ИНФОРМАЦИОННОЕ ПРАВО
И ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

DEFINITION OF ARTIFICIAL INTELLIGENCE
IN THE CONTEXT OF THE RUSSIAN LEGAL SYSTEM:
A CRITICAL APPROACH

© 2022   V. V. Arkhipov[1, *], A. V. Gracheva[1, **], V. B. Naumov[2, 3, ***],
T. A. Polyakova[2, ****], A. V. Minbaleev[2, 4, *****]

[1]*Saint Petersburg State University*
[2]*Institute of State and Law of the Russian Academy of Sciences, Moscow*
[3]*Saint Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS)*
[4]*Kutafin Moscow State Law University (MSAL)*

*E-mail: vladislav.arkhipov@gmail.com
**E-mail: anastasia.v.gracheva@gmail.com
***E-mail: nau@russianlaw.net
****E-mail: polyakova_ta@mail.ru
*****E-mail: alexmin@bk.ru

Abstract. The article addresses the normative definition of artificial intelligence as provided in the Russian legislation by 2021 in the context of general approach to this term. Authors identify principal legally significant features of artificial intelligence that include autonomy of functioning (complete or relative), the ability to find solutions (including without predetermined algorithms) and "inorganic" nature that imply − strictly from formal legal, and not factual, perspective − potential risk and, subsequently, potential danger of artificial intelligence to be understood in formalized legal sense. The authors take into consideration examples of approaches to defining the term in the world, raise criticism of the existing definition and suggest more concise approach based on defining artificial intelligence as an information system (in formal legal sense) which provides the function of forming algorithms for solving problems different from those laid down by the system developers, including artificial neural network.

ОПРЕДЕЛЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В КОНТЕКСТЕ РОССИЙСКОЙ ПРАВОВОЙ СИСТЕМЫ: КРИТИЧЕСКИЙ ПОДХОД

© 2022 г.   В. В. Архипов[1, *], А. В. Грачева[1, **], В. Б. Наумов[2, 3, ***], Т. А. Полякова[2, ****], А. В. Минбалеев[2, 4, *****]

[1]Санкт-Петербургский государственный университет
[2]Институт государства и права Российской академии наук, г. Москва
[3]Санкт-Петербургский федеральный исследовательский центр Российской академии наук (СПИИРАН)
[4]Московский государственный юридический университет им. О.Е Кутафина (МГЮА)

*E-mail: vladislav.arkhipov@gmail.com
**E-mail: anastasia.v.gracheva@gmail.com
***E-mail: nau@russianlaw.net
****E-mail: polyakova_ta@mail.ru
*****E-mail: alexmin@bk.ru

Аннотация. В статье рассматривается нормативное определение искусственного интеллекта, сформулированное в российском законодательстве к 2021 г. в контексте общего подхода к данному термину. Авторы выделяют основные юридически значимые признаки искусственного интеллекта, к которым относятся автономность функционирования (полная или относительная), способность находить решения (в том числе без заранее заложенных разработчиком алгоритмов) и «неорганическая» природа, предполагающая — строго с формально-юридической, а не фактической точки зрения — потенциальный риск и, соответственно, потенциальную опасность искусственного интеллекта для понимания его в формально-юридическом смысле. Авторы рассматривают примеры существующих в мире подходов к определению данного термина, подвергают критике действующее определение и предлагают более емкий подход, основанный на определении искусственного интеллекта как информационной системы (в формально-юридическом смысле), обеспечивающей функцию формирования алгоритмов решения задач, отличных от заложенных разработчиками системы, в том числе с помощью искусственной нейронной сети.

Ключевые слова: право, искусственный интеллект, определение, нейронная сеть.

1. Introduction

1.1. Scope, purpose and methodology of the paper

The normative definition of artificial intelligence (hereinafter, "**AI**") was established in Russia at the level of the National Strategy for the Development of Artificial Intelligence for the period up to 2030, approved by the Presidential Decree No. 490 of 10 October 2019 "On the development of artificial intelligence in the Russian Federation" (hereinafter, the "**National Strategy**")[1]. According to Paragraph "a" Part 5 of the National Strategy, "*artificial intelligence [is] a set of technological solutions that allows to imitate the cognitive functions of a human being (including self-learning and search for solutions without a predetermined algorithm) and obtain results when performing specific tasks, comparable, at least, with the results of human intellectual activity. The set of technological solutions includes information and communication infrastructure, software (including that which uses machine learning methods), processes and services to process data and search for solutions*". In addition, the National Strategy enshrines a number of

---

[1] See: National Strategy for the Development of Artificial Intelligence for the period up to 2030, approved by the Presidential Decree No. 490 of 10 October 2019 "On the development of artificial intelligence in the Russian Federation" // "ConsultantPlus" legal reference system (in Russian) (accessed: 2021 August 23).

basic principles for the development and use of AI technologies, namely the principles of protecting human rights and freedoms, security, transparency, technological sovereignty, the integrity of the innovation cycle, reasonable economy, and supporting competition. For the purposes of this paper and taking into account the application of the formal-dogmatic method, it should be noted that the National Strategy has not only legal goals, but also a number of others, including economic and political. As a consequence, individual principles are aimed not only and not so much at the legal aspects of AI, but at the priorities of politics, economics and/or other institutional areas. The principles of protection of human rights and freedoms, security and transparency should be attributed to the principles that are directly related to the field of law. At the same time, the principle of protection of human rights and freedoms is a universal and general legal principle.

That said, the authors of this work believe that the definition of AI, presented in the National Strategy and subsequently reproduced in other normative documents of the Russian Federation, although not being fundamentally flawed (in the light of general approaches to the understanding of artificial intelligence), can be improved and perhaps made more concise. This is due to the need for a clearer definition and delineation of the use of AI in various aspects of public life.

The main problems with the abovementioned definition are the lack of specificity of artificial intelligence as a high-risk technology, and the legislator's focus on artificial intelligence as solely an imitation of human cognitive abilities, resulting in tasks close to the results of the human brain. This approach blurs the legal concept and causes difficulties at the stage of law enforcement. Thus, the purpose of this article is to reconstruct the concept of artificial intelligence for the Russian legal system based on the following principal methods of research:

1. Dialectical development is a method that helps trace the evolution of the formal definition of AI from the purely technical approach, which is fundamental to today's legal definition of AI, to a more versatile approach that takes into account not only the purposes of technical standardization and regulation, but also legal and ethical frameworks. The expansion of the scientific concepts of AI, the experience of making definitions in this category in other countries' laws, and the emerging Russian and International Law enforcement practices determine a novel approach to defining AI. This, in turn, creates a methodological need for a better legal definition of AI.

2. Comparative legal method based on the comparison of legal phenomena of the same order, belonging to different legal systems, in order to identify similarities and differences between them. The use of the comparative legal method makes it possible to establish which of the recommendations already developed in foreign legal science and practice can be used in the study focused on the domestic legal reality.

Such borrowings should not be arbitrary and should be directly or indirectly justified, including in a deductive way (if in foreign doctrine and practice they are a direct consequence of the same principles that apply in the Russian Federation, including cases of obligations arising from common for the objects of comparison international treaties) or essential similarity of social and economic conditions, to which such principles or recommendations are oriented.

3. Formal-legal (dogmatic) method, which includes two ways of reasoning: deductive-axiomatic and hypothetic-deductive. The deductive-axiomatic method is based on accepting the provisions of positive law as axioms and consistently applying the techniques of legal interpretation. The hypothetic-deductive method is similar to the deductive-axiomatic method, but it involves positing a hypothesis about the possible end result of legal interpretation, which is refuted or confirmed through legal argumentation. The use of the formal-legal method allows to determine the positive legal regulation of certain public relations at a particular moment, as well as to identify gaps in the law and (or) lack of formal certainty of the studied legal norms in relation to the subject of research.

1.2. Literature review in relation to the paper

The main challenge in the legal definition of AI posed by this paper is to identify all the specific criteria that determine the concept and its high-risk nature. For this purpose, the authors address the legal and policy debates. G. Hallevy defines AI "as having some human skills and being capable of an imperfect and incomplete imitation of the human mind"[2]. J. Turner speaks of the possibility of having moral rights for the artificial intelligence[3]. J.P. Günther argues, that the main purpose of AI is to reproduce human procedures of solving problems on computers, i.e. human intelligence[4].

J. Schuett deduces three possible types of definitions of AI (through the Turing test, through the notion of intelligent agent, and the field of intelligent machine development) and concludes that the legal definitions of AI are highly over-inclusive, vague, not comprehensive and impractical[5].

Other authors[6] argues that the definition of AI is still necessary for legal practice and therefore it should not be based on a comparison of AI with human cognitive skills

---

[2] *Hallevy G.* When robots kill: artificial intelligence under Criminal Law. Boston (2013).

[3] See: *Turner J.* Robot rules: regulating artificial intelligence. London (2019).

[4] See: *Günther J. P.* Roboter und rechtliche Verantwortung. München (2016).

[5] See: *Schuett J.* Defining the scope of AI regulations, https://arxiv.org/abs/1909.01095 (accessed: 2021 August 23).

[6] See: *Simmler M., Markwalder N.* Roboter in der Verantwortung? // Zeitschrift für die gesamte Strafrechtswissenschaft. 129(1), 20—47 (2017); *Scherer M. U.* Regulating artificial intelligence systems: risks, challenges, competencies, and strategies // Harv. JL & Tech 29(2), 353—400 (2015).

but rather on AI features and technical characteristics, in particular the criterion of autonomy, and risks associated with AI technologies as such. The authors of this paper argue for this position and follow its argument.

In addition, it should be noted that one of the criteria proposed by legal scholars to define artificial intelligence − the imitation of human cognitive skills − is a rather ambiguous issue in various fields of scientific knowledge other than law. Therefore, the authors take into account the following debates.

For instance, experts in semiotics and semantics emphasize that the identification of artificial intelligence with the human mind is a common linguistic and semantic error. There are two, often related, processes: "personification of a mechanism (robot, automaton, computer) and objectification, depersonification of an organism / person (a human being)"[7].

Some specialists in the field of information technology claim that "attempts to formalize and generalize these efforts, although potentially of great scientific value, have not led to breakthrough results, as they proceeded from the idea that the process of solving "intellectual tasks" must imitate (i.e. be similar to) similar human activities in every respect"[8]. On the contrary, they emphasize that "the greatest practical successes have been achieved using methods that are not inherent to man in principle, but are based on the use of 'brute force computing', or in other words, the ability to quickly go through various solutions with the aid of high-performance computers"[9].

Psychologists also explain that claims that machines have and imitate cognitive skills and properties are a very dangerous idea in the long run. For example, McDermott requires artificial intelligence researchers to demonstrate that their software actually implements these psychological properties (in particular, attributing joy to a robot hoover) before using psychological terms in scientific articles[10]. D. Proudfoot even introduces the notion of a "forensic anthropomorphism problem", which is this: the anthropomorphisation of artificial intelligence risks introducing bias (in favour of the machine) into judgements about machine intelligence[11].

---

[7] *Volkov V.V.* Artificial Intelligence and the human mind: futuristic synecdoche and reality (linguistic and linguo-mental aspects) // Vestnik Rossijskogo universiteta druzhby narodov. Ser.: Teoriya yazyka. Semiotika. Semantika 11(4), 745−759 (2020).

[8] *Kalyaev I., Zaborovskij V.* Artificial Intelligence: from metaphor to technical solutions // Control Engineering Rusia 83(5), 26−31 (2019).

[9] Ibid.

[10] See: *McDermott D.* Artificial intelligence meets natural stupidity. SIGART Newsletter 57, 4−9 (1976).

[11] See: *Proudfoot D.* Anthropomorphism and AI: Turing's much misunderstood imitation game // Artificial intelligence 175(5−6), 950−957 (2011).

2. Understanding AI in the context of law

2.1. General approach to the term "artificial intelligence"

The words "artificial intelligence" has firmly entered the modern language, but they can be used in different, albeit related, senses, and therefore need to be specified. Some authors relate the appearance of these words in the sense of designating one of the disciplines of computer science to 1956[12]. Subsequently, one of the main meanings of these words became not so much a designation of the field of knowledge, but rather a designation of various technologies involving automated computer actions. In its broadest sense, the words "artificial intelligence" are used more as a rhetorical device designed to dramatize the use of computer technology, behind which an ordinary algorithm or computer program is hidden, and therefore in its strict scientific sense it is redundant in this usage. At the same time, with a certain approach the use of these words can make sense.

R. Susskind once pointed out as a shortcoming of these words that it is "used in many ways and is often wielded as no more than a rather blunt marketing weapon or as a part of an alerting headline or tweet"[13]. At the same time, he suggests that there are two possible approaches to defining AI in a more or less rigorous sense: "architectural" and "functional". The definition of AI in terms of the architectural approach implies a reference to the technology used (during the "first wave" of AI development such technologies were based on the principle of building explicit software algorithms, and during the "second wave" − on neural networks and machine learning). From a functional perspective, AI is defined by the fact of using software systems to perform those tasks that traditionally required the application of human thinking[14] regardless of technology. While we agree with the author that it is the functional approach to defining AI that is most relevant to legal research, we emphasize that it does not override the need to consider the architectural features of individual AI technologies, as they can have a significant impact on legal issues.

One such technology, in particular, is neural networks that involve machine learning. We believe that we should agree with those authors who associate the general interest in the problems of AI precisely with these technologies[15] − this is confirmed by individual studies in the field of legal informatics, noted in the next section of this paper. We believe that for the purposes of legal research it is not advisable to focus on the "architectural" technical definition of neural networks in detail − for the purposes

---

[12] See: *Hunter D.* The death of the legal profession and the future of law // UNSW Law Journal 43(4), 1199−1225 (2020).

[13] *Susskind R.E.* Online courts and the future of justice. Oxford (2019).

[14] See: ibid.

[15] See: *Hunter D.* Op. cit.

of this study it is the "functional" approach that defines the legally significant principles of this technology that matters. The following is fundamentally important in it: unlike the "first wave" expert systems (according to R. Susskind's conditional periodization), neural networks do not have a predetermined algorithm in the strict sense, but "learn" from big data, selected the be processing by such a system. At the same time, neural networks can operate on any data, including legal data, in order to establish dependencies between them[16].

As D. Hunter notes, "*at its core, deep learning is a statistical method for classifying patterns, based on large amounts of sample data, using neural networks with multiple layers*"[17]. In other words, by default a neural network not only does not have a predetermined algorithm understood literally (which could, for example, reproduce the logic of legal reasoning, as was considered in the legal analysis of the particularities of the possible use of expert systems during the "first wave" of AI), but it also does not operate in a way that can be correlated with the logic of any decision in the strict sense: machine learning-based neural networks, at first sight, do not even remotely simulate the logical reasoning. The principle underlying the functioning of a neural network can thus be described as prediction based on *sui generis* statistical analysis. This seems important in order to understand how exactly a neural network can "make a decision" − on the principle of predicting what the corresponding decision should be, based on already available examples. Of course, a combination of the "first" and the "second wave" AI technologies is not technically excluded, but if we imagine a hypothetical technological solution based solely on a neural network with machine learning, the conclusions it provides can be likened to learning a foreign language not in a deductive way in school, but by imitating and mimicking foreigners during personal immersion in a language environment[18].

2.2. Principal legal features of artificial intelligence

The totality of legal norms aimed at regulating social relations in connection with and about artificial intelligence will have to represent an inter-branch institute (implying a number of institutional legal principles). As noted, for example, by D.A. Lipinsky and O.E. Repeteva, "a legal institution is characterized by the presence of some obligatory features: homogeneity of regulated sphere of social relations; an independent subject of legal regulation; a single function; the presence of norms similarly regulating a given relation; legal unity of norms, etc."[19]. According to A.V. Polyakov and E.V. Timoshina, "[i]nstitution of law is a stable group of legal norms regulating a certain kind

of qualitatively homogeneous public relations"[20]. At this point any statements about "the law of AI" as a part of positive law in general or as a certain already established legal institution in particular will obviously be premature due to the fact that even in cases where such legal regulation exists, it is usually fragmentary in nature, which follows from the application of the comparative legal method. Nevertheless, using the method of theoretical modeling, we can note that the relations in connection with and about the application of AI technologies will meet the signs of a legal institution, and this legal institution will be inter-branch. Institute of "law of AI" meets the above substantive criteria (in particular, homogeneity [potentially] regulated sphere of social relations and the presence of an independent subject of legal regulation) and, subject to a systematic approach to legislative technique, will meet the formal-legal criteria, such as the presence of norms, similarly regulating this relation, legal unity of norms, etc.

Besides, the relative isolation of AI as a technology makes it an independent thing under law that pertains to the functioning of multiple special entities under law; this is an important trait of AI. Besides, whether AI can be an entity under law is a question currently subject to much discussion. A number of rights and responsibilities are associated with AI. Therefore, we are observing the emergence of relatively well-defined legal relations that pertain to the creation and operation of AI. In addition, this newly emerging institution has a system of institutional principles of its own; these principles are indicative of the risky nature of the technology and reflect the need to enshrine in law the constraints and prohibitions, assumptions and special experimental permits applicable to the use of AI. Some of such principles stem indeed from the fact that AI is a digital technology; however, others are AI-specific.

The interdisciplinary nature of the institution is determined by the "architectural" legally significant features of AI, which include the autonomy of functioning (complete or relative), the ability to find solutions (based on both predetermined algorithms and without them − the latter feature must be considered as more important since it permits to distinguish AI from mere automated solutions with traceable internal logic) and the "inorganic" nature. We believe that these two qualities are on the same logical level and cannot be reduced to one another, while the other qualities (e.g., potential danger in certain areas or opaqueness of the sequence of search for solutions in certain contexts) are the consequence of these qualities. For example, the relative autonomy of operation may be characteristic of many devices and mechanisms, including vending machines well known to civil law. However, vending machines are not necessarily characterized by the ability to search for solutions (if they do not use AI in the sense of this study). The ability to search for solutions is

---

[16] See: *Hunter D.* Op. cit.

[17] Ibid.

[18] See: *Susskind R.E.* Op. cit.

[19] *Lipinsky D.A., Repeteva O.E.* On the inter-branch institution of legal liability // Yuridicheskaya Mysl' 4(96), 24−32 (2016).

[20] *Polyakov A.V., Timoshina E.V.* General theory of law. St. Petersburg (2005).

Definitions of artificial intelligence in different legal systems

Country	Definition
United Kingdom	"AI is understood as technologies with the ability to perform tasks that would otherwise require human intelligence, such as visual perception, speech recognition, and language translation"[21].
USA	"AI means a variety of information processing techniques and technologies used to perform a goal-oriented task and the means to reason in the pursuit of that task"[22]. "The term "artificial intelligence" includes the following (a) Any artificial systems that perform tasks under varying and unpredictable circumstances, without significant human oversight, or that can learn from their experience and improve their performance. Such systems may be developed in computer software, physical hardware, or other contexts not yet contemplated. They may solve tasks requiring human-like perception, cognition, planning, learning, communication, or physical action. In general, the more human-like the system within the context of its tasks, the more it can be said to use AI. (b) Systems that think like humans, such as cognitive architectures and neural networks. (c) Systems that act like humans, such as systems that can pass the Turing test or other comparable test via natural language processing, knowledge representation, automated reasoning, and learning. (d) A set of techniques, including machine learning, that seek to approximate some cognitive task. (e) Systems that act rationally, such as intelligent software agents and embodied robots that achieve goals via perception, planning, reasoning, learning, communicating, decision making, and acting"[23].
European Union	"'artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with"[24].
Germany	"AI systems are human-designed intelligent systems consisting of hardware and/or software components that aim to solve complex problems and tasks in interaction with and for the digital or physical world"[25].
China	"AI has become a new engine for economic development. As the core driver of a new round of industrial change, artificial intelligence will further unleash the enormous energy accumulated in previous scientific and technological revolutions and industrial changes, and create a new powerful engine that will reconfigure all aspects of economic activities such as production, distribution, exchange and consumption, forming new intelligent demands in various fields from macro to micro, giving rise to new technologies, new products, new industries, new business models and new modes, triggering major changes in the economic structure"[26].
France	"the part of intelligence delegated under the term AI is the learning intelligence based on the analysis of data provided to the AI, (machine learning or deep learning for example) requiring data processing capacities. The key technological elements of AI are therefore, on the one hand, the collection, veracity, reliability, organization, security and traceability of data and, on the other hand, the construction of learning algorithms adapted to provide simulation, prediction and control models for industrial systems"[27].
Estonia	"'Kratt' in the framework of this project, meaning practical applications based on AI technologies (in the narrow meaning of AI) performing a specific function"[28].

not possible without autonomy but distinguishes AI and devices with AI from other software and from other devices, respectively. The feature of inorganic nature is formal, but logically necessary, because autonomy and the ability to search for solutions without this feature can themselves characterize humans as well.

Further, from the combination of the features of autonomy and the ability to find solutions, follows another fundamental quality — the potentially risky nature of AI technologies. It is too early to qualify this quality directly as "real [not potential] risky nature" or to qualify AI as a "source of increased danger", because (a) not every AI technology has a real, and not potential, risky nature, and (b) not every example of risky

nature can be legally qualified without doubt as a source of increased danger. Both the real risky nature and the danger are determined by the field of use of AI technologies, which, in turn, determines whether the use of AI in a given case is socially and legally significant. So, for example, the use of artificial intelligence technologies in

---

[21] Industrial Strategy. Building a Britain fit for the future, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/664563/industrial-strategy-white-paper-web-ready-version.pdf (accessed: 2021 August 23).

[22] NIST. U.S. LEADERSHIP IN AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools, https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf (accessed: 2021 August 23).

[23] H.R.4625 − FUTURE of Artificial Intelligence Act of 2017, https://www.congress.gov/bill/115th-congress/house-bill/4625/text (accessed: 2021 August 23).

[24] Proposal for a regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (accessed: 2021 August 23).

[25] Deutscher Bundestag Drucksache 19/23700 19. Wahlperiode 28.10.2020 Unterrichtung der Enquete-Kommission Künstliche Intelligenz − Gesellschaftliche Verantwortung und wirtschaftliche, soziale und ökologische Potenziale, https://dserver.bundestag.de/btd/19/237/1923700.pdf#page=51&zoom=100,82,166 (accessed: 2021 August 23).

[26] Plan for the Development of New Generation Artificial Intelligence (Guo Fa [2017] No. 35), http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm (accessed: 2021 August 23).

[27] Definition Intelligence Artificielle dans les systems industriels de production, http://www.industrie-dufutur.org/content/uploads/2019/02/Intelligence-Artificielle-AIF-Janvier-2019.pdf (accessed: 2021 August 23).

[28] Report of Estonia's AI Taskforce, https://f98cc689-5814-47ec-86b3-db505a7c3978.filesusr.com/ugd/7df26f_486454c9f32340b28206e140350159cf.pdf (accessed: 2021 August 23).

"ordinary" (not associated with professional cybersports with high investments and not associated with the circulation of virtual items for real money) computer games may conditionally qualify only as potentially risky in nature (i.e. simply a game with AI, as a rule, rather than something leading to socially significant consequences). The same can be said, for example, of the use of AI technologies to create works of art. However, the use of AI technologies in securities trading is already characterized by its real risky nature, although this quality cannot yet be called "dangerous" in many cases (in the sense of the genuine concept of sources of increased danger, for example). Finally, the use of AI technologies in self-driving cars would already be "dangerous" in the complete sense of the word. Nevertheless, this is consistent with the proposed general concept of the key legally significant qualities of AI technologies: autonomy, solution-seeking ability and inorganic nature define a potentially risky nature that may become "real" and even apparently imply "increased danger".

2.3. Examples of approaches to defining the term in the world

Many countries (e.g. USA, China, France, etc., see *table 1* below) mention in draft, industrial and/or strategic documents that artificial intelligence may develop to a "strong" form, but despite this they conceptually proceed from its "weak" version (according to J. Searle) to address today's problems. This primarily affects the configuration of possible legal relations, where the subjects of law are understood to be exclusively persons recognized as such by law who are directly or indirectly associated with AI, and the object is either AI in general, or specific methods or types of systems.

As for regulation, not all countries see the need to establish definitions of the object of legal relations. One can find (1) a broad approach to the definition, (2) a narrow one detailing all technologies and methods, (3) a refusal to formulate a precise definition. A different classification of the definition of the object could be: (1) linguistic approach (referring to a dictionary or field of scientific knowledge), (2) technical approach (referring to a technical definition), (3) sociological approach (understanding AI through a set of consequences and effects on society).

2.4. Criticisms of the existing legal definition of artificial intelligence

Among other examples from Russian legal system, a representative definition of AI that is consonant with the one from the National Strategy, is contained in the Federal Law "On conducting an experiment on the establishment of special regulation in order to create the necessary conditions for the development and implementation of artificial intelligence technologies in the subject of the Russian Federation – the city of federal significance Moscow and

Amendments to Articles 6 and 10 of the Federal Law 'On Personal Data'" of 24 April 2020 No. 123-FZ[29].

Based on the definition presented in the Federal Law of 24 April 2020 No. 123-FZ, AI, which is subject to regulation, is characterized by the following: (1) the ability to imitate human cognitive functions (self-learning and search for solutions without a predetermined algorithm), (2) the ability to obtain, when performing specific tasks, results comparable to or superior to the results of human intellectual activity.

First, here AI is defined through its similarity to human cognitive abilities, which in itself is quite controversial. A. Turing wrote about AI's imitation of human cognitive abilities, proposing a mental experiment that later became known as the Turing test[30]. However, e.g. B. Whitby relates this test to the history of science due to its emphasis on imitation of human behavior, which does not correspond to the modern challenges of AI[31]. Jordan Pollack said that the big mistake scientists make is assuming that human intelligence is the greatest intelligence there is, and so he urged the science of AI to focus on "meaningless intelligence"[32]. Turing's test for detecting artificial intelligence's imitation of human abilities, as D. Proudfoot suggests, can only uncover the human tendency to anthropomorphize artificial systems, and no more[33].

Moreover, the notion of "intelligence" and the possibilities of its imitation through various mechanisms has changed considerably in the course of the development of scientific knowledge. "Practically until the middle of the nineteenth century, simple counting tasks seemed to be quite intelligent"[34]. Therefore, it seems difficult to formulate an exhaustive and consistent list of legal criteria for classifying artificial intelligence as a machine that imitates human cognitive abilities.

As a result, if we turn to the experience of foreign countries, the reference to human cognitive abilities is often found only in strategic or advisory documents. The only exception is the US bill[35], which proposes to introduce, among other things, such definitions of artificial intelligence as "a set of methods, including machine learning, that aims to approximate some cognitive task", "systems that think like humans, such as cognitive architectures and neural networks". However, most countries define artificial

---

[29] See: "ConsultantPlus" legal reference system (in Russ.) (accessed: 2021 August 23).

[30] See: *Turing A.* Can machines think? Moscow (1960).

[31] See: *Whitby B.* The Turing test: AI's biggest blind alley? // Machines and thought: the legacy of Alan Turing (1), 53–62 (1996).

[32] *Pollack J.B.* Mindless intelligence // IEEE Intelligent Systems. (21)3, 50–56 (2006).

[33] See: *Proudfoot D.* Op. cit.

[34] *Kalyaev I., Zaborovskij V.* Op. cit.

[35] See: H.R.4625 – FUTURE of Artificial Intelligence Act of 2017, https://www.congress.gov/bill/115th-congress/house-bill/4625/text (accessed: 2021 August 23).

intelligence through specific technical methods and the results that can be obtained from them.

Secondly, it is not clear from the definition what are the signs of the results of human intellectual activity, how to make a comparison for sufficient intelligence of such systems. Moreover, self-learning and finding solutions without a predetermined algorithm are not human cognitive functions. For example, the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders lists the following as cognitive functions: integrated attention, perception and psychomotor function, learning and memory, speech, executive functions, and social intelligence[36]. These criteria are used, for example, in assessing the level of dementia. The concepts of "self-learning" and "search for solutions without a predetermined algorithm" belong to the field of information technology, outside the scope of which the definition of human cognitive functions is concerned.

Third, the related notion of "artificial intelligence technology" in this law is an example of a circulus vitiosus (vicious circle), a logical fallacy in which a statement is inferred from itself: "[A]rtificial intelligence technologies are technologies based on the use of artificial intelligence (including computer vision, natural language processing, speech recognition and synthesis, intelligent decision support, and advanced artificial intelligence techniques)". At the same time, it remains unclear what methods are advanced, this attribute is not disclosed, in contrast to the National Strategy.

Fourth, if we turn to foreign definitions in the bills under consideration today in the United States and Europe, we find that artificial intelligence is described through a set of specific methods that are used to create a technical solution. For example, in the European Union's proposal, "'artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches (i.e., machine learning, deep learning, logic, knowledge-based approaches, and statistical methods) and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with".

Thus, for the purposes of normative legal regulation, it is proposed to provide a refined concept of AI instead of the one borrowed from the National Strategy. At the same time, such notion should be based on a set of considerations that (1) reflect the attitude to the specific methods and approaches to the creation of AI to be regulated (machine and deep learning and/or statistical approaches and/or logical approaches), (2) take into account different types of AI classifications (hardware/software, inorganic/organic/mixed), and (3) contain a clear typification by the degree of its autonomy. Among other things, it is advisable to introduce the concept of "the

result of artificial intelligence", which can be understood as information, forecasts, recommendations and decisions, depending on the purpose and purpose of the system. In view of the necessity for the legal definition to be clear, concise and corresponding to the principle of formal certainty, such definition itself, however, must not necessarily be overloaded with the considerations that precede it – in contrast, it can be expressed in quite a short way.

Fifth, the National Strategy's definition of AI is based on a technological view of this technology; to be applied properly, this definition needs to be systematically interpreted in conjunction with several other definitions: technological solution, information and communication infrastructure, machine learning methods, etc. Some of these are set forth in the National Strategy, some are made in standards, and some are not legally enshrined in any social or technical regulations.

3. Proposals based on the research results

Taking into account all of the considerations provided above, it is acceptable to consider several possible options to improve the definition of AI. Thus, the first option, which can be provisionally characterized as "closed" is: "*Artificial intelligence [is] an information technology, which includes an artificial neural network*". That said, the term "information technology" is used in the meaning of the Federal Law of 27 July 2006 No. 149-FZ "On Information, Information Technologies and Protection of Information"[37]. This approach can be justified by the following methodological considerations: the object should be associated with the most general known object in terms of positive law; the assumption that the neural networks are the only currently significant technology, involving "self-learning"; in the definition a neural network should be specified, strictly speaking, as artificial (because natural one belongs to human beings).

At the same time, the main critical argument to this approach is its rigid connection to a particular technology, while the interpretation of the term denoting it can also vary. In this regard, the second option seems more promising, which can be conditionally described as "open": "*Artificial intelligence [is] an information system* [in the terminology of the Federal Law No. 149-FZ "On Information, Information Technologies and Protection of Information"], *which provides the function of forming algorithms for solving problems different from those laid down by the system developers, including the artificial neural network*". Main methodological comments: information system implies a database, which in technical sense is always present in practice (because it implies a data set for machine learning); neural networks are still not the only possible technology, correlated with the concept of AI, so descriptive functional approach is

---

[36] See: Diagnostic and statistical manual of mental disorders. American Psychiatric Association. 2013.

[37] See: "ConsultantPlus" legal reference system (in Russ.) (accessed: 2021 August 23).

preferred in the definition, which can "capture" the subsequent technologies.

In addition to the definition of AI, it seems necessary to consider a possible approach to the definition of a robot. Taking into account the research conducted, the principal methodological imperative is that the system of definitions for AI and robotics in this context is built for robots without a biological component (i.e., so-called "cyborgs" are beyond the scope of this study). Robots should also be distinguished from automata ("simple automata") characterized by the absence of uncertainty in the operation of algorithms. At the same time, it seems correct to connect the definition of robot with the object interacting with the material world, otherwise the category "robot" would be redundant in relation to the category of AI. The common usage of the word "robot" is not decisive in this case. For the purposes of this paper's proposal, the category "robot" should be associated with *physical component of AI control*. Under this approach, any robot implies AI, but not any AI implies a robotic component.

The challenges of making a legal construct of such concepts as artificial intelligence or robot result in the proposed definitions failing to fully cover all the peculiarities of such and carry a risk of being interpreted too broadly in implementation and in law enforcement[38]. It is important to decide which level of regulation, in the system of which branches of law and what kind of legal acts should contain such definitions.

Many researchers maintain the position that in Russia it is necessary to promptly adopt the corresponding federal law. Thus, it is proposed to develop a draft law "On the basis for regulating the use of artificial intelligence and robotics technologies"[39], where, along with the terminological base, it is proposed to include sections and chapters defining:

Principles of legal regulation, including principles of legal regulation of machine learning;

State policy and functions of the authorized body in the field of artificial intelligence and robotics;

(3) International cooperation and definition of the scope of the law;

(4) The legal status of entities in the field of creation and use of artificial intelligence and robotics technologies;

(5) Access and processing of confidential information, including for the purposes of machine learning and technology security;

(6) Identification of technologies and systems of artificial intelligence and robotics; (7) Fundamentals of standardization and development of self-regulation;

(8) Physical and information security;

(9) Liability of subjects of legal relations, including the basics of liability insurance.

At the same time, in the current conditions, it seems premature to endow AI systems and robots with limited legal personality until the appearance of a "strong AI". Nevertheless, defining these concepts in a Federal Law seems to be the most probable but not the only solution. A Federal Law could make a general definition to describe AI as a type of information system. The subvarieties could be described in substatutory documents. Perhaps the concepts of artificial intelligence and robot should be defined in technical standards, as those adapt faster to technological progress, whereas legal frameworks could simply refer to the applicable technical standards. Other solutions could exist, too.

REFERENCES

1. Definition Intelligence Artificielle dans les systems industriels de production, http://www.industrie-dufutur.org/content/uploads/2019/02/Intelligence-Artificielle-AIF-Janvier-2019.pdf (accessed: 2021 August 23).

2. Digital transformation: challenges to law and the vectors of scientific research / T.A. Polyakov, A.V. Minvaleev [and others]. Moscow, 2020 (in Russ.).

3. *Günther J. P.* Roboter und rechtliche Verantwortung. München (2016).

4. *Hallevy G.* When robots kill: artificial intelligence under Criminal Law. Boston (2013).

5. *Hunter D.* The death of the legal profession and the future of law // UNSW Law Journal 43(4), 1199−1225 (2020).

6. *Kalyaev I., Zaborovskij V.* Artificial Intelligence: from metaphor to technical solutions // Control Engineering Russia 83(5), 26−31 (2019) (in Russ.).

7. Legal and ethical aspects related to the development and application of artificial intelligence systems and robotic technology: history, modern status, and development prospects / ed. by V.B. Naumov. St. Petersburg, 2020 (in Russ.).

8. *Lipinsky D.A., Repeteva O.E.* On the inter-branch institution of legal liability // Yuridicheskaya Mysl' 4(96), 24−32 (2016) (in Russ.).

9. *McDermott D.* Artificial intelligence meets natural stupidity. SIGART Newsletter 57, 4−9 (1976).

10. *Pollack J.B.* Mindless intelligence // IEEE Intelligent Systems. (21)3, 50−56 (2006).

---

[38] See: Digital transformation: challenges to law and the vectors of scientific research / T.A. Polyakov, A.V. Minvaleev [and others]. Moscow, 2020; *Polyakova T.A., Minbaleev A.V., Krotkova N.V.* New vectors of Information Law development in the conditions of the civilizational crisis and digital transformation // State and Law, No. 5, 75−87 (2020).

[39] Legal and ethical aspects related to the development and application of artificial intelligence systems and robotic technology: history, modern status, and development prospects / ed. by V.B. Naumov. St. Petersburg, 2020.

11. *Polyakov A. V., Timoshina E. V.* General theory of law. St. Petersburg (2005) (in Russ.).

12. *Polyakova T. A., Minbaleev A. V., Krotkova N. V.* New vectors of Information Law development in the conditions of the civilizational crisis and digital transformation // State and Law, No. 5, 75–87 (2020) (in Russ.).

13. *Proudfoot D.* Anthropomorphism and AI: Turing's much misunderstood imitation game // Artificial intelligence 175(5–6), 950–957 (2011).

14. Report of Estonia's AI Taskforce, https://f98cc689-5814-47ec-86b3-db505a7c3978.filesusr.com/ugd/7df26f_486454c9f32340b28206e140350159cf.pdf (accessed: 2021 August 23).

15. *Scherer M. U.* Regulating artificial intelligence systems: risks, challenges, competencies, and strategies // Harv. JL & Tech 29(2), 353–400 (2015).

16. *Schuett J.* Defining the scope of AI regulations, https://arxiv.org/abs/1909.01095 (accessed: 2021 August 23).

17. *Simmler M., Markwalder N.* Roboter in der Verantwortung? // Zeitschrift für die gesamte Strafrechtswissenschaft. 129(1), 20–47 (2017).

18. *Susskind R. E.* Online courts and the future of justice. Oxford (2019).

19. *Turing A.* Can machines think? Moscow (1960) (in Russ.).

20. *Turner J.* Robot rules: regulating artificial intelligence. London (2019).

21. *Volkov V. V.* Artificial Intelligence and the human mind: futuristic synecdoche and reality (linguistic and linguo-mental aspects) // Vestnik Rossijskogo universiteta druzhby narodov. Ser.: Teoriya yazyka. Semiotika. Semantika 11(4), 745–759 (2020) (in Russ.).

22. *Whitby B.* The Turing test: AI's biggest blind alley? // Machines and thought: the legacy of Alan Turing (1), 53–62 (1996).

Сведения об авторах

АРХИПОВ Владислав Владимирович —
доктор юридических наук, доцент, заведующий кафедрой теории и истории государства и права Санкт-Петербургского государственного университета; 199034 г. Санкт-Петербург, Университетская наб., д. 7–9
SPIN-код: 1715-5893, AuthorID: 197523
ORCID: 0000-0003-4566-1454

ГРАЧЕВА Анастасия Вячеславовна —
преподаватель кафедры теории и истории государства и права Санкт-Петербургского государственного университета; 199034 г. Санкт-Петербург, Университетская наб., д. 7–9
SPIN-код: 1350-8296, AuthorID: 1025107
ORCID: 0000-0002-2320-8053

НАУМОВ Виктор Борисович —
доктор юридических наук, доцент, главный научный сотрудник сектора информационного права и международной информационной безопасности Института государства и права Российской академии наук; 119019 г. Москва, ул. Знаменка, д. 10; главный научный сотрудник Лаборатории прикладной информатики и проблем информатизации общества Санкт-Петербургского федерального исследовательского центра РАН; 199178 г. Санкт-Петербург, 14-я линия В.О., д. 39
SPIN-код: 5729-5413, AuthorID: 6101
ORCID: 0000-0003-3453-6703

Authors' information

ARKHIPOV Vladislav V. —
Doctor of Law, associate Professor, Head of the Department of Theory and History of State and Law of St. Petersburg State University; 7–9 Universitetskaya Embankment, 199034 St. Petersburg, Russia

GRACHEVA Anastasia V. —
lecturer at the Department of Theory and History of State and Law of St. Petersburg State University; 7–9 Universitetskaya Embankment, 199034 St. Petersburg, Russia

NAUMOV Victor B. —
Doctor of Law, associate Professor, chief researcher of the Information Law and International Information Security Sector of the Institute of State and Law of the Russian Academy of Sciences; 10 Znamenka str., 119019 Moscow, Russia; chief researcher of the Laboratory of applied informatics and problems of society informatization of the St. Petersburg Federal Research Center of the Russian Academy of Sciences; 39, 14th line of V.I., 199178 St. Petersburg, Russia

ПОЛЯКОВА Татьяна Анатольевна — доктор юридических наук, профессор, заслуженный юрист РФ, главный научный сотрудник, и.о. заведующей сектором информационного права и международной информационной безопасности Института государства и права Российской академии наук; 119019 г. Москва, ул. Знаменка, д. 10
SPIN-код: 4224-3174, AuthorID: 732015
ORCID: 0000-0003-3791-2903

POLYAKOVA Tatyana A. — Doctor of Law, Professor, Honored Lawyer of the Russian Federation, chief researcher, Acting Head of the Information Law and International Information Security Sector of the Institute of State and Law of the Russian Academy of Sciences; 10 Znamenka str., 119019 Moscow, Russia

МИНБАЛЕЕВ Алексей Владимирович — доктор юридических наук, доцент, заведующий кафедрой информационного права и цифровых технологий Московского государственного юридического университета им. О.Е. Кутафина (МГЮА); 123001 г. Москва, ул. Садовая-Кудринская, д. 9; главный научный сотрудник сектора информационного права и международной информационной безопасности Института государства и права Российской академии наук; 119019 г. Москва, ул. Знаменка, д. 10
SPIN-код: 7148-1527, AuthorID: 651824
ORCID: 0000-0001-5995-1802

MINBALEEV Alexey V. — Doctor of Law, associate Professor, Head Department of Information Law and digital technologies of Kutafin Moscow State Law University (MSAL); 9 Sadovaya-Kudrinskaya str., 123001 Moscow, Russia; chief researcher, Information Law and International Information Security Sector of the Institute of State and Law of the Russian Academy of Sciences; 10 Znamenka str., 119019 Moscow, Russia