Assessment of risks and threats to the development of artificial intelligence technologies

Elena Shevchenko*, *Angelika* Efremova, *Dmitriy* Titarenko, *Aleksey* Voloshin, and *Inga* Artyukhova

V.I. Vernadsky Crimean Federal University, Simferopol, Russian Federation

Abstract. The paper analyses the risks and threats caused by the development and implementation of artificial intelligence technologies. We believe that these risks and threats need research in the long term. Importantly, the use of intelligent information systems has a twofold effect: it can lead to both positive and negative results. The paper also considers the influence of artificial intelligence technologies on the various activities. It proposes a classification of risks and threats caused by the development and implementation of artificial intelligence technologies by the main spheres of human activity.

1 Introduction

Artificial intelligence as a research area emerged in the 1950s thanks to the development of information and communication technologies and computer machines. This made scientists confident that it was possible to create artificial intelligence. Now the prospects for creating artificial intelligence seem to be achievable for both scientists and practitioners. Computer science and information technology are developing rapidly, with discoveries every day in different parts of the world.

As an example, Alan Turing was the first to publish a research paper on artificial intelligence in 1950, entitled "Computing Machinery and Intelligence" and looking into the problem of the machine's ability to think. Later, in 1997, IBM's Deep Blue chess supercomputer won a six-game match against the world chess champion Garry Kasparov. Then, in 2016, Google's AlphaGo beat Go's absolute champion Lee Sedol. As you know, this game belongs to the category of very complex, highly intelligent games. The technology of deep learning allowed achieving this victory [1].

Currently, in artificial intelligence, several scientific areas are forming and developing, such as knowledge representation, automation of reasoning, knowledge acquisition, machine learning and automatic hypothesis generation, data mining and processing of image information, multi-agent systems, dynamic intelligent systems and planning, natural language processing, fuzzy models and soft computing, and development of tools [2].

---

* Corresponding author: shevchenko_lena_@list.ru

Artificial intelligence technologies have found their application in various information systems, known as intelligent information systems. The use of intelligent information systems has a twofold effect: on the one hand, it gives significant economic benefits to those who use it, and on the other hand, it can bring significant negative aspects for society as a whole, various risks and threats. Moreover, in the future, these negative aspects may exceed the total positive effect of artificial intelligence technologies.

We need to think about the risks and threats of artificial intelligence right now and design measures to overcome and prevent them so that in the future we do not face a crisis without a possible change for the better. This problem is complex and needs a system analysis methodology.

Like any discovery, artificial intelligence can be used for different purposes. For example, dynamite was created for mining, and no one thought that it could find itself inhumane use.

The purpose of the article is to analyze and classify possible risks and threats of the development of artificial intelligence technologies in various spheres of human activity.

2 Materials and methods

The study is based on the specific scientific literature on artificial intelligence. Using the methods of analysis and synthesis, we have designed a classification of risks and threats of the development and implementation of artificial intelligence technologies.

We have grouped the risks and threats of the development and implementation of artificial intelligence technologies according to the main spheres of human activity: economic, social, political, cultural, spiritual, and military.

3 Results

Currently, a national economy's competitiveness is impossible without modern information technologies. In its turn, artificial intelligence is the crown of information technology development. Economies will have to add capacities in this direction as well as in the military sector since lag in this area threatens with economic lag in the future.

The use of artificial intelligence technologies leads to a significant positive economic effect: it reduces by 30% the time spent on the formation of warehouse stocks, leads to a 20% increase in electricity production, speeds up the delivery of materials in industrial production by 30%, increases the productivity of the healthcare sector by 30-50% [3]. We can provide many more such facts.

Besides, large corporations (in particular, transnational companies) are currently interested in artificial intelligence technologies, which, as you know, can significantly influence many areas of human activity, including the economy and politics.

There are the following groups of subjects in the global market for artificial intelligence [3]:

— companies that develop artificial intelligence technologies and have their own data (for example, Google, Microsoft);

— companies whose activities are related to the processing of large amounts of data (for example, IBM, Oracle);

— innovative companies that solve specific problems but do not have their own data (for example, Two Sigma Investments).

E. N. Smirnov and S. A. Lukyanov point out that "artificial intelligence development can not only modify business but also modify the competition in the world market, contribute to further differentiation of the countries in the world by levels of economic and techno-

logical development ...” [3]. Moreover, in the current hyperglobalization, the competitive advantages of individual countries, markets, and companies can be lost as quickly as acquired [3].

The development and implementation of artificial intelligence technologies can give rise to significant risks and threats in various spheres of human activity. As K. M. Manheim and L. Kaplan [4] note, artificial intelligence is just a technology with many possible applications. “Rather it is how the tool is used, by whom, and for what purpose that generate concern” [4].

Table 1 shows the classification of risks and threats of the development and implementation of artificial intelligence technologies. We have grouped these risks and threats by the main spheres of human activity: economic, social, political, cultural, spiritual, military, and informational.

We have identified several risks and threats that can belong to all basic areas of human activity:

— the emergence of complete dependence on computers. Any serious failure in such a system can lead to its full stop, and in some cases to destruction;

— the unpredictability of self-conscious systems (intelligent robots), and the influence of various accidents on them. A system endowed with artificial intelligence can face failures and unforeseen factors, which can change the behavior of such a system and lead to the loss of control;

— errors and failures in the work of intelligent information systems. Any information system is created and programmed by people, who can make various kinds of errors in the process of its creation, therefore failures are possible during the operation of these systems;

Table 1. Classification of risks and threats of the development and implementation of artificial intelligence technologies.

Sphere of human activity	Risks and threats of the development and implementation of artificial intelligence technologies	
Economic	Malicious and criminal use of artificial intelligence to obtain various material benefits. Financial risks arising from the use of artificial intelligence technologies. Differentiation of countries by development level.	The emergence of complete dependence on computers. The unpredictability of self-conscious systems (intelligent robots), and the influence of various accidents on them. Errors and failures in the work of intelligent information systems. Unreliable, unsafe, or unsatisfactory results. Threats to information confidentiality. The impossibility of bringing to justice the parties responsible for negative consequences.
Social	Unemployment. Growing poverty and the emergence of a strong economic stratification of society (increased marginalization and inequality). Increased discrimination. Physical security threat.	
Political	Loss of control over the national economy. Loss of sovereignty. A decline in democracy. Manipulation of political systems.	
Cultural	The disunity of people (disintegration of social ties). Loss of natural human skills.	

Table 1. Continued.

Spiritual	Loss of spirituality. Moral degradation.	The creation of an artificial superintelligence that will surpass humans and destroy the human civilization.
Military	The threat to national security. The use of combat robots can lead to additional human losses, among civilians as well.	
Informational	Fake news reports with realistic fabricated video and audio. Automated, hyper-personalized disinformation campaigns. Automating influence campaigns. Denial-of-information attacks. Manipulation of information availability.	

[a] Formed by the authors based on [1; 3–11].

— unreliable, unsafe, or unsatisfactory results. "Irresponsible data management, negligent design and production processes, and questionable deployment practices can, each in their own ways, lead to the implementation and distribution of AI systems that produce unreliable, unsafe, or poor-quality outcomes" [9];

— threats to information confidentiality. Global digitalization of all spheres of human activity can lead to the loss of confidentiality and information leaks, as well as to complete control over a human personality;

— the impossibility of bringing to justice the parties responsible for negative consequences;

— the creation of an artificial superintelligence that will surpass humans and destroy the human civilization (the darkest scenario). However, the current state of science and technology does not yet allow this scenario to come true. Still, we should keep this option in mind due to the rapid development of technology. See, for example, [11].

Below we shall consider the risks and threats of the development and implementation of artificial intelligence technologies that belong to specific areas of human life.

Economy. Artificial intelligence technologies may have criminal uses: cybercrime, particularly in finance; terrorism, including an international one. Here we should take into account not only the current level of technology development but also the future one. These technologies may fall into the hands of dishonest people. As D. Yu. Bazarkina and Y. N. Pashentsev note, "readiness to meet unexpected hidden risks is crucial" [5]. There are also numerous risks of using artificial intelligence technologies in the financial sector. Tom C.W. Lin underlines that "… the proliferation of artificial intelligence in finance increases the dangers of systemic risks and major financial accidents" [10]. Moreover, the development of artificial intelligence technologies may result in further differentiation of countries by development level.

Artificial intelligence technologies can pose significant risks and threats in the social sphere. Among them, we would like to highlight an increase in unemployment and poverty, a strong economic stratification of the society (worsened marginalization and inequality), growing discrimination, and a threat to physical security.

In politics, we can list the following risks and threats: loss of control over the national economy, loss of sovereignty, a decline in democracy, and manipulation of political systems.

As K. M. Manheim and L. Kaplan note, "Artificial intelligence could erase many practical advantages of democracy, and erode the ideals of liberty and equality. It will further concentrate power among a small elite if we don't take steps to stop it" [4].

In culture, artificial intelligence technologies can contribute to the disunity of people (disintegration of social ties), and the loss of natural human skills. "Excessive automation, for example, might reduce the need for human-to-human interaction, while algorithmically enabled hyper-personalization, by limiting our exposure to worldviews different from ours, might polarize social relationships" [9]. "Humans may gradually become passive elements of an emerging global socio-technical system – a system composed of machines, algorithms, sensors and actuators, AI programs, and humans interacting in the globally present Internet, and the Internet that is ever present due to mobile technologies and ambient intelligence" [11].

In the spiritual realm, there can be a loss of spirituality and moral degradation.

The development of artificial intelligence can lead to significant and, in some cases, even catastrophic risks in the military sphere: a threat to national security (artificial intelligence used against any sovereign state); constantly improving unmanned weapon systems based on artificial intelligence can have not only defense but also offense applications; the use of combat robots can lead to additional human losses, including among civilians.

In the information sphere, we can distinguish the following risks and threats: fake news reports with realistic fabricated video and audio; automated, hyper-personalized disinformation campaigns; automating influence campaigns; denial-of-information attacks; manipulation of information availability [6].

Existing and future threats associated with artificial intelligence technologies require measures to overcome and prevent them. The paper [6] proposes the following measures:

— closer cooperation between the government and the designers of artificial intelligence to prevent the possible illegal use of artificial intelligence technologies;

— socially responsible behavior of artificial intelligence designers; they must be constantly aware of possible risks and threats;

— application of the world's best security practices (information, economic, social, etc.);

— involvement of the best scientists and experts.

Dr. David Leslie talks about the need of artificial intelligence ethics, i.e. "a set of values, principles, and techniques that employ widely accepted standards of right and wrong to guide moral conduct in the development and use of AI technologies" [9].

B. Cheatham, K. Javanmardian and H. Samandari also advocate the need to manage the risks of artificial intelligence technologies at the company (firm) level. Besides, they suggest the following principles [7]:

— using a comprehensive approach to risk identification;

— introducing reliable risk management tools;

— strengthening control over risks.

In investment decision-making in the field of artificial intelligence, it is necessary to adhere to the ethical principle. As shown in [12; 13], companies' compliance with ethical standards in business operations helps to increase their competitiveness and has a positive effect on national competitiveness.

4 Discussion

A very important question is: can this problem be solved without government intervention? In our opinion, the regulatory role of the state in the field of artificial intelligence is very important since, in the pursuit of profit, companies can ignore aspects of security, including the future one. This, in particular, is stated in [14].

5 Conclusion

The risks and threats caused by the development and implementation of artificial intelligence technologies can group by the main areas of human activity: economic, social, political, cultural, spiritual, military, and informational. It is also possible to single out several risks and threats related to all of the above areas: the emergence of complete dependence on computers; the unpredictability of self-conscious systems; errors and failures in the work of intelligent information systems; unreliable, unsafe, or unsatisfactory results; threats to information confidentiality, etc.

Preventing risks and threats caused by the development of artificial intelligence technologies will require efforts not only by the state but also by businesses, representatives of science, and society as a whole. The state should adopt special laws to ensure human safety and rights, and companies should conduct socially responsible policies and comply with the ethical principle in their activities.

References

1. E. M. Proydakov, Digital Economy, **3(3)**, 50 (2018)

2. G. S. Osipov, *Artificial Intelligence Lectures* (LENAND, Moscow, 2018)

3. E. N. Smirnov, S. A. Lukyanov, Region Economy, **15(1)**, 57 (2019)

4. K. M. Manheim, L. Kaplan, Yale J. of Law and Technology, **21**, 106 (2019)

5. D. Yu. Bazarkina, Y. N. Pashentsev, Russia in Global Affairs, **17(1)**, 147 (2019)

6. M. Brundage, et al., *The malicious use of artificial intelligence: forecasting, prevention, and mitigation* (2018)

7. B. Cheatham, K. Javanmardian, H. Samandari, *Confronting the risks of artificial intelligence* (McKinsey Quarterly, 2019)

8. J. Johnson, *Artificial intelligence & future warfare: implications for international security. Defense & Security Analysis* (2019)

9. D. Leslie, *Understanding Artificial Intelligence Ethics and Safety: a Guide for the Responsible Design and Implementation of AI Systems in the Public Sector. The Alan Turing Institute* (2019)

10. D. Lin, Fordham Law Review, **88** (2019)

11. A. Nowak, P. Lukowicz, P. Horodecki, "Assessing artificial intelligence for humanity. will AI be our biggest ever advance — or the biggest threat?" in *IEEE Technology and Society Magazine* (2018)

12. M. Bakumenko and A. Sigal, Int. J. of Risk Assessment and Management, **23(1)**, 80 (2020)

13. M. Bakumenko, A. Sigal, Int. J. of Risk Assessment and Management, **21(1/2)**, 111 (2018)

14. C. Wheelus, X. Zhu, IoT, **1**, 259 (2020)