежи (енгл. Fakecheck), када би корисници отишли на дати линк, били су преусмерени на страну са логотипом банке, где се од њих тражило да одговоре на неколико питања о инвестицијама и да попуне образац дајући своје име, електронску адресу и број телефона (Dulneva, Milukova, 2021). Овакве преваре очигледно веома лако могу довести до стреса и панике међу превареним људима, нарочито у критичној ситуацији. Како се дипфејк технологије и даље унапређују и јављају се све ефикасније шеме манипулативног утицаја, то ће и њихов психолошки утицај бити све већи.

Интелигентни биометријски систем идентификације једна је од области која се најактивније развија у Русији. Главни психолошки ризици повезани са злоупотребом ових технологија јесу крађа биометријских података и њихова употреба у криминалне сврхе, нарочито за стварање дипфејка или за узнемиравање и уцењивање људи. Овде је очигледно реч о кршењу права на приватност и личну безбедност. Данас се технологије за препознавање лица обично користе у системима видео-надзора у неколико градова Русије. Године 2018. полиција их је користила за успостављање реда и мира на Светском купу у Русији (NtechLab, n.d.). Осим тога, Русија има један од најразвијенијих система „безбедног града" (енгл. Safe City) на свету и обједињену мрежу градских камера у већини градова (Koleganov, Kuvshinov, Pigina, Fedotov & Shedrov, 2021). Стога је само питање тренутка када ће владе почети да шире географску примену технологија вештачке интелигенције ради личне идентификације како би се осигурала јавна безбедност, али то исто тако може подразумевати појаву нових психолошких ризика од злоупотребе вештачке интелигенције.

Систем бесконтактног плаћања такође функционише помоћу технологија вештачке интелигенције које обрађују биометријске податке корисника. На пример, све линије московског метроа приликом наплате карата користе систем за препознавање лица. Иако нема разлога да се ова технологија сматра непоузданом, изгледа да ипак постоји неповерење јавности у ову врсту технологије. То доказује чињеница да се

за годину и по свеобухватног функционисања пројекта, за њега пријавило мање од 2% Московљана (Deptrans Moskvy, 2022). Поред тога, негативан став према примени вештачке интелигенције у систему планирања превоза намерно подржава известан број противника технолошког напретка и активиста за људска права. Они тврде да су ове технологије смишљене да би влада и предузећа успоставили контролу над грађанима Русије. Описана ситуација је сликовит пример претње првог нивоа, према класификацији датој на почетку рада. У овом случају, непримерен негативни став према технологијама вештачке интелигенције доводи до кашњења у технолошком развоју земље и до све већег неповерења у владу.

Међутим, овакав тренд постепено губи значај у Русији. Руско искуство показује да правилна државна политика помаже у смањивању утицаја претње првог нивоа. Према недавно спроведеној анкети Руског центра за истраживање јавног мњења (VCIOM) о ставу Руса према технологијама вештачке интелигенције, ниво поверења људи у ове технологије, као и степен свести о њима, у активном је порасту, а грађани су све објективнији у процени капацитета вештачке интелигенције и ризика повезаних са њеном применом (VCIOM, 2022). Па ипак, ризик од све већих друштвених тензија због увођења технологија вештачке интелигенције и даље је висок, а у појединим областима људи нису спремни да вештачкој интелигенцији повере чак ни помоћне функције. Чини се да постоји недостатак одговарајуће процене капацитета вештачке интелигенције и претњи повезаних са њом, не само у Русији већ широм света. То значи да државе треба да наставе образовни рад и да стварају услове за најбезбеднију могућу примену технологија вештачке интелигенције.

Немогуће је проценити пораст претњи од злоупотребе технологија вештачке интелигенције у Русији уколико се не узму у обзир екстерни ризици у овој области. Амерички сектор високе технологије доказао се као моћна алатка за сукобљавање са Русијом у сајберпростору. Бред Смит, председник и касније потпредседник компаније *Microsoft*, пише сасвим отворено о улози његове компаније у Украјини.

> „Украјинска влада је успешно спровела своје цивилне и војне операције тако што је брзо реаговала активирањем своје дигиталне инфраструктуре у јавном „облаку", одакле су је преузимали центри података широм Европе. Ово је подразумевало хитне и ванредне кораке у читавом технолошком сектору, укључујући и *Microsoft*. Иако је рад технолошког сектора био од виталног значаја, такође је важно мислити и о дугорочнијим лекцијама добијеним из оваквих напора" (Microsoft, 2022).

Генерал Пол Накасоне, директор Националне безбедносне агенције, у свом интервју за *Sky News* у јуну 2022. године потврдио је да су Сједињене Државе обавиле операције офанзивног хаковања као подршку Украјини: „Спровели смо низ операција у читавом спектру; офанзивне, дефанзивне [и] информационе операције" (Martin, 2022). Ове операције нису могуће без ангажовања високотехнолошке сфере. Самим тим је високотехнолошка агенда у Сједињеним Државама, која је данас незамислива без пуне примене технологија вештачке интелигенције, постала отворено подређена војним и политичким интересима и захтевима психолошког ратовања.

Значајна је примена технологија вештачке интелигенције створених на Западу у актуелном војном сукобу у Украјини. Америчка стартап технологија за препознавање

лица (*Clearview AI*) пружила је техничку подршку Украјини. Алатке ове вештачке интелигенције могу да препознају лица на снимцима, да их упореде са базом података ове компаније у којој има 20 милијарди слика са јавних мрежа, и на тај начин препозна потенцијалне шпијуне и погинуле људе. Алатке вештачке интелигенције такође имају значајну улогу у украјинском пропагандном рату и у обради критичних информација о сукобу. Програм америчке компаније *Primer* може да врши препознавање говора, транскрипцију и превод. Програм пресреће и анализира руске податке, укључујући и разговоре између руских војника у Украјини. Швајцарски шифровани сервис за ћаскање по имену *Threema* омогућава корисницима у Украјини да шаљу ове податке војсци не откривајући свој идентитет (Global Times, 2022).

У кратким цртама, комплекс разних интерних и екстерних развојних фактора одређује постојеће и потенцијалне ризике од злоупотребе вештачке интелигенције у вези са подацима о личности, која би могла да утиче на људску психу. Правилно је рећи да употреба података о личности у сврху развоја вештачке интелигенције крши бројна основна начела заштите података (Datatilsynet, 2018) зацртана и у Општој одредби за заштиту података (енгл. GDPR), која је постала репер, и у законима појединих држава, укључујући Русију и њен Савезни закон о подацима о личности (2006). Аутори овог рада нуде неколико решења за претходно наведен проблем. С једне стране, неопходно је осигурати усклађеност са начелима тако што ће се ограничити домен примене технологија вештачке интелигенције и, у исто време, значајно унапредити њени алгоритми, на пример, да би се елиминисало одлучивање засновано на предрасудама. С друге стране, такође је могуће створити „компромисно законодавство" у којем ће бити развијени приступи подацима о личности и приватности узимајући у обзир све већу улогу технологија вештачке интелигенције. У пракси, постоје случајеви када владе уводе тзв. меку регулативу о заштити података како би се створили услови за делотворнију примену и развој вештачке интелигенције. На пример, Москва примењује овакав „експериментални режим" још од 2002. године. То, између осталог, омогућава коришћење података о личности за развој вештачке интелигенције, али под условом да су подаци деперсонализовани (Lukackij, 2020). Упркос томе што овај приступ понекад наилази на критике, он руској влади даје прилику да осигура развој вештачке интелигенције корисне и безбедне за друштво и да истовремено препозна осетљиве области које захтевају додатну законодавну регулативу.

4. ЗАКЉУЧАК

Русија се суочава са интерним и екстерним претњама које са собом носи злоупотреба вештачке интелигенције у области психолошке безбедности. Осим тога, ове претње се упадљиво повећавају упоредо са растом међународних тензија, активног хибридног рата који против Русије воде Сједињене Америчке Државе и њихови савезници. Очигледно је да се са убрзаним развојем вештачке интелигенције у разним земљама повећава и вероватноћа коришћења дословно свих врста вештачке интелигенције у незаконите сврхе. Стога се препоручује успостављање регионалне и међународне сарадње како би се заједно развиле мере за супротстављање злоупотреби вештачке интелигенције када је реч о подацима о личности, која је опасна по

безбедност свих држава. Осим тога, међудржавна сарадња такође је потребна да би се одредио међусобни однос између података о личности и вештачке интелигенције и успоставили међудисциплинарни стандарди. Исто тако, пресудно је не само одредити случајеве у којима би се употреба података о личности сматрала прекршајем већ исто тако разрадити мере њихове заштите, укључујући ограничавање употребе и даљег развоја вештачке интелигенције у појединим случајевима.

Анализа показује да у великом броју случајева злоупотребе вештачке интелигенције подаци о личности служе и као предмет и као средство незаконитих радњи. У већини ситуација људи могу да претрпе не само материјалну и физичку штету већ и психолошку. Самим тим, изузетно је важно узети у обзир психолошку безбедност појединаца и друштва у састављању законодавства и јавне политике у области вештачке интелигенције и заштите података о личности. Владе би у овај процес такође требало да укључе психологе и социологе. На крају, подједнако је важно бавити се психолошким утицајем одмах и без одлагања јер ће овај проблем с временом постајати све озбиљнији.

REFERENCES / ЛИТЕРАТУРА

Adamov, D. (2022, October 20). "Agency": They tried to create using deepfake technology a Russian billionaire, who allegedly bought the cement company Holcim. *RTVI*. Available at: https://rtvi.com/news/agentstvo-rossijskogo-milliardera-yakoby-kupivshego-czementnuyu-kompaniyu-holcim-pytalis-sozdat-s-pomoshhyu-tehnologii-dip-fejk/?ysclid=ldbo99k8e8657309437 [In Russian]

Aiken, M. (2017). *The Cyber Effect: An Expert in Cyberpsychology Explains How Technology Is Shaping Our Children, Our Behavior, and Our Values — and What We Can Do About It.* New York: Random House.

Alvarez, R., Hall T., Hyde S. (eds.) (2009). *Election Fraud: Detecting and Deterring Electoral Manipulation.* Washington DC: Brookings Institution Press.

Armistead, L. (2010). *Information Operations Matters. Best Practices.* Washington, DC: Potomac Books.

Bazarkina, D., Pashentsev, E. (2019). Artificial Intelligence and New Threats to International Psychological Security. *Russia in global affairs,* (1), 147–170. https://doi.org//10.31278/1810-6374-2019-17-1-147-170

Bazarkina, D., Pashentsev, E., Simons, G. (eds.) (2020). *Terrorism and Advanced Technologies in Psychological Warfare: New Risks, New Opportunities to Counter the Terrorist Threat.* New York: Nova Science Publishers.

Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B., Dafoe, A., Scharre, P., Zeitzoff, T., Filar, B., Anderson, H., Roff, H., Allen, G., Steinhardt, J., Flynn, C., Ó HÉigeartaigh, S., Beard, S., Belfield, H., Farquhar, S., Lyle, C., Crootof, R., Evans, O., Page, M., Bryson, J., Yampolskiy, R., Amodei, D. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. Oxford: Future of Humanity Institute, University of Oxford.

Brusnitsyn, N. (2001). *Information Warfare and Security*. Moscow: Vita. [In Russian]

Caldwell, M., Andrews, J., Tanay, T., Griffin, L. (2020). AI-enabled future crime. *Crime Science,* (9), 14. https://doi.org/10.1186/s40163-020-00123-8

Datatilsynet. (2018). Artificial intelligence and privacy Report. *Datatilsynet*. https://www.datatilsynet.no/

Deptrans Moskvy. (2022, December 20). Starting from January 2023, we will start connecting 330 more turnstiles to the biometric payment system — it will become even more convenient to use public transport. *Deptrans Moskvy.* Available at: https://t.me/DtRoad/22191 [In Russian]

Dulneva, M., Milukova, Ya. (2021, September 6). "Hugged everyone!": the image of Oleg Tinkov was used in deepfake advertising. *Forbes*. Available at: https://www.forbes.ru/milliardery/439255-vseh-obnal-obraz-olega-tin-kova-ispol-zovali-v-dipfejk-reklame [In Russian]

Edison Research. (2022). The Smart Audio Report. *NPM*. Available at: https://www.nationalpublicmedia.com/insights/reports/smart-audio-report/

Executive Order on the development of AI in Russia №490. (2019, October 10). Kremlin. Available at: http://static.kremlin.ru/media/events/files/ru/AH4x6HgKWANwVtMOfPDhcbRpvd1HCCsv.pdf [In Russian]

Federal Law "On Personal Data" № 152-FZ. (2006, July 27). Kremlin. Available at: http://www.kremlin.ru/acts/bank/24154 [In Russian]

Glazkov, B., Mitkin, A., Semibratov, N., Krasovskiy, P., Naumtseva, E., Skvirskaya, O., Karev, I. (2023). *Monitoring of global digitalization trends 2022*. Moscow: Rostelecom. Available at: rostelekom_monitoring_2022_rasterize (rt.ru) [In Russian]

Global Times. (2022, November 2). From commercial satellites to social media, Western tech companies are deeply involved in the Russia-Ukraine conflict. Teller Report. Available at: https://www.tellerreport.com/news/2022-11-02-from-commercial-satellites-to-social-media--western-tech-companies-are-deeply-involved-in-the-russia-ukraine-conflict.HJSuXB1Bo.html

Grudin, R. (2006). *American Vulgar: The Politics of Manipulation Versus the Culture of Awareness.* Emeryville, CA: Shoemaker and Hoard.

Hammond, P. (2008). *Media, war and postmodernity*. London: Routledge.

Higdon, N., Huff, M., Nader, R. (2019). *United States of Distraction: Media Manipulation in Post-Truth America (And What We Can Do About It)*. San Francisco, CA: City Lights Publishers.

Ilyina, N. (2021, September 1). Cheating by correspondence: vulnerabilities in bank chatbots allow money theft. *Izvestiya*. Available at: https://iz.ru/1214668/natalia-ilina/obman-po-perepiske-uiazvimosti-v-bankovskikh-chat-botakh-pozvoliaiut-krast-dengi [In Russian]

Jacobs, R., Shapiro, R. (2002). *Politicians Don't Pander: Political Manipulation and the Loss of Democratic Responsiveness.* Chicago: University of Chicago Press.

Kefeli, I., Yusupov, R. (eds.) (2017). *Psychological and Cognitive Security*. Saint Petersburg: Petropolis. [In Russian]

King, S., Roth, R. (2006). *Broken Trust: Greed, Mismanagement & Political Manipulation at America's Largest Charitable Trust.* Honolulu, HI: University of Hawaii Press.

Koleganov, S., Kuvshinov, D., Pigina, S., Fedotov, A., Shedrov, I. (2021). "Safe city" in the world regions: comparative analysis with the Russian concept. *Civil Security Technologies*, vol. 3 (69), 56-60. https://doi.org/10.54234/CST.19968493.2021.18.3.69.11.55 [In Russian]

Kuteynikov, D., Izhaev, O., Zenin, S., Lebedev, V. (2019). The nature and legal features of cyber-physical, cyber-biological and artificial cognitive systems. *Rossijskoe pravo: obrazovanie, praktika, nauka*, vol. 3 (111), 75-81. https://doi.org//10.34076/2410-2709-2019-3-75-81 [In Russian]

Lepskij, V., Rajkov, A. (2022). *Socio-humanitarian aspects of digital transformations and artificial intelligence*. Moscow: Kogito-Tsentr. [In Russian]

Lukackij, A. (2020). Cybersecurity of digital transformation. In: S. Kirushin, E. Borisov, V. Opredelenov (eds.) *Tutorial 4CDTO. About digital transformation and digitalization.3.0. Moscow: 4CIO.* [In Russian]

Martin, A. (2022, June 1). US military hackers conducting offensive operations in support of Ukraine, says head of Cyber Command. Sky News. Available at: https://news.sky.com/story/us-military-hackers-conducting-offensive-operations-in-support-of-ukraine-says-head-of-cyber-command-12625139

Microsoft. (2022, November 2). Defending Ukraine: Early Lessons from the Cyber War. Microsoft Corporation. Available at: https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RE50KOK

Mihalevich, E. (2022). Malicious use of artificial intelligence was discussed at UNESCO Conference in Khanty-Mansiysk. *RIAC*. Available at: https://russiancouncil.ru/analytics-and-comments/columns/cybercolumn/zlonamerennoe-ispolzovanie-iskusstvennogo-intellekta-obsudili-na-konferentsii-yunesko-v-khanty-mansi/ [In Russian]

NtechLab. N.d. Video analytics for FIFA 2018. *NtechLab*. Available at: https://ntechlab.ru/success-stories/fifa/ [In Russian]

Ozkaya, E. (2018). *Learn Social Engineering: Learn the art of human hacking with an internationally renowned expert.* Birmingham: Packt Publishing.

Panda Mediacenter. (2021, August 10). Deepfake Fraud: Security Threats Behind Artificial Faces. *Panda Mediacenter*. Available at: https://www.pandasecurity.com/en/mediacenter/technology/deepfake-fraud/

Pashentsev, E. (2020). AI and Terrorist Threats: The New Dimension for Strategic Psychological Warfare. In: Bazarkina D., Pashentsev E. and Simons G. (eds.) *Terrorism and Advanced Technologies in Psychological Warfare: New Risks, New Opportunities to Counter the Terrorist Threat.* New York: Nova Science Publishers.

Pashentsev, E. (2021). *Experts on the Malicious Use of Artificial Intelligence and Challenges to International Psychological Security. Edition of the International Center for Social and Political Studies and Consulting.* Moscow: LLC "SAM Polygraphist". Available at: https://www.researchgate.net/publication/356781944_Experts_on_the_Malicious_Use_of_Artificial_Intelligence_and_Challenges_to_International_Psychological_Security_Report_by_Evgeny_Pashentsev_Center_for_Social_and_Political_Studies_and_Consulting_Decemb

Pashentsev, E., Bazarkina, D. (2020). Malicious Use of Artificial Intelligence New Psychological Security Risks in BRICS Countries. *Russia in global affairs*, (4), 154–177. Available at: https://eng.globalaffairs.ru/wp-content/uploads/2020/12/154-177.pdf?ysclid=lbdw0ab3nv36962900

Program "Digital economy of Russia" № 1632-p. (2017, July 28). The Russian Government. Available at: http://static.government.ru/media/files/9gFM4FHj4PsB79I5v7yLVuPgu4bvR7M0.pdf [In Russian]

Sber AI. (2021). Digital Petr. *Sber AI.* Available at: https://sber.ru/digital-petr/ [In Russian]

Semenihina, A. (Producer), Salimianov, V. (Director). (2022). *PMZHejson*. Russia: Agenda Media Group. Available at: https://www.youtube.com/watch?v=TEHt2e1RSug&list=PLWTwWADrHvpkgv3cKyjomdfhESt5711OZ&index=1 [In Russian]

Simons, G. (2016). *Mass media and modern warfare*. London: Routledge.

Singer, P., Brooking, E. (2019). *LikeWar: The Weaponization of Social Media*. Mariner Books, Boston, MA

Statista Research Department. (2021). Report "Voice assistants home usage in Russia 2021". *Statista*. Available at: https://www.statista.com/statistics/1258819/voice-assistants-home-usage-russia/

Stupp, C. (2019). Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case. *The Wall Street Journal*. Available at: https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402

Tortoise. (2022). Report "The global AI Index". *Tortoise*. Available at: https://www.tortoisemedia.com/intelligence/global-ai/#

The Presidential Decree "About the Strategy of Scientific and Technical Development in Russia" №642. The Russian Government. (2016, December 1). Available at: http://government.ru/docs/all/109256/ [In Russian].

The Russian Government. (2023, January 23). The Government signed the final package of cooperation agreements on the development of high-tech areas. *The Russian Government.* Available at: http://government.ru/news/47551/

Uskov, M., Balashova, A. (2021, November 11). The authorities announced an attack on "Gosuslugi" after reports of an anti-vaxxer bot. *RBC*. Available at: https://www.rbc.ru/technology_and_media/11/11/2021/618d42109a7947252fe7d448 [In Russian]

VCIOM. (2022, December 28). Report "Artificial Intelligence: threat or better tomorrow?" *VCIOM*. Available at: https://wciom.ru/analytical-reviews/analiticheskii-obzor/iskusstvennyi-intellekt-ugroza-ili-svetloe-budushchee?ysclid=lct0f6jfee970504040 [In Russian]

Veldkamp, D. (2022). Cyber Awareness 2022: Consider Deepfakes, NFTs, and More. *InfoSystems.* Available at: https://infosystems.biz/cybersecurity/cyber-awareness-2022-consider-deepfakes-nfts-and-more/.

Welch, M. (2011). *Irregular pen and limited sword: PSYWAR, PSYOP, and MISO in counter-insurgency.* U.S. Fort Leavenworth, KS: Army Command and General Staff College.

Wodzicki, M., Majewski, M., MacRae, M. (2020). *Digital Banking Maturity 2020*. Deloitte. Digital. Available at: https://www2.deloitte.com/content/dam/Deloitte/ce/Documents/financial-services/ce-digital-banking-maturity-2020.pdf

Woolley, S., Howard, P. (2018). *Computational Propaganda: Political Parties, Politicians, and Political Manipulation on Social Media.* New York: Oxford University Press.