Problems of ethical regulation of the use of artificial intelligence technologies

Aleksey Vladimirovich Minbaleev[2][1], *Tatyana Anatolyevna* Polyakova[1] and *Oleg Sergeevich* Makarov[3]

[1]Kutafin Moscow State Law University, Department of Information law and Digital Technologies, Moscow, Russia
[2]Institute of State and Law of the Russian Academy of Sciences, Information Law and International Information Security Sector, Moscow, Russia
[3]Belarusian Institute for Strategic Studies, Minsk, Republic of Belarus

Abstract. The purpose of the article is to analyze the problem of ethical regulation of artificial intelligence technologies in a modern complex socio-cultural environment and active digital transformation of society. In the world today, the practice of private ethical initiatives is very common, associated with the joining of efforts of a number of companies to develop the use of artificial intelligence technologies and consolidate the basic principles of working with such technologies. The research used such scientific methods as dialectical, logical, historical, predictive, systems analysis, content analysis, as well as private scientific methods, such as comparative legal and the method of legal modeling. The study results in proposals for introducing ethical regulators into the general mechanism for regulating digital technologies, obtained on the basis of an analysis of the main directions and approaches to ethical regulation, as well as identifying key problems of ethical regulation in the field of artificial intelligence that have to be solved by society and the state. The authors support the position of researchers and governments of a number of states that normative regulation should take precedence over ethical, the tasks of any state should be to provide normative support for the use of artificial intelligence technologies and protect the rights of citizens. The novelty of the work lies in the conclusion that at the level of the Government of the Russian Federation today it is advisable to develop Model Rules for the use of artificial intelligence technologies in government and local authorities, as well as in their subordinate organizations. The rules can enshrine ethical principles and norms based on the provisions of acts of international organizations and recommendations of the professional expert community. The rules can also reflect recommendations for their use by non-state legal entities, individual citizens. It is concluded that such an approach would eliminate the possibility of numerous ethical codes and regulations from both the public and commercial sectors and would not be redundant.
Keywords: artificial intelligence, ethical regulation, digital law, mechanisms for regulating digital technologies

---

[1] Corresponding author: alexmin@bk.ru

1   Introduction

One of the key problems in the regulation of artificial intelligence technologies (hereinafter referred to as AI) is the problem of the possibility and limits of ethical regulation of these relations. The possibility of ethical regulation of digital technologies, including AI, is being discussed at various venues in the world and in Russia [1]. Many believe that ethical norms are much more perceived in the technical sphere than legal ones, and are more accessible [2-4].

Indeed, ethical standards are much more quickly adopted, implemented and used in the circulation and use of AI technologies. At the same time, the truth is that a certain organizing principle is lost in the management of relations in the field of AI on the part of the state, which should always be. The way out of this situation can be the initiation of ethical regulation by the state.

In recent years, Canada, China, Denmark, the EU Commission, Finland, France, India, Italy, Japan, Mexico, the Nordic-Baltic region, Singapore, South Korea, Sweden, Taiwan, the United Arab Emirates, and Great Britain have issued national strategies in the field of AI and the promotion of AI technologies. However, each of them focuses on different aspects of AI policy: research, talent development, skills and education, public-private interaction, ethics and coexistence, standards and regulations, and data and digital infrastructure [5].

2   Methods

The research used such scientific methods as dialectical, logical, historical, predictive, systems analysis, content analysis, as well as private scientific methods, such as comparative legal and the method of legal modeling.

The method of systems analysis, as well as prognostic methods, allow us to identify the main directions and patterns in the field of ethical regulation and allow us to determine the prospects for its development. The method of legal modeling, used in the context of the rapid development of the digital economy, outstripping the development of legislation, makes it possible to develop the main mechanisms, methods, approaches to ethical regulation of the use of AI, as well as to identify the main factors and conditions affecting the development of new mechanisms for the regulation of AI-based on a combination of ethical and other regulators. Epistemological, prognostic, and axiological methods made it possible to identify the features, tendencies and conditions for the formation of new scientific directions related to the ethical regulation of AI.

All these methods will make it possible to effectively ensure the implementation of the research goal – to identify the main problems of ethical regulation of the use of AI technologies.

3   Results and discussion

The National Strategy for the Development of AI for the Period up to 2030 states that "to stimulate the development and use of AI technologies, it is necessary to adapt the regulatory framework in terms of human interaction with AI, and develop appropriate ethical standards" [6]. Thus, the President of the Russian Federation has set the task of developing both normative and ethical regulations of the use of AI technologies.

Various governments and nongovernmental organizations, international organizations have today put forward principles, frameworks and guidelines for the ethics and governance

of AI. First, this was due to the formulation of voluntary guidelines for those who develop or implement AI systems. For example, one of the latest newest principles was adopted by the Government of Singapore in the second edition of its AI Governance Model. This framework "provides practical guidance for private sector actors on key ethical and governance issues when deploying AI solutions" [7, p. 4]. It is based on two main principles: AI decision-making processes must be explainable, transparent and fair, and AI decisions must be human-centered [8].

One of the world's most recent analyzes of the ethical regulation of AI by the Law Reform Committee of the Singapore Law Academy, the principles include law and core interests; consideration of the effects of AI systems; respect for values and culture; management of risks; well-being and safety; accountability; transparency; ethical use of data [7, p. 32].

The analysis of these principles indicates that the world uses standard approaches to the construction of its system. Their differentiation is carried out mainly depending on the scope of use of AI technologies [9, p. 188]. Thus, specific principles are clearly distinguished in the field of health care, education, transport, energy, including nuclear, and in other areas, especially those subject to compulsory licensing and certification.

The problems of ethical regulation in the field of AI give rise to a significant number of questions that have to be solved by society and the state:

- the growth of AI undoubtedly opens up incredible opportunities for states, individuals and society as a whole, but it is fraught with serious ethical risks [10, 11]. For example, Minter Ellison's 2019 Cyber Risk Outlook report examines some of the cyber risks associated with AI and recommends actions that organizations can take to mitigate that risk;

- problems of unemployment and social stratification as a consequence of robotic automation;

- problems of replacing human AI in the process of his own life (4, 12, 13);

- the problem of AI superpowers.

Societies today also face ethical issues such as collecting data without proper consent, the confidentiality of personal data, inherent selection bias, the risk of profiling and discrimination, and the opaque nature of some AI solutions. It is also important to talk about the reputational concerns of the public fear that corporations are massively exploiting and misusing huge amounts of consumer data to gain insight into consumers and to provide an unfair digital competitive advantage. One of the clear systemic risks of AI is the problem or phenomenon of the "black box" [14]. Almost all branches of law pose the problem of automated solutions and management of the results of the functioning of AI technologies, since it is very often impossible to predict the result of its functioning unambiguously.

AI systems around the world are falling under regulatory scrutiny for ethical violations. For example, in the United States, a number of companies are under special scrutiny for allegedly developing an algorithm that encouraged doctors and nurses to pay more attention to white patients than black patients. Thus, in a special study [15], it was revealed that an algorithm was developed to help health systems target patients who will have the greatest future health care needs, predicting how likely it is that people will use a lot of medical services and accumulate high costs in the future.

Such systems of inequality of people, given by AI algorithms, are the most painful for the healthcare sector, but they are far from limited to it. Thus, credit rating algorithms based on a social scoring system [16] predict income-related outcomes, thus including differences in employment and wages [17]. Policing algorithms predict measurable crime, which also reflects the increased focus on certain groups [18]. Hiring algorithms predict hiring decisions or supervisory ratings that are influenced by racial and gender bias [19]. Even

retail algorithms that set prices for goods at the national level punish poorer households, which are subject to higher prices as a result.

All these facts testify to the constant violations of human rights in a number of countries of the world, which are implemented using AI technologies and cause a number of ethical discussions.

4 Conclusion

Today, most researchers of the problems of regulating digital relations note the need to develop a complex application of various regulators – legal, ethical, self-regulation, technical [4, 20, 21]. At the same time, the position of researchers [14] and the governments of a number of states (for example, Australia) should be supported that regulatory regulation should take precedence over ethical, the tasks of any state should be to ensure the use of AI technologies and protect the rights of citizens.

In this regard, at the level of the Government of the Russian Federation today it is advisable to develop Model Rules for the use of AI technologies in government and local authorities, as well as in their subordinate organizations. The rules can enshrine ethical principles and norms based on the provisions of acts of international organizations and recommendations of the professional expert community. The rules can also reflect recommendations for their use by non-state legal entities, individual entrepreneurs and individuals when using AI technologies. We believe that such an approach would eliminate the possibility of numerous ethical codes and regulations from both the public and commercial sectors and would not be redundant.

References

1. Ethics and Digital: Ethical Issues of Digital Technologies. Analytical report. Accessed on: February 20, 2021. [Online]. Available: http://ethics.cdto.center

2. A.V. Neznamov, V.B. Naumov, Legal Research, **8**, 14-25 (2017)

3. K. Nikolskaia, V. Naumov, *Artificial Intelligence in Law*, in 2020 International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon), Vladivostok, Russia, 1-4 (2020). https://doi.org/10.1109/FarEastCon50210.2020.9271095

4. V.B. Naumov (ed.), Pravovye i eticheskie aspekty, svyazannye s razrabotkoi i primeneniem sistem iskusstvennogo intellekta i robototekhniki: istoriya, sovremennoe sostoyanie i perspektivy razvitiya: monografiya [Legal and ethical aspects related to the development and application of artificial intelligence systems and robotics: history, current state and development prospects: monograph] (NP-Print, Saint Petersburg, 2020)

5. T. Dutton An Overview of National AI Strategies. [Online]. Available: https://medium.com/politics-ai/an-overview-of-national-ai-strategies-2a70ec6edfd.

6. National strategy for the development of artificial intelligence for the period up to 2030. Approved by the Decree of the President of the Russian Federation dated 10.10.2019, Collected Acts of RF, **41**, Art. 5700 (2019)

7. Applying Ethical Principles for Artificial Intelligence in Regulatory Reform, Part of the Impact of Robotics and Artificial Intelligence on the Law series (Singapore Academy of Law, Singapore, 2020)

8. Model Artificial Intelligence Governance Framework, Second Edition (2020). Accessed on: February 10, 2021. [Online]. Available: https://www.pdpc.gov.sg/-/media/Files/PDPC/PDFFiles/Resource-for-Organisation/AI/ SGModelAIGovFramework2.pdf

9. V.V. Blazheeva, M.A. Egorova, Digital Law: textbook (Prospect, Moscow, 2020)

10. I. Kaftannikov, V. Zhernova, A. Minbaleev (2019). Problems of structuring risks and ensuring legal relations in IoT, in Proceedings of the 1st International Scientific Conference "Modern Management Trends and the Digital Economy: from Regional Development to Global Economic Growth" (MTDE 2019) (2019). https://doi.org/10.2991/mtde-19.2019.14

11. K.Y. Nikolskaya, S.A. Ivanov, V.A. Golodov, A.V. Minbaleevand, G.D. Asyaev, *Review of modern DDoS-attacks, method sand means of counteraction*, in 2017 International Conference "Quality Management, Transport and Information Security, Information Technologies" (IT&QM&IS), Saint Petersburg, 87-89 (2017). https://doi.org/10.1109/ITMQIS.2017.8085769

12. K. Nikolskaia, V. Naumov, *Ethical and Legal Principles of Publishing Open Source Dual-Purpose Machine Learning Algorithms*, in 2020 International Conference Quality Management, Transport and Information Security, Information Technologies (IT&QM&IS), Yaroslavl, Russia, 56-58 (2020). https://doi.org/10.1109/ITQMIS51053.2020.9322897

13. D.V. Bakhteev, Iskusstvennyi intellekt: etiko-pravovye osnovy: monografiya [Artificial intelligence: ethical and legal foundations: monograph] (Prospect, Moscow, 2021)

14. J. Black, A.D. Murray, European Journal of Law and Technology, **10(3)** (2019). Accessed on: February 10, 2021. [Online]. Available: http://eprints.lse.ac.uk/102953/4/722_3282_1_PB.pdf

15. Z. Obermeyer, B. Powers, Ch. Vogeli, S. Mullainathan, Science, **366(6464)**, 447-453 (2019). https://doi.org/10.1126/science.aax2342.

16. A.A. Lazarov, Information Law, **4**, 41-44 (2020)

17. S. Barocas, A.D. California Law Review, **104(3)**, 671-732 (2016). Accessed on: February 10, 2021. [Online]. Available: www.jstor.org/stable/24758720

18. K. Lum, W. Isaac, Significance, **13**, 14-19 (2016). https://doi.org/10.1111/j.1740-9713.2016.00960.x

19. I. Ajunwa The Paradox of Automation as Anti-Bias Intervention. . [Online]. Available: https://ssrn.com/abstract=2746078.

20. E. Gromova, T. Ivanc, BRICS Law Journal, **7(2)**, 10-36 (2020). https://doi.org/10.21684/2412-2343-2020-7-2-10-36

21. V.V. Blazheev, M.A. Egorova, Pravovoe regulirovanie iskusstvennogo intellekta v usloviyakh pandemii i infodemii [Legal regulation of artificial intelligence in a pandemic and infodemic] (Prospect, Moscow, 2020)