JURISPRUDENCE

Original article
UDC 342.9
doi 10.46741/2713-2811.2026.74.2.001

Artificial Intelligence: Concept and Problems of the Use

LARISA I. BELYAEVA
Academy of Management of the Ministry of Internal Affairs of Russia, Moscow, Russia, beliaeva_055@mail.ru, https://orcid.org/0000-0002-8245-2435

IGOR' V. GOROSHKO
University of Prosecutor's Office of the Russian Federation, Moscow, Russia
Academy of Management of the Ministry of Internal Affairs of Russia, Moscow, Russia
ggoroshko@agprf.org, https://orcid.org/0000-0002-0749-992X

Abstract

Introduction: the article is devoted to the disclosure of an artificial intelligence concept and issues of its use. The relevance is due to the ever-growing interest in the problems of understanding the essence of artificial intelligence, prospects and opportunities for its inclusion in the lives and activities of individuals, government and public structures. The historical roots of this phenomenon are considered. Special attention is paid to characteristics of the current state of artificial intelligence. *Purpose* is to analyze the current practice and prospects for using artificial intelligence. *Methods*: historical, comparative, observations and descriptions of formal and dialectical logic, legal and technical analysis and method of interpretation of legal norms. *Results*: the analysis of the essence of artificial intelligence and its use indicates the inevitable occurrence of problems in this area. Hence, it is crucial to identify them and develop measures for their prevention and elimination. This will ensure the protection of Russian citizens and foreigners in the Russian Federation, society and the state itself from negative consequences of using artificial intelligence. *Conclusions*: it is necessary to improve legislation in the field of development and use of artificial intelligence based on existing legal provisions. Legal regulation should ensure the legality and transparency of the use of artificial intelligence in all spheres of human life and activity.

Keywords: artificial intelligence; science; artificial neural networks; algorithms; technologies; humanization; ethical principles; responsible behavior.

5.1.1. Theoretical and historical legal sciences.

5.1.2. Public law (state law sciences).

For citation: Belyaeva L.I., Goroshko I.V. Artificial intelligence: concept and problems of the use. *Penitentiary Science*, 2026, vol. 20, no. 2 (74), pp. 118–124. doi 10.46741/2713-2811.2026.74.2.001.

Artificial intelligence is currently one of the most dynamically developing and discussed information technologies. It penetrates deeper and deeper into all spheres of society and human life. It seems that the day is not far off when we will witness the headlines, "AI has made a legitimate, reasonable and fair decision", "AI has successfully apprehended a criminal", or "AI has saved a drowning man". At the same time, skeptics warn us against becoming too involved in AI or even scare us with extremely negative consequences of its use. One way or another, AI is of interest to the widest segments of the population, and with interest there is a growing need for its detailed analysis, since, as the history of mankind shows, the successful application of scientific achievements in everyday life is possible only if they are correctly and comprehensively understood.

The term "artificial intelligence" appeared in the middle of the last century, when artificial intelligence and human intelligence were not directly related to each other.

The author of this term, J. McCarthy, interpreted AI as the science and engineering art of creating intelligent machines, especially intelligent computer programs that are designed to solve complex computational problems rather than imitating human intelligence [1].

Indeed, according to the Soviet encyclopedic dictionary of 1987, "intelligence" (from Latin "intellectus" – cognition, understanding, reason) means the ability of thinking, rational cognition, that is, cognition carried out by the mind (in contrast, for example, from sensory cognition, which occurs through the human senses).

In other words, a computer program, in order to be considered intelligent, should ideally be able to implement the full range of cognitive methods, operate with concepts, generate judgments and conclusions, and do this as a person based on previous experience.

In real life, as a rule, specific algorithms for understanding written text, sounding speech, face recognition, and behavioral analysis are unknown. However, this does not prevent us from drawing conclusions and making appropriate decisions based on general patterns, sometimes without thinking much about the mechanisms of their adoption.

It is obvious that seventy years ago, the existing level of development of science and technology did not allow the creation of such computer programs. Therefore, at the first stage, AI was considered primarily as a scientific discipline, within the framework of which hardware and software complexes were developed designed to automate processes in areas characterized by the presence of well-known rules and creative human nature. Mathematics and various games were of interest. Computers that had appeared by that time were actively used for these purposes. So, in 1956, the first program for automatic proof of theorems "Logical-Theorist" appeared, and in 1957 – the first program for playing chess NSS. In the early 1960s, LISP, one of the first most widely used artificial intelligence languages, was developed, allowing symbolic information to be manipulated.

At the same time, research was conducted on modeling the perception of information by the human brain. The year of 1958 witnessed the appearance of the first artificial neural network, a single-layer perceptron, and with its help, the possibility of solving the classification problem using machine learning methods.

The further development of AI was related to the processing of knowledge (data and rules) and the construction of expert systems.

Unlike traditional programs, expert systems are designed to solve tasks having the following characteristics:

– the task cannot be strictly formulated in a numerical form;

– the algorithm for solving the problem is not known in advance, but is built by the expert system itself based on heuristic rules.

The undoubted advantage of the expert system compared to other programs was that after accumulating knowledge, not only a specialist in the subject area, but also any user could work with such a system, which predetermined the popularity of these systems.

At the same time, the results of their work significantly depended on the qualifications of the expert who filled the expert system with knowledge, on the level of his professional training, awareness, and experience. Therefore, these systems could be barely called intelligent [2].

The current state of AI is characterized by a high level of development of machine learning and its subset, deep learning, which has become a key focus in recent years.

In general, machine learning refers to the process of configuring AI parameters (optimizing them) based on the available input data set, carried out in order to find stable patterns in this data necessary to forecast likely results of their development or generate new data.

The AI models used today can be divided into two main groups: generative and discriminative.

Generative models are designed to create new data (text, images) that could be obtained from the input set. Discriminative models focus on finding possible values of the output data, assuming their close dependence on the input data. They analyze the input data and try to determine which class or category it belongs to, or forecast the next value based on the identified dependencies and patterns in the data.

Artificial neural network (ANN) is widely used in the creation of generative and discriminative models.

The artificial neural network is a computer model that to a certain extent reproduces the work of the human brain and implements some of its principles, in particular, such as parallel computing and distributed information processing. The key element of such a network is an artificial neuron, which is an elementary computer that converts input data into output data, taking into account the specified parameters, and transmits the result to the next neuron. As a rule, neurons are arranged in layers interconnected by various configurations. A structure organized in this way resembles the structure of the human brain.

The important characteristic of an artificial neural network is its ability to learn based on input data. The learning process consists of adjusting the parameters in order to achieve greater comparability of the output datasets with the inputs on which the network was trained. The main types of learning in this case are:

– supervised training. In this case, the model parameters are configured based on labeled data that contains reference values. The output data obtained during the processing of input sets is compared with the reference values. In case of the discrepancy, the model parameters are adjusted, and the input set is fed back into the model. The process is repeated until the minimum discrepancy between the reference and output sets is reached. In this case, the parameters of the artificial neural network are configured and it is ready to work;

– unsupervised learning. In this type of learning, non-labeled data is used when setting parameters. The ANN independently determines characteristics of the data and combines them into groups (structures) based on them, essentially solving the clustering problem;

– reinforcement learning. In reinforcement learning, the network interacts with the external environment using an iterative performance evaluation mechanism through a system of rewards and penalties. With a positive assessment, the model updates its parameters to get a bigger reward, in the case of a negative one (penalty), the parameters are adjusted. The learning process continues until an acceptable result is achieved.

With an increase in the number of neurons, layers, and the type of their interconnections in a neural network, the learning process also becomes more complicated. Networks such as fully connected (perceptrons), convolutional, and recurrent networks use deep learning algorithms that allow step-by-step processing of large amounts of data circulating between network layers and extract very complex dependencies from this data.

It can be stated that in its development, artificial intelligence has gone from theoretical research to the creation of computer systems that demonstrate elements of intellectual behavior (communication, learning, reasoning, decision-making). Accordingly, there appears a great number of AI definitions.

For example, N.N. Chernogor believes that AI is a technology that determines the ability of a certain information system, without direct human involvement, to correctly interpret external data (external information), refine the database (databases) with their consideration, learn from the mistakes made and use the data obtained to

achieve specific goals, solutions tasks through flexible adaptation in a poorly defined situation [3, p. 9].

Here is another definition that, in our opinion, most fully reflects modern achievements in the field under consideration: artificial intelligence is a software and hardware complex that provides support and/or effective decision-making in a dynamic, unstable environment at a set time, based on deliberately incomplete and unclear information [4].

The algorithms used in the creation of such complexes have found wide application in projects of computer vision, natural language processing, speech recognition and synthesis, social communications, predictive analytics, etc

At the same time, modern AI embodiments today solve mainly specialized tasks in specific areas of human activity (healthcare, industry, transport) without comprehensive human control. Therefore, such artificial intelligence is called narrowly focused or weak.

As a perspective, the issue on the agenda is the development of a general (or universal) intelligence capable of perceiving the environment as a human (in its comparative, cognitive, emotional and social aspects), functioning equally effectively in various fields, acting in conditions of uncertainty, surpassing human efforts in results. Mathematicians, engineers, programmers, and psychologists are exploring the possibility of creating it and making successful attempts in this direction (generative AI models, such as ChatGPT).

It can be argued that today AI technologies are at a transitional stage of their evolution: from weak to strong artificial intelligence, demonstrating elements of cognitive behavior, the results of which can exceed the results of human work.

The presence of cognitive functions in strong artificial intelligence implies its possession of the following abilities:

– to get information from the environment;

– to perceive (comprehend) this information;

– to learn based on perceived information, to transform it into knowledge;

– to accumulate knowledge;

– to reason (to consider various options of actions);

– to make decisions based on these considerations and forecast their consequences;

– to independently adjust decisions in case of failure (or insufficient achievement) of the result (i.e. to self-study);

– to generate new information.

We emphasize that, although ensuring excellence in results is only a promising characteristic of artificial intelligence, today's AI systems significantly exceed human capabilities in a number of indicators. So, in 2022, several American companies demonstrated artificial neural network models on the market that allow creating images based on text with a high degree of realism.

Another AI system is capable of converting developers' ideas, formulated as text in natural language, into program codes of several programming languages at once.

The AlphaTensor AI system, based on reinforcement learning, finds new, effective algorithms for manipulating matrices, a process that researchers have been working on for decades and which is important for the development of digital infrastructure. And the list of similar examples can be continued.

Today, key directions of AI technology development in the world are the following:

– intelligent vision, designed to analyze and extract information from photo and video images. Within the framework of this technology, a whole set of tasks is solved, including evaluating human intentions based on body position images; semantic segmentation (correlating individual pixels of an image with selected categories – person, car, apartment); image generation, etc.;

– natural language processing. The technology includes text understanding, summarization, sentimental analysis, text generation, and machine translation. Summarization involves extracting a short summary from the source text, preserving the main essence and key ideas. Sentiment analysis is used to determine the emotional coloring of texts and to classify text data as positive, negative or neutral, depending on the author's opinion or feelings expressed in them.

The natural language processing technology is also implemented in large language models, which are becoming increasingly popular;

– speech recognition and synthesis. The technology implies the ability of artificial intelligence systems to recognize spoken words

and convert them into text. This technology has advanced a lot in recent years. Currently, many computer applications are capable of providing real-time text messaging. To do this, they support dictation devices that can easily convert speech into writing.

The leading positions in the world in the development of artificial intelligence systems, investments and research in this field belong to the USA and China [5].

According to the TheGlobal AI Index 2024 rating compiled by TortoiseMedia, the level of AI development in individual countries depends on a set of factors that are divided into seven groups:

– availability of competencies for implementing AI ("Talent" ("personnel") indicator);

– availability of infrastructure (communications, capacities) ("Infrastructure");

– regulatory conditions and degree of public support ("Work environment");

– availability of research on AI ("Research");

– availability of innovative developments (research results), including patent activity ("Development");

– level of government support for investments in AI ("Government Support");

– activity of startups and businesses in the field of AI ("Commercial sector").

The United States and China take a lead by all these indicators [6].

According to the analytical company "Precedence Research", the volume of the AI market in the United States in 2025 was 173.6 billion dollars, which was almost 30 billion higher than in 2024 [7]. OpenAI alone invested 41 billion dollars in artificial intelligence technologies.

The volume of the AI market in China in 2025 was smaller – about 142 billion dollars [8]. However, China is significantly ahead of the United States in the number of patents granted in the field of AI. So, 69.7% of the AI patents issued accounts for China and 14.2% for the United States [9].

"Precedence Research" forecasts a 20% AI annual growth rate from 2025 to 2034, with the Asia-Pacific region being the fastest growing market.

Other countries, striving to keep up with the development of artificial intelligence, focus on creating favorable conditions for business and research, making the country attractive for external investment, and developing regulations to increase their effectiveness.

Undoubtedly, it is impossible to stop AI development. It makes our lives more dynamic, versatile and diverse. But at the same time, it also creates a number of problems that require attention and resolution. In particular, there is concern that if an AI interacts with a person without appropriate restrictions (or rules), then such interaction can lead to negative consequences for the individual, including mental disorders, since such a system of relationships does not know and does not understand such important connecting elements of human communication as empathy, love, duty, responsibility, etc. [10–20].

We believe that by introducing advanced technologies into everyday life, it is necessary to take into account humanization of these processes, preservation and development of a person's personality, his/her emotional and spiritual world. In this regard, it is important to develop a system of ethical principles that establish guidelines in the process of human-AI system interaction.

There are already examples of such solutions. In 2021, at the international forum "Ethics of Artificial Intelligence: the Beginning of Trust", the Code of Ethics in the field of Artificial Intelligence was signed, which establishes general ethical principles and standards of behavior that should guide participants in artificial intelligence relations in their activities, as well as mechanisms for their implementation.

The idea of developing such a code was submitted a year earlier by President of the Russian Federation Vladimir Putin at the plenary session of the international online conference "AI Journey". By the end of 2025, the Code of Ethics in the field of Artificial Intelligence had been signed by over 900 Russian and foreign companies, scientific and educational organizations.

The main purpose of the code is to establish common ethical principles and standards of conduct that should guide participants in artificial intelligence relationships when creating, implementing and using technologies at all stages of the life cycle.

In 2024, Russian member companies of the Alliance for Artificial Intelligence Association, leading universities and scientific organizations

in the country (HSE, Skoltech, Innopolis University, Moscow Institute of Physics and Technology, National Research Lobachevsky State University of Nizhny Novgorod, ITMO University) signed a Declaration on Responsible Development and Use of Services Based on Generative AI, which specifies the provisions regarding generative artificial intelligence. It contains recommendations for both developers and those who use it for their own purposes.

Users are advised not to use generative AI to cause harm or for other illegal purposes and not to purposefully create incorrect information, for example, contrary to the law or generally accepted standards of morality.

An important point concerns the situation when generative models are supposed to be used where the manifestation of a person's creative contribution is obviously required. In this case, users are also not recommended to do this.

It seems that the ideas of responsible behavior laid down in the Code of Ethics in the Field of Artificial Intelligence and the Declaration on Responsible Development and Use of Services Based on Generative AI can be scaled up and implemented in everyday life when implementing human–AI system interaction.

In conclusion, we draw attention to an interesting paradox: in the context of the rapid development of artificial intelligence and its widespread introduction into various spheres of human activity, the requirements for the quality and level of natural intelligence, its ability to critical thinking and information analysis are increasing. Therefore, the task on the agenda is not to neglect these requirements, not to ignore them, so as not to make personal and public well-being dependent on decisions made by a computer. It is important to remember that AI algorithms are not perfect and error-prone, as they are based on assumptions or constantly updated training samples.

REFERENCES

1. McCarthy J. *What is artificial intelligence?* Available at: http://www-formal.stanford.edu/jmc/whatisai.html) (accessed April 14, 2025).
2. *Iskusstvennyi intellekt: v 3 kn. Kn. 1. Sistemy obshcheniya i ekspertnye sistemy: spravochnik* [Artificial intelligence: in 3 books. Book 1. Communication systems and expert systems: handbook]. Ed. by Popov. E.V. Moscow, 1990. 464 p.
3. Chernogor N.N. Artificial intelligence and its role in the transformation of modern law and order. *Zhurnal rossiiskogo prava = Journal of Russian Law*, 2022, vol. 26, no. 4, pp. 5–15. (In Russ.).
4. Larina E.S., Ovchinskii V.S. *Iskusstvennyi intellekt. Bol'shie dannye* [Artificial intelligence. Big Data]. Moscow, 2018. 416 p.
5. *Rossiya zanyala 29-e mesto v mirovom reitinge razvitiya II – lidiruyut SShA i Kitai* [Russia ranks 29th in the global AI development ranking, with the United States and China leading the way]. Available at: https://www.semafor.com/article/11/20/2024/stanford-names-us-the-top-ai-ecosystem-china-follows (accessed April 14, 2026).
6. *TortoiseMedia. TheGlobalAIIndex 2024*. Available at: https://ict.moscow/analytics/the-global-ai-index-2024/ (accessed April 15, 2026).
7. *What is U.S. Artificial Intelligence (AI) Market Size?* Available at: https://www.precedenceresearch.com/us-artificial-intelligence-market (accessed April 15, 2026).
8. *China core AI industry scale exceeds 1 trln yuan in 2025*. Available at: https://english.www.gov.cn/archive/statistics/202512/26/content_WS694e383cc6d00ca5f9a0847a.html (accessed April 14, 2026).
9. *AI Index Report The 2025 AI Index R 9. The 2025 AI Index Report*. Available at: https://hai.stanford.edu/ai-index/2025-ai-index-report (accessed April 14, 2026).
10. Ibragimov R.S., Suragina E.D., Churilov D.Yu. Ethics and regulation of artificial intelligence. *Zakon = Law*, 2021, no. 8, pp. 85–95. (In Russ.).
11. Kozyreva A.A., Pirozhkova T.V. Application of artificial intelligence in justice. *Administrator suda = Court Administrator*, 2021, no. 2, pp. 12–16. (In Russ.).
12. Kurochkin S.A. On the prospects of using artificial intelligence in civil and arbitration proceedings. *Arbitrazhnyi i grazhdanskii protsess = Arbitration and Civil Procedure*, 2023, no. 2, pp. 37–42. (In Russ.).

13. Makutchev A.V. Modern possibilities and limits of artificial intelligence introduction into the justice system. *Aktual'nye problemy rossiiskogo prava = Actual Problems of Russian Law*, 2022, vol. 17, no. 8 (141), pp. 47–58. (In Russ.).
14. Malina M.A. World justice and artificial intelligence*. Mirovoi sud'ya = Magistrate Judge*, 2021, no. 4, pp. 17–21. (In Russ.).
15. Savina V.S. Artificial intelligence and modern art: problems and prospects. *Pravo i biznes = Law and Business*, 2021, no. 3, pp. 8–10. (In Russ.).
16. Kharitonova Yu.S., Savina V.S., Pan'ini F. Civil liability in the development and application of artificial intelligence and robotics systems: basic approaches. *Vestnik Permskogo universiteta. Yuridicheskie nauki = Bulletin of Perm University. Legal Sciences*, 2022, no. 4, pp. 683–708. (In Russ.).
17. Tsvetkov Yu.A. Artificial intelligence in justice. *Zakon = Law*, 2021, no. 4, pp. 91–107. (In Russ.).
18. Shakhnazarov B.A. Legal regulation of relations using artificial intelligence. *Aktual'nye problemy rossiiskogo prava = Actual Problems of Russian Law*, 2022, no. 8, pp. 63–72. (In Russ.).
19. Blinov V.S. New technologies – old problems. Artificial intelligence as a legal entity. *Pravo i biznes = Law and Business*, 2022, no. 4, pp. 3–9. (In Russ.).
20. Taran K.K. Prerequisites for the legal regulation of the results of intellectual activity created using artificial intelligence. *Pravo i ekonomika = Law and Economic*s, 2023, no. 1, pp. 20–26. (In Russ.).

INFORMATION ABOUT THE AUTHORS

LARISA I. BELYAEVA – Doctor of Sciences (Law), Professor, Honored Lawyer of the Russian Federation, Honorary Employee of the Ministry of Internal Affairs of Russia, Honored Professor of the Academy of Management of the Ministry of Internal Affairs of Russia, Honorary Doctor of the Vologda Institute of Law and Economics of the Federal Penitentiary Service of Russia, professor at the Department of Criminal Policy of the Academy of Management of the Ministry of Internal Affairs of Russia, Moscow, Russia, beliaeva_055@mail.ru, https://orcid.org/0000-0002-8245-2435
IGOR' V. GOROSHKO – Doctor of Sciences (Technology), Professor, Head of the Department of Legal Statistics and Information Support of Prosecutorial Activities of the University of Prosecutor's Office of the Russian Federation, Moscow, Russia, professor at the Department of Information Technology of the Academy of Management of the Ministry of Internal Affairs of Russia, Moscow, Russia, ggoroshko@agprf.org, https://orcid.org/0000-0002-0749-992X

Received April 23, 2026