THE PROBABILITY OF MILITARY AGGRESSION OF AUTONOMOUS AI: ASSUMPTIONS OR IMMINENT REALITY (ANALYZING THE FACTS OF RUSSIAN WAR AGAINST UKRAINE)

Kostenko O.V.,
Doctor of Philosophy (Ph.D.) in Law,
Head of the Scientific Laboratory of Digital Transformation
Theory and Law of the Scientific Center for
Digital Transformation and Law of the State Scientific Institution
«Institute of Information, Security and Law of the
National Academy of Legal Sciences of Ukraine»,
Ukraine, ORCID ID: https://orcid.org/0000-0002-2131-0281
antizuk@gmail.com

Kostenko O.V. The probability of military aggression of autonomous AI: assumptions or imminent reality (analyzing the facts of russian war against Ukraine).

The application of modern technologies with artificial intelligence in all spheres of human life is growing exponentially. There is a real concern that this growth could become uncontrollable. The lack of public, state and international control over artificial intelligence technologies creates large-scale risks of using such software and hardware that inevitably affect or, conversely, intentionally harm humanity. The events of recent weeks and Russia's war against democratic Ukraine confirm the thesis that the uncontrolled use of AI, especially in the military sphere, can lead to complete non-compliance with the moral norms of the use of controlled AI or the spontaneous emergence of aggressive autonomous AI.

The development of legal regulation of the use of technologies with artificial intelligence is being carried out rather slow, in comparison to the rapid development of AI technologies, which simultaneously cover all areas of public relations. Therefore, control over the creation and use of AI should be carried out not only by technical regulation (requirements, technical standards, regulations, assessments of compliance with technical standards, control of compliance with technical regulations) but also by the developing of standard laws and specific changes in the information, civil, criminal and other spheres of law, by creating comprehensive legislation and the state supervisory authority in this area.

Key words. Artificial intelligence, autonomous artificial intelligence, robot, war against Ukraine, neural networks, ethics, morality, international supervisory body.

Костенко О.В. Вірогідність військової агресії автономного AI: припущення чи не-далека реальність (аналізуючи факти війни росії проти України).

Застосування сучасних технологій зі штучним інтелектом у всіх сферах життєдіяльності люди-ни зростає у геометричній прогресії. Існує цілком реальне занепокоєння, що це зростання може стати неконтрольованим. Відсутність суспільного, державного та міжнародного контролю за тех-нологіями штучного інтелекту формує масштабні ризики застосування таких програмно-апаратних засобів, які ненавмисно заподіють або навпаки навмисно нанесуть шкоду людству. Події останніх тижнів та війна Росії проти демократичною Укра-їни підтверджують тезу про те, що неконтрольо-ване застосування AI, перш за все у військовій сфері, може призвести до навмисного нехтування моральними нормами застосування керованого AI або спонтанного виникнення агресивного авто-номного AI.

Розробка правового регулювання застосування технологій зі штучним інтелектом наразі відбува-ється вкрай повільно, по відношенню до стрімкого розвитку технологій AI, які одночасно охоплюють всі сфери суспільних відносин. Тому контроль за створенням та використанням AI необхідно здійс-нювати не тільки суто технічним регулюванням (вимоги, технічні стандарти, регламенти, оцінки відповідності технічним стандартам, контроль від-повідності вимогам технічних регламентів) але й розробкою Типових законів та конкретних змін в інформаційному, цивільному, кримінальному та інших галузях права, шляхом створення комплек-сного законодавства та міждержавного наглядо-вого органу в даній сфері.

Не викликає сумніву доцільність створення міжнародного наглядово-керуючюого органу, на який покласти завдання нагляду та контролю за застосуванням морального AI та розповсюджен-ням автономного AI військового або подвійного призначення. Запровадити єдину міжнародну ав-

томатизовану систему сертифікації та ліцензування розробок з алгоритмами штучного інтелекту.

Доцільно створити єдину міжнародну систему фіксації та реагування на інциденти в системах з алгоритмами штучного інтелекту. Розпочати формування міжнародної Типової моделі технічних, біологічних, фінансових, економічних, політичних та військових загроз застосування (санкціонованого та несанкціонованого) систем на основі AI.

The problem of the morality of autonomous artificial intelligence has been a source of concern for scientists and researchers since the rapid application of rapid mathematical algorithms for processing large data sets. In 1956, at a scientific conference at Dartmouth College (USA), the notion of «piecewise intelligence» was formed with the help of the attending researchers. Today, these are state-of-the-art, sophisticated and high-speed algorithms that work with big data to form output predictions with a given timeliness, which are used for decision-making in various fields of science, technology and social sphere. Problems of artificial intelligence and the metaverse are highlighted in previous work, which formed the key postulates of social relations with the use of technologies of the metaverse, artificial intelligence, artificial neural networks, robotic systems [1]. Let's consider the problems of autonomous AI.

Today, numerous scientific discussions are going on about the use of AI. One of the key discussions is the morality of autonomous AI.

At the same time, there are two hypotheses among scientists concerning the morality of autonomous AI. According to the first hypothesis, autonomous AI can be exclusively moral from a philosophical point of view and exist in order with humanity, while fulfilling the task of positive sustainable development of society. The second hypothesis does not exclude the possibility of forming an autonomous AI (artificially or spontaneously), which will act independently, that is, without receiving instructions from humans, and possessed by the purpose of domination over the human race or its destruction.

It is quite obvious that scientists are focused on research and development of positively moral AI. Currently, the models of moral AI are diverse: mixed, capable of using different algorithms and methods of data processing; endowed with the functions of «making» moral decisions [2]; based on the OSI model, which perform a set of ethical rules one by one "from the top-down" [3]; anti-OSI with compliance with certain ethical rules «bottom up» [4]; training systems with reinforcement, i.e. not only data processing, but also with additional testing or selection of different strategies for achieving maximum results, etc.

It stands to mention that the main efforts of scientists are focused on technical and legal restrictions on AI. Thus, the expert Mady Delvaux in the Report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103 (INL)) [5] emphasized that technological advances in robotics stimulate the implementation by robots of certain autonomous and cognitive functions unique to humans, such as the ability learn from experience, to make independent decisions. However, this can encourage robots to destructive actions that have harmful consequences and for which, by human analogy, the legal responsibility must be incurred. At the same time, manufacturers, owners, software developers, users, military commanders, etc. are still responsible for the actions of non-autonomous or partially autonomous robots. Until now, the only legal act that defines the liability for damage caused by works is Council Directive 85/374 / EEC «On the Approximation of Laws, Regulations and Administrative Provisions of the Member States on the Liability for Unsafe Products» [6].

The European Parliamentary Research Service (EPRS) and the European Parliament Legal Affairs Committee are developing recommendations to improve the civil and ethical aspects of robotics, propose setting up a register of robots and the appropriate EU Agency for Robotics, as well as the establishment of civil liability for damages caused by the robots in accordance with the actual level of the instructions given according to the degree of autonomy. It is planned to carry out technical, ethical and regulatory expertise in the field of robotics and develop a «Code of Ethical Conduct for Robotics Engineers» and a «Code of Ethics Committees for Research» [7].

In September 2016, the British Standards Institution issued the standard BS 8611: 2016 «Robots and robotic devices. Guide to the ethical design and application of robots and robotic systems»[8]. Thus, in 2019, the Expert Group on Responsibility and New Technologies for New Technologies, established by the European Commission, in the statement "Responsibility for Artificial Intelligence and Other Digital Technologies" stated[' that the responsibility regimes for AI and IoT implementation should provide sufficient protective measures to minimize the risks of damage that these technologies can cause [9].

In addition, in 2016, the Executive Office of the President of the National Science and Technology Council Washington, D.C. 20502, the Technology Committee has prepared a review of «Preparing for the future of artificial intelligence» on the prospects for the use of AI. The document states that in the short term, the use of AI is the automation of tasks that previously could not be automated. Concerns

have also been expressed about the unintended consequences of the use of AI, as it is considered unacceptable to use certain AI predictions regarding human development and at the same time to ignore the decisions that are taken by people on the same issues. In addition, the use of artificial intelligence to control the equipment of the physical world needs to be regulated, especially when it is directly related to human security [10].

In 2017, the European Parliament published its set of general recommendations on how to implement ethics in robots, as well as proposals to define their legal status as an «electronic person with special rights and responsibilities». Google company presents «Perspectives on Issues in AI Governance», which covers the general problems of using artificial intelligence at the conceptual level [11].

In early 2022, Stanford University published its annual Artificial Intelligence Index Report 2022, which includes academic, private, and nonprofit research, a survey of robotics researchers from around the world, global artificial intelligence legislation in 25 countries, and a section on in-depth analysis of technical indicators of AI ethics.

An analysis of the legislative activity in 25 countries shows that the number of bills containing AI and the number of adopted laws increased from 1 in 2016 to 18 in 2021. Spain, the United Kingdom and the United States have passed the largest number of bills related to artificial intelligence.

In our opinion, the fact that robotic armaments are becoming cheaper and more affordable for mass use is quite worrying, namely that the average price of robotic weapons has decreased four over the past six years from $ 50,000 per unit in 2016 up to $ 12,845 in 2021 [12].

Therefore, technological improvements can provide greater accuracy in the use of ammunition and greater «humanity» of military operations. It is high-precision ammunition that allows the completion of the war with less waste of ammunition and casualties while remotely piloted vehicles can provide reduced risks for to troops.

At the same time, it should be noted that scientists do not know for sure what AI algorithms are currently being developed and applied in the military sphere. It is considered very risky avoiding the direct human control over the controlled and autonomous weapons systems.

Also in November 2012, US Department of Defense Directive No. 3000.09 «Autonomy in Weapon Systems» came into force, which establishes the Department of Defense's policy and legal responsibility for the development and use of autonomous and semi-autonomous functions in weapons systems, including manned and unmanned platforms and establishes recommendations designed to minimize the likelihood and consequences of failures in autonomous and semi-autonomous weapon systems that could lead to unintentional losses [13].

According to some publications and information reports, we can state that the use of AI is currently being implemented in some types of autonomous weapons. For example, Heron Systems successfully conducted virtual air battles between artificial intelligence and an F-16 fighter pilot of the US Air Force's F-16 fighter jet at the US Department of Defense's Advanced Research Projects Test Site (DARPA) with a complete overpowering of artificial intelligence aircraft [14].

Recently, Collaborations Pharmaceuticals, a private pharmaceutical company, changed several commands in an artificial intelligence algorithm for research purposes to speed up the MegaSyn molecule generator to eliminate toxic reactions during searching for new drugs. Instead, the MegaSyn molecule generator has specifically revealed 40 thousand variants of potential biological and chemical weapons based on well-known warfare weapons, including «Novichok». Many new toxic molecules have also been developed that are potentially even more harmless than those already known [15].

Moral AI and the requirements for its ethical application can become declarative and unrealistic if dictatorial regimes use AI technology. According to the facts of the current war launched by Russia against democratic Ukraine, the vast majority of conventions, international regulations and ethical norms can be brutally ignored. Therefore, international organizations are becoming powerless against violations of the world order and decades of legal and moral rules of coexistence, because in this case, the norms of the rule of law will be enforced. Russia barbarically destroys civilian objects, the entire city of Ukraine is almost in the center of Europe and does not pay attention to the international efforts of the global world community. In fact, Russia today is gaining the status of an aggressive enabler, guided only by its own ideas about the world order. In such a situation, the use of AI in the military sphere can occur not only uncontrolled but also carry out the tasks of destroying people. Russia's aggression against Ukraine violates many issues of the future structure of society. It is in countries with dictatorial regimes that one such issue is the possibility of the emergence and existence of aggressive autonomous AI. Imagine the possibility of autonomous aggressive AI in the near future, taking into account the development of digital humanity in the metaverse.

We propose to consider the hypothesis that aggressive autonomous AI with a certain degree of probability may exist. In our opinion, it is unlikely that it will look like a supercomputer created and operating in a certain place. The most rationally

autonomous AI will look like a group of independent computing clusters, such as bot-nets distributed in space, which are configured to perform specific tasks-algorithms. Cluster groups may be grouped according to the OSI model, in which each level is characterized by its own set of AI clusters. Several clusters will undoubtedly be formed for strategic decisions, as well as for the use of mechanical and biotechnical android-like robots. This is the most efficient and least vulnerable configuration to external factors.

What is the purpose and tasks of aggressive autonomous AI? It is believed that the autonomous AI will use data from the historical development of mankind to form the goal and measures to achieve it, because other data on the development of civilization, such as extraterrestrial, does not exist. Unfortunately, the history of mankind, and especially in recent years, is full of negative examples of hostility and aggression. Numerous military simulators, military AI, electronic tools for analytical assessment of past and current military conflicts actually create a modern library on the basis of which aggressive autonomous AI can independently formulate strategies and tactics for destructive actions against humanity.

Thus, to choose a strategy to attack humanity, autonomous AI has a fairly broad base of both historical examples of military action and today's war. Therefore, the strategy, tactics and methods of warfare against humanity are not a big mystery. There are many wars and military conflicts in the history of mankind. There are enough reasons why it arose, but in general they are threefold: a) the capture of territory; b) the capture of resources: c) capture of labour and reproductive power. It can be assumed that for aggressive autonomous AI these bases will also be basic. Items «a» and «b» are likely to be combined in the capture of large cities and megapolises, because there are huge resources that are concentrated in a small area: ferrous and nonferrous metals, data transmission infrastructure, computers, electronic gadgets, data centers, servers, IT devices and many other things necessary to support various clusters, including mechanical and biotechnical android-like robots. Therefore, an autonomous AI does not need to physically produce crude digging minerals from the bowels of the ground and create a full production cycle. Autonomous AI does not need to create an army of robots of 5 billion units. Effective autonomous robotic weapons under the control of AI can be many times more effective than any army by the compact forces required to maintain the robot ecosystem in proper condition. However, the autonomous AI model has one vulnerable disadvantage it is the energy supply. Destruction of AI energy infrastructure is the destruction of AI clusters or the autonomous AI itself. Item «c» will

be ignored, as the author believes that in the near future humanity will take decisive steps to control the development and use of all types and forms of AI, which will allow preventing its use for the sake of humanity. The author also hopes that Russia's aggression against democratic Ukraine will force humanity to create strong international, rapid and effective mechanisms to block and destroy all forms and methods of using AI to create uncontrolled autonomous weapons and use AI in the military to aggression against the world and humanity in general.

Is there a possible scenario in which an aggressive autonomous AI will attack humanity? Well, in general, such a scenario should be considered at least in order to prevent it from happening in proper time.

Conclusions. Taking into account the main directions of development of technologies with artificial intelligence, we can assert that today there is a tendency to reasonably assess the risks of using AI in various spheres of human activity. The moral and ethical problems of AI are increasingly being raised at all stages from the development of technical tasks to practical use. Today, legislative activity is focused on creation legal rules and barriers of uncontrolled distribution of AI. However, the state of penetration of AI developments and autonomous armaments in military defense agencies, their purpose, readiness for use, number and military potential remain unknown. It is also unknown whether the military defense agencies are selecting such developments from the AI on the basis of autonomy and causing damage to the enemy or civilian infrastructure. Besides, it is also unknown whether regulations are being developed to restrict the use of AI directly in the military sphere.

Considering the above, we believe it appropriate to establish an international supervisory and controlling body, which will be responsible for supervision and control over the use of moral AI and dessemination of autonomous AI for military or dual purpose. Introduce a single international automated system for certification and licensing of developments with artificial intelligence algorithms. It is advisable to create a single international automated system for capturing and responding to incidents in the systems with artificial intelligence algorithms. Initiate the formation of an international standard Model of technical, biological, financial, economic, political and military threats to the use of (sanctioned and non-sanctioned) AI-based systems [16].

At the same time, it is also necessary to intensify the work of international organizations to develop and create a Model Law «On Artificial Intelligence», which will form the basis of relevant national legislation. In turn, scientists need to form a unified categorical and conceptual apparatus in the sphere

of AI in the shortest terms and ensure its maximum dissemination for simultaneous application in the legislative jurisdictions of different states.

REFERENCES:

1. Kostenko O. V. 2022. Electronic Jurisdiction, Metaverse, Artificial Intelligence, Digital Personality, Digital Avatar, Neural Networks: Theory, Practice, Perspective. World Science. 1(73). DOI: https://doi.org/10.31435/rsglobal_ws/30012022/7751.

2. Wallach, W., Allen, C. 2005. Android Ethics : Bottom-up and Top-down Approaches for Modeling Human Moral Faculties. https://www.semanticscholar.org/paper/Android-Ethics-%3A-Bottom-up-and-Top-down-Approaches-Wallach-Allen/a65e0e2c951c447f827107786c4b6deeb2b71226.

3. Еталонна модель взаємодії відкритих систем (EMBBC, OSI – Open System Interconnect). Рівневі протоколи. URL: http://osvita-plaza.in.ua/publ/45-1-0-435. (Last accessed: 02.04.2022).

4. Anderson, M., Anderson, S.L. 2008. ETHEL: Toward a Principled Ethical Eldercare Robot. URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.177.5971.

5. Mady Delvaux. REPORT with recommendations to the Commission on Civil Law Rules on Robotics URL: https://www.europarl.europa.eu/doceo/document/A-8-2017-0005_EN.html.

6. Cost of non-Europe in robotics and artificial intelligence. EPRS. URL: http://www.europarl.europa.eu/committees/en/juri/robotics.html?Tab=Introduction.

7. Civil law rules on robotics – European Parliament. URL: http://www.europarl.europa.eu/RegData/etudes/ATAG/2017/599250/EPRS_ATA(2017)599250_EN.pdf.

8. BS 8611:2016 Robots and robotic devices. Guide to the ethical design and application of robots and robotic systems. URL: https://www.en-standard.eu/bs-8611-2016-robots-and-robotic-devices-guide-to-the-ethical-design-and-application-of-robots-and-robotic-systems/

9. Report from the Expert Group on Liability and New Technologies - New Technologies Formation. Liability for artificial intelligence and other emerging digital technologies. Luxembourg Publications Office of the European Union, 2019. URL: https://op.europa.eu/en/publication-detail/-/publication/1c5e30be-1197-11ea-8c1f-01aa75ed71a1/language-en. https://data.europa.eu/doi/10.2838/25362.

10. Preparing for the future of artificial intelligence. Executive office of the President National Science and Technology Council Washington, D.C. October. 12. 2016. URL: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

11. Google. Perspectives on Issues in AI Governance. URL: https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf.

12. Artificial Intelligence Index Report 2022. Stanford HAI. URL: https://aiindex.stanford.edu/wp-content/uploads/2022/03/2022-AI-Index-Report_Master.pdf.

13. Department of Defense USA Directive Number 3000.09 Autonomy in Weapon Systems. URL: https://www.esd.whs.mil/portals/54/documents/dd/issuances/dodd/300009p.pdf.

14. V Mynysterstve oborony SShA provely vyrtualnыi boi yskusstvennoho yntellekta s pylotom ystrebytelia F-16. Alhorytm uspыtыval maksymalnыe vozmozhnosty cheloveka y samoleta. URL: https://techno.nv.ua/innovations/algoritm-pobedil-pilota-vvs-v-boyu-50108456.html.

15. Shtuchnyi intelekt dlia poshuku likiv vyhadav 40 tysiach variantiv khimichnoi zbroi. URL: https://tech.24tv.ua/shtuchniy-intelekt-dlya-poshuku-likiv-vigadav-40-tisyach-variantiv_n1915202.