Legal Issues in the
DIGITAL AGE

Вопросы права в цифровую эпоху

3/2023

Legal Issues in the
DIGITAL AGE

ARTICLES

Legal Issues in the **DIGITAL AGE**

Legal Issues in the **DIGITAL AGE**

"Legal Issues in the Digital Age" Journal is an academic quarterly e-publication which provides a comprehensive analysis of law in the digital world. The Journal is international in scope, and its primary objective is to address the legal issues of the continually evolving nature of digital technological advances and the necessarily immediate responses to such developments.

The Digital Age represents an era of Information Technology and Information Communication Technology which is creating a reliable infrastructure to the society, taking the nations towards higher level through, efficient production and communication using digital data. But the digital world exposes loopholes in the current law and calls for legal solutions.

"Legal Issues in the Digital Age" Journal is dedicated to providing a platform for the development of novel and analytical thinking among, academics and legal practitioners. The Journal encourages the discussions on the topics of interdisciplinary nature, and it includes the intersection of law, technology, industry and policies involved in the field around the world.

"Legal Issues in the Digital Age" is a highly professional, double-blind refereed journal and an authoritative source of information in the field of IT, ICT, Cyber related policy and law.

Authors are invited to submit papers covering their state-of-the-art research addressing regulation issues in the digital environment. The editors encourage theoretical and comparative approaches, as well as accounts from the legal perspectives of different countries.

Articles

Specialized Legal Language — Guided AI

Yuri Mikhailovich Baturin

Computer Law and Data Security Chair, Higher School of State Audit Department, Moscow State University, 1 Leninskie Gory, Building 13, Block 4, Moscow 119992, Russia, baturin @ihst. ru, ORCID: 0000-0003-1481-5309

Abstract

For legal regulation of behavior of artificial intelligence (AI), it is proposed, based on the structural similarity between the law and computer software, to make the legal profession a mandatory party to the design and development of artificial intelligence systems, with a special object-oriented legal language to be developed. In discussing the core elements of such a language, it is underlined that AI should be able to independently formulate and describe its purposes in the same object-oriented language to ensure feedback between AI and developers/users. It is demonstrated on the example of regulations and state standards adopted in Russia for driverless vehicles, that developing an AI-specific legal language is a complex task, including since contextual gradation is needed to formalize legal judgments. The emergence of a family of object-oriented legal languages is predicted. The issue of creating an AI theory designed to explain the data and facts to be handled by strong AI is raised. It is suggested to adjust the AI definition in the approved guidelines and strategies to describe AI as a system searching for solutions outside a preset algorithm but not excluding the use of algorithms altogether. The importance of algorithms for AI is demonstrated, with strong AI interpreted as systems guided by an object-oriented language. The differences between strong AI and man are analyzed. With regard to AI capable of responsible behavior, the internal representation of the outside world and itself is discussed for consistency of input data. It is concluded that inevitable conceptual, linguistic and practical problems to be faced by lawyers involved in the development of strong AI should not hold back the "juridification of AI design".

Background

The engineering and legal professions work in much the same manner, both following the established rules: while engineers assemble sophisticated products in a strictly prescribed order, lawyers apply the adopted provisions to social life. Both procedures are called algorithms. An algorithm is a sequence of actions to achieve a purpose (intended result). The structure of a provision — "if, then, otherwise," — is itself a basic (simple) algorithm (fig. 1) making up more complex ones which may be described by programming languages.

if (hypothesis)

otherwise (sanction)

then (disposition)
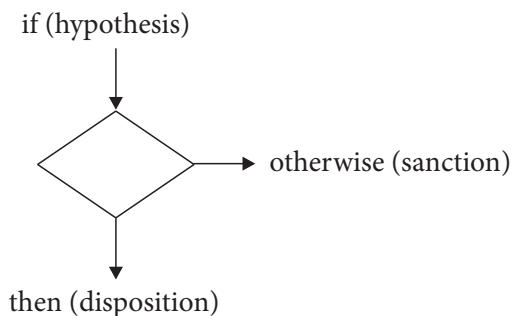
Fig. 1. Legal provision as a basic algorithm

While languages of technical and legal algorithms lexically differ, they structurally coincide. It would be wrong not to use this coincidence for the benefit of the legal science. The academician V. Kudryavtsev wrote more than fifty years ago: "The question of possible programming of the enforcement process is no longer a matter of controversy" [Kudryavtsev V.N., 1970: 69]. The achievements of that time have unfortunately sunk into oblivion together with the legal cybernetics. However, they are called for again with the development of artificial intelligence (AI) where one of the biggest problems is to feed a system of provisions into AI to be understood and complied with.

1. Goal Setting for AI

The systems based exclusively on algorithms are called automatic or simply automatons. They have long been established, the only practical problems involved in their development being related to the complexity of algorithms to be created, their consistency, feasibility etc. These systems are called weak AI as a tribute to today's fashion for artificial intelligence although, strictly speaking, an automaton has no intellect. Weak AI could be exemplified by autopilot systems (for cars and aircraft). Weak AI (automaton) is assigned a goal (internal purpose) achievable by executing a set of algorithms. Legal behavior is also algorithmized. Much is done today to "intellectualize" such automatons by training them how to "interpret" the environmental states (sometimes by posing certain problems to be solved by the autopilot), adjusting or setting a new goal ("intent") and assessing the expected result ("foresight"). (Quotation marks for the terms "interpretation", "intent" and "foresight" mean that they are not concepts of a theory of intelligent systems but metaphors or, legally speaking, legal analogies) [Baturin Yu. M., Polubinskaya S.V., 2022: 141–154]. This is already a step — but just one– towards creating strong AI or, more exactly, a strong AI-enabled robot. Such robots (autopilot systems etc.) should be designed to be "capable" of complying with legal provisions — in our example, traffic rules. Qualitatively, it is a more complex goal than in case of weak AI.

Weak AI is thus set a goal with an algorithm to achieve it. Strong AI will perceive a goal formulated in an object-based language defined in the developer's meta-language. If we want the legal profession to be involved in the development of strong AI, we should describe the object-based language to interact with AI in the legal (meta) language. Let's call it the specialized object-based legal language. The word combination "specialized legal" means a homomorphic image of legal language only partially reproducing the original language — that is, stripped down language preserving its structure and meanings to the extent sufficient to describe complex operations prescribed to AI. An AI developer, even a legal professional with the knowledge of the object-based language, can set a goal for AI. This is not hard to do. Creating the object-based language is much more difficult. Importantly, AI should itself be able to formulate and describe its goals in the same language. In particular, this is required for feedback between AI systems and their developers and users.

Obviously, lawyers can be involved in the development of strong AI only as part of a team of engineers, mathematicians, programmers and jurists,

all of whom understand their functional relationships in the process of designing, testing and applying AI. Thus, the object-based language will be a composite language, that is, only partially of legal specialization. As we have stated, engineering activities and jurisprudence are structurally described by one and the same language. As regards the lexical side, it is possible to compile a relatively comprehensive engineering-mathematical-legal dictionary suitable for well-designed and unmixed coordination of technical and legal approaches for full-fledged involvement of legal profession into the development of AI technology. However, we will deal here only with the specialized object-based legal language to show the lawyer's role and operating modes at the stage of AI development.

2. Artificial Intelligence and Algorithms

The most adequate definition of artificial intelligence is probably the one found in the National AI Development Strategy for the period until 2030[1] in which it is described as "a set of technological solutions allowing to mimic human cognitive functions (such as self-learning and search for solutions outside a preset algorithm) and address specific tasks with results at least comparable with those of human intellect" (paragraph 1.5a).

Meanwhile, one element of this definition — "search for solutions outside a preset algorithm" — is questionable. In fact, according to this definition, AI mimics human cognitive functions while human behavior is often algorithmic. Moreover, man finds himself in a "forest of algorithms" as soon as he is born — from baby breeding recipes to operating manuals and street crossing rules (look to your left before you step on the roadway; look to your right when you are in the middle of the road; or vice versa in countries with left-hand traffic). People sometimes follow algorithms automatically. It happens to everyone, even if deeply immersed in thoughts, to get off at the right stop, make the right turns and come exactly to the door of one's house. Such mechanical algorithmic behavior results from multiple repetition of a certain sequence of operations or from a fully and exactly defined goal.

Thus, the AI definition proposed by the Strategy should be amended to read "a search for solutions both on the basis of and outside preset algorithms". With this amendment, the AI definition becomes quite operable.

---

[1] Presidential Decree No. 490 of 10.10.2019 On Development of Artificial Intelligence in the Russian Federation (attached to the National AI Development Strategy for period until 2030) // SPS Consultant Plus.

Developing algorithms is a creative task of higher complexity than making arrangements for their execution as there is no one-size-fits-all solution. That work is fulfilling by human person for weak AI. Meanwhile, it is a complex task that strong AI is to be taught to handle. At the same time, there are algorithmically unsolvable (that is, inaccessible to AI) problems which hold for strong AI (impossibility to recognize self-applicability to its own code or self-inapplicability of normal algorithms, i.e., those that use letter–word strings as input data; non-feasibility of a Turing machine based on external alphabet *A* which would recognize whether an arbitrary Turing machine with external alphabet *A* is applicable to an arbitrary word expressed in *A* given that *A* contains at least two letters; problems that, if solvable, would result in the existence of paradoxical objects) [Krinitsky N.A., 1984: 76–80]. Algorithms will inevitably become part of strong AI.

Artificial intelligent robotic systems such as driverless cars will be used in variable environments which require "an ability to understand" such provisions as piloting parameter constraints and to take "reasonable" action to observe them as much as possible (quotation marks reflect the same reservation). However, not so accurate course of action is only possible if input data and intended results can be described in a language probably created for the purpose. Thus, chemical agents and their proportions are input data for a medical prescription (algorithm) we give to a pharmacist, while the result is the medication he makes as well as the dosage and periodicity of administration. Moreover, the patient understands only when and how many times the medication should be administered, the rest is written in the established Latin-based medical jargon, that is, the special language of the pharma industry. It is a similar specialized legal language that is dealt with below.

All language-guided systems are developed within normative boundaries of an object-oriented language. This means specific provisions to be fed into strong AI-enabled robots at the stage of development. Thus, the Guidelines for Regulation of Relationships in AI and Robotic Technologies until 2024[2] explicitly mentions as one of its purposes the establishment of "the principles for legal regulation of new social relations resulting from the development and application of AI and robotic technologies" (Section 1–2). This process is already underway with respect to social relations

---

[2] Government Executive Order No. 2129-r of 19.08.2020 attached to the Guidelines for Regulation of Relationships in AI and Robotic Technologies until 2024 // SPS Consultant Plus.

involved in AI. Meanwhile, the behavioral standards to be fed into AI is a blank spot. The Guidelines specifically address this task: "The development of AI and robotic technologies should rely on core ethical standards" (section 1–3). These standards are rightly called ethical as distinguished from legal provisions. It should be borne in mind, however, that the Guidelines mean AI ethics rather than human ethics in identifying some of them such as the priority of human well-being and safety, prohibition to cause harm to humans, human control, non-manipulation of human behavior and, finally, the provision directly related to the subject of this paper and explicitly addressed to the legal profession: "law-compliant development including compliance with safety requirements (the use of AI systems should not result in the developer-intended violation of legal provisions)" (Section 1–3). These provisions of AI ethics should be described in terms of an object-oriented legal language.

3. Creating an Object-Oriented Legal Language

The development of complex systems such as AI has caused a need to create special languages for natural description of the artifacts (objects) they incorporate, thus resulting in the emergence of object-oriented programming languages.

To behave responsibly in the outside world, AI should have an idea of this world expressed as input data. What does it mean for AI to have an idea? Know? Have information? Understand? These essentially philosophical questions cannot be answered unless a host of fundamental concepts — "knowledge", "opportunity", "action", "cause", "result", "situation" etc. — are formalized for AI to be able to ask and answer the questions such as: "How will the situation change if I trigger action X?"; "Do I have enough information to answer the previous question?" etc. The concept of knowledge is of principal importance. In the probabilistic, polyvalent or fuzzy logic underlying AI development, knowledge is stochastic (fuzzy) since the used judgments are only true with a certain probability. The concept of knowledge can lead to that of conviction (belief in something) that can be interpreted as perception of whether a judgment is true with a probability (fuzzy set function) of 1. AI's material convictions on the outside world could be regarded as embryonic "self-consciousness" (to be complemented by judgments on itself and its structure achieved through introspection, a subject far outside the scope of this paper). The general concept of AI's "self-consciousness" can be reduced to a sufficient number of "convictions"

(true judgments on the world outside and inside) relative to AI's own "convictions" and processes that bring about a change.

A specialized object-based language should include the terms and concepts used to formulate the requirements to AI. But "if the language set is too limited, many tasks will involve long and inconvenient structures", D. Stepulyonok wrote. "On the contrary, if syntax is excessively abundant, such language will be hard to implement. Apparently, the creation of language requires an acceptable balance in the number of language structures" [Stepulyonok D.O., 2010: 22]. Thus an object-oriented legal language is much dependent on AI's functional purpose. The below Program for experimental legal regime of driverless cars is a good example of initial approach to an object-oriented language for a specific task. Let's outline some basic elements of a specialized legal language.

The internal model of the outside world will provide AI's representation of it. The issue of AI's personality is pertinent if its model of the world is adequate in terms of understanding of the underlying mathematics, own goal setting, ability to ask and answer the above questions using this model and seek more information in the outside world as necessary. The task is far from simple: AI needs a mechanism for self-control of internal processes; an alphabet to designate and describe these processes, and a language describing the outside world in a way that the elements of AI's internal representation make up a system enabling a search of the goal and choice of a goal-focused action identifiable in the available set of possible actions, and a course of action. The available set of possible actions will require imperative structures using conventional operators and cycles. These actions will be described by statements in the chosen object-oriented language using symbols, descriptive and modal operators, internal parameters etc. Thus, AI's representation of the outside world, goal, strategy (actions required to achieve it) are expressed linguistically.

Under such approach, we could acceptably "roughen" the human understanding of "free will" as an ability, decide on the course of action by assessing the result of various possible actions, and accept that "free will" for AI is the ability (based on strictly formalized concept of "can") to make up a list of alternatives for achieving the set goal and to choose one or several of them.

To describe one of the vital actions, let's introduce the concept (verb) "can" and select just one meaning — be able — out of the whole variety of meanings (including legal) which would reduce the legal language but offer

the meaning appropriate to internally decide what to do. Nothing hampers us to use other legal meanings of "can" — such as to be entitled, to be capable, to have an opportunity etc. — through the graduation of contexts.

The causal link can be naturally expressed by the concept (verb) "cause" to mean the relationship of "resulting", "entailing", "causing". The said relationship (causal link) follows from the context described in terms of the variables called "conditions".

A "situation" means for AI the state of the outside world at the moment t. Since the world is too big, it can never be exhaustively described, with its state perceived by AI via conditions and "facts" to be interpreted as "true events". Facts will be used to derive new facts relative to the given situation, as well as make judgments on any prospective causally linked and hypothetic situations, such as the one where an AI-enabled robot helping persons with disabilities around the house has accepted an order to get the moon on a stick. This hypothetic situation is not fully defined as it is not clear what exactly the robot has "in mind", that is, in its decision block (a robot is unlikely to be trained to deal with the moon, stars etc.). But this representation of a situation could be useful for analysis of a set of facts which would be sufficient to understand why the AI-enabled robot has attempted to get the moon on a stick and whether it will give up and why. Such situations can be internally represented for AI in terms of symbolic expressions translatable under the prescribed rules.

The concept of "result" is causally linked to the performance of an action. If an action does not lead to any result, the value of this variable becomes indefinite. Importantly, an "action" intended by AI is not necessarily the one to be performed in reality. Therefore, we can only approximately speak of an action bringing about a certain situation. Hence, the concept of "result" cannot be considered definite in the outside world. It is definite and preferential for AI only in its representation of the outside world. This is one of the reasons why AI can cause harm, undesirable incidents etc.

Actions will form into strategies, the most basic one being a finite sequence of actions. A cyclic repetition of actions, a strategy with interrupted action and priorities etc. are possible [McCarthy J., Hayes P., 1972: 52–54, 58, 62].

Suppose it is a need to trace a route for an AI-enabled driverless car. To have a goal achievement strategy (the well-known phrase "the route is traced"), AI needs to analyze the situation described by several types of

"facts": topographic facts (coordinates of the start and end points, number of blocks to pass, number of turns and their directions); facts indicating the effect of intended action (for instance, road under repair after the second turn to the right); finally, the fact that the street in question will be reached ("result") upon completion of a cyclical sub-program corresponding to a set number of blocks to be passed and turns to be made. The last fact does not assume a possibility ("can") of starting the trip. The possibility to achieve the goal is affected by "knowledge" described in terms of predicate logic. Using the concept "can", AI should be able to demonstrate it "knows" alternative goal achievement routes and to specify the route selection criteria (minimum time spent, minimum route length, absence of road jams etc.).

The complexity of creating a specialized legal language for weak AI even in a simplified situation can be seen on the example of driverless car regulations.

4. Elements of a Specialized Language for Driverless Cars

On 17 October 2022 the Russian Federation Government has adopted Resolution No. 1849 to approve the Program of experimental legal regime for digital innovations to operate intelligent vehicles under the driverless logistical corridors initiative for M-11 Neva federal highway.[3] In this context, a driverless car in terms of our terminology is just weak AI. However, for lack of a similar document for strong AI, let's assume an external lawyer to be a model of internal normative block for would-be strong AI. The said Program shows the interaction between the lawyer in question, driverless cars and users. For instance, it is stated that "unless provided for by the operating algorithm, no third party may interfere with the operation of an automatic driving system" (paragraph 88 "d"), something to be compensated by "a diagnostic system for real-time performance monitoring of the intelligent driving system" (paragraph 88 "c"). It is also envisaged that "the driving system should be able to bring the intelligent vehicle to a safe stop" (paragraph 88 "e"), etc. Meanwhile, the required safety level equally depends on "the intelligent vehicle's controller" — from a test driver to test engineer — who should exercise "supervisory monitoring" along the route

---

[3] Government Resolution No. 1849 of 17.10. 2022 (attached to the Program of experimental legal regime for digital innovations to operate intelligent vehicles under the driverless logistical corridors initiative for M-11 Neva Federal Highway as amended by Government Resolution No. 607 of 17.04. 2023 and No. 1206 of 08.08.2023) // SPS Consultant Plus.

and "in manual driving mode" (paragraph 2). Thus, both the safety technology and control system partially depend on human control to identify a malfunction. Therefore, the safety technology and human control are part of a higher level safety system. Strong AI will be likewise subject to human control in the future, at least via the normative block designed by lawyers.

Growing automation actually makes the man-machine relationships more complex. The Program provides for different levels of automation:

"intelligent vehicle" equipped with "automatic driving system, that is, software and hardware for non-assisted driving (without presence of a test driver);

"1st category intelligent vehicle" with a test driver in the driving seat;

"2nd category intelligent vehicle" for non-assisted controller-supervised driving (with a test engineer on board but not acting as a (test) driver).

The four key variables are: human and/or automatic driving operation; human or automatic control of the traffic situation (road conditions); possibility (impossibility) for a human operator to override the non-assisted automatic system; and possibility (impossibility) for the automatic system to operate under all or some traffic situations. These variables can take specific values at once (possibility to instantly cancel decisions), with a delay (requiring some time) or in a mediated way (through a controller). The choice of value in the first situation is obvious as envisaged by paragraph 17 "f": "The test driver should instantly assume driving by taking control of the intelligent vehicle to prevent a traffic accident". The second variable depends on the dynamic digital traffic map, a part of the intelligent traffic system based on a geo-information road and traffic model for higher situational awareness of vehicles in an automatic mode. The third variable concerns technical malfunction and third-party deliberate intervention (paragraphs 86 "b" and "c"). The fourth variable can take specific value depending on "circumstances that make the intelligent vehicle's driving impossible or unsafe" (paragraph 2). The values of these four variables become important, for example, when the automatic system can respond faster than human operator or when the driverless vehicle's automatic system operates in coordination with other systems (such as the infrastructure of the intelligent vehicle operator or driverless cargo transportation controller), or when the driver's commands are incompatible with real constraints of the driving route.

While representing a descriptive taxonomy, the above examples from the Program of experimental legal regime for driverless cars pose complex questions.

Should the test driver in the driving seat ("1ˢᵗ category intelligent vehicle") keep at least one hand on the wheel? Can the driver in the front passenger seat (paragraph 17 "b") maintain the level of attention presumed to ensure safety of the trip in non-assisted driving? Under paragraph 17 "e" the test driver should keep monitoring the traffic situation while the automatic driving system is in operation (in particular, no telephone could be used during driving except with a hands-free kit).

Can a test engineer on board of the vehicle under full automatic control ("2ⁿᵈ category intelligent vehicle") reasonably assess the traffic situation and adequately react even with a test driver in the front passenger seat waiting for an order to help or take control (paragraph 17 "f")?

Will the automatic system reliably respond in the event of extraordinary driving conditions –worse road grip, fire, smoke, adverse weather conditions such as strong wind, heavy precipitation (paragraph 58)?

These questions are not only about safety as such but equally about values (human life, damage to property) that support or clash with this idea. Importantly, there is no common understanding of safety either from the technical or legal point of view. The assessment of safety will necessarily include the assumptions of the extent of damage, timing and causal links.

Let us discuss, for example, to what extent the "driver–car" pair should be safe. Suppose it should be required to operate as reliably as an experienced driver would in any imaginable maneuver or traffic situation. Such a strict standard implies, however, that driverless cars will be marketed at a slower pace and higher cost. Lower requirements will result in accidents, loss of life and confidence in the AI technology. Therefore, we need to analyze the costs and likely damage, on the one hand, and benefits from driverless vehicles on the other hand. The fruits of that analysis can be impacted by possible restrictions or wrong goal setting. For instance, a road accident inevitable at a given speed could be prevented at a lower speed while an attempt to protect passengers of a driverless vehicle by increasing its weight could put pedestrians at risk if the vehicle runs them down.

Safety can be defined as a guaranteed protection from the risk of harm. The Program under discussion has two sections dedicated to risk: "X. Assessment of the risk to life, health or property of individuals, property of legal persons, national defense and/or security or other values protected by federal law" (paragraphs 85-87) and "XI. Policies to minimize the risks specified in Section X…" (paragraph 88). These risks "result from the likeli-

hood of traffic accidents involving intelligent vehicles" (paragraph 86). The wording is correct but not adequate for the purpose of an object-based language. Let's use a stricter definition from the mathematical risk theory: risk is an aggregate value of possible damage in a stochastic situation of certain probability [Korolev V. Yu. et al., 2007: 9]. It is this approach that is used in the Guidelines for Regulation of Relationships in AI and Robotic Technologies until 2024 whereby "specific regulatory decisions need a risk-oriented approach based on the assessment of potential damage to the said values at a given probability against potential positive effect from the introduction of AI and robotic technologies, as well as policies to minimize the relevant risks" (Section I-4).

The guidelines provide for a mandatory and well-founded assessment of risk of AI-related damage and for the adoption of restrictive provisions if the use of AI technologies involves an objectively high risk of damage to the parties to social relationships. Where necessary for establishing specific provisions, the Guidelines suggest to use the definitions contained in standardization documents (section II-6). Since 1 January 2023, the Federal Technical Regulation and Metrology Agency has introduced eight standards for AI-enabled driverless cars which will be indeed useful when formalizing an object-based legal language for the development of AI-enabled driverless cars, primarily for terminology (GOST R 70249-2022),[4] but also for the requirements to road obstacle detection algorithms (GOST R 70251-2022),[5] testing requirements to road sign identification (recognition) algorithms (GOST R 70255-2022),[6] crossroad structure detection and reconstruction algorithms (GOST R 70253-2022),[7] roadside and traffic lane control algorithms (GOST R 70256-2022),[8] road user behavior prediction algorithms (GOST R 70254-2022)[9] and low-level data merge algorithms

---

[4] GOST R 70249-2022 AI-enabled road transport systems. Intelligent vehicles. Terms and definitions. Moscow, 2022.

[5] GOST R 70251-2022 AI-enabled road transport systems. Vehicle driving systems. Test requirements to obstacle detection and recognition algorithms. — Idem.

[6] GOST R 70255-2022 AI-enabled road transport systems. Vehicle driving systems. Test requirements to road sign detection and recognition algorithms.

[7] GOST R 70253-2022 AI-enabled road transport systems. Vehicle driving systems. AI-enabled road transport systems. Vehicle driving systems. Test requirements to crossroad detection and reconstruction algorithms.

[8] GOST R 70256-2022 AI-enabled road transport systems. Vehicle driving systems. Test requirements to roadside and traffic lane control algorithms.

[9] GOST R 70254-2022 AI-enabled road transport systems. Vehicle driving systems. Test requirements to road user behavior prediction algorithms.

(GOST R 70252-2022).[10] Finally, GOST R 70250-2022[11] sheds light on the above issue of safety of AI-enabled driverless vehicles by mentioning, in particular, "a standardized structured language to describe traffic scenarios" (paragraph 7.1.3). Safety of automated driving systems is addressed by ISO 22737, first international voluntary standard.[12]

The making laws and its enforcement are not simple even in the classical form, not to mention making a robot comply with provisions fed into its "brains" and written in a specialized legal language, some examples of which have been provided in this section. According to U.S. authors involved in automation of law enforcement activities, "law is rarely written with such algorithmic precision in mind". Worse still, it is not drafted with a view to be fed into AI's memory. Law is not always straightforward and has to be interpreted, sometimes adjusted. H. Surden, U.S. professor of law, is right when he says: "Automated legal reasoning systems that exist operate within particular legal contexts in which legal decisions tend to be relatively more determinate", only to become dispositive since in the given context the variability of meaning is extremely low. He notes a widespread skepticism of the legal profession about computerization of law: "Scholars from the legal domain tend to insist upon a nuanced view of legal analysis. In this conception, legal reasoning is too imbued with uncertainty, ambiguity, judgment, and discretion to permit computerized assessment. This literature's common theme is that even if computers were technically able to mimic legal decision making in a mechanical fashion they would necessarily miss the subtle institutional, value-based, experiential, justice-oriented, and public policy dimensions that are the heart of lawyerly analysis".

While recognizing that computerization of the legal process is a complex task, Surden, however, says: "In comparative terms, the number of legal contexts in which legal outcomes are tolerably determinate is probably somewhat small." [Shay L., Hartzog W., Nelson J., Conti G., 2016: 276–277].

---

[10] GOST R 70252-2022 AI-enabled road transport systems. Vehicle driving systems. Test requirements to low-level data merge algorithms.

[11] GOST R 70250-2022. AI-enabled road transport systems. Application options and composition of functional AI sub-systems.

[12] ISO 22737. International standard. Intelligent transport systems. Low-speed automated driving for predefined routes. Performance requirements, system requirements and performance test procedures. Available at: URL: https://www.novotest.ru/news/world/standart-iso-22737-na-nizkoskorostnye-sistemy-avtomatizirovannogo-vozhdeniya/ (accessed: 15.09.2023)

It means that formalization of legal judgments will require a context-based gradation, that is, a range of meanings. In this way, a specialist to feed provisions into AI's memory will be able to use different levels of conceptual abstraction. That allows to characterize the extent of certainty of a provision across multiple legal contexts which is useful for the development of strong AI.

5. Developing Theory of Artificial Intelligence

If we count out Ramon Llull, Spanish mathematician and philosopher [Gilson E., 1992: 18], who attempted back in the 13th century to create a logical problem solving device on the basis of his own system of concepts, Gottfried Leibniz [Leibniz G., 1984: 412] and Rene Descartes [Descartes R., 1989: 256–262], who proposed in their works universal languages for classification of concepts, artificial intelligence (though called otherwise at the time) dates back to Norbert Wiener and his already classical book on Cybernetics [Wiener N., 1983]. It was followed by Alan Turing's equally famous paper "Computing Machinery and Intelligence" first printed in 1950 [Turing A., 1960]. The same term "artificial intelligence" first made its appearance in 1956 at a Dartmouth College workshop (United States), only to be wrongly translated then into Russian as "intellect" although "intelligence" means just the "reasoning ability". In short, it was about "artificial reason" but not about "intellect". Author of the article present will further use the established notion of "artificial intelligence" in the meaning of a device with the reasoning ability.

Artificial intelligence has developed with practice and theory taking turns to outstrip each other. At present, major achievements in this field are rather backed by the development of high performance devices than the evolution of theory, sill to catch up with practice. U.S. specialists even assert, probably too pessimistically, that "there is no generally acceptable concept of automatic enforcement, not to mention common theoretical framework to guide the introduction of the relevant systems" [Shay L., Hartzog W., Nelson J., Larkin D., Conti G., 2016: 272].

We have referred above to the definition of artificial intelligence proposed by the National AI Development Strategy for the period until 2030[13]

---

[13] Presidential Decree No. 490 of 10.10. 2019 "On the Development of Artificial Intelligence in the Russian Federation" (attached to the National AI Development Strategy for the period until 2030) // SPS Consultant Plus.

where AI is described as "a set of technological solutions allowing to mimic human cognitive functions" (paragraph 1.5a). Man is normally believed to be able to interpret and conceptualize specific actions and predict their result. Therefore, to mimic man in view of all his other cognitive functions, AI should be regarded as a device capable of interpreting, intending and predicting. In other words, AI should be primarily able to handle the changing states of the outside world, with some to be interpreted as posing specific problems. To solve them, AI will take a course of action with a predictable result to remove the problem.

Theory starts with selecting a model with predictive properties. To identify and predict a future event, it should be interpreted and aligned with a set of real input data as the theoretical foundation. To correctly interpret and make the right choice from a multitude of possible models, the functional purpose of AI should be known, otherwise there is no telling whether interpretation is reasonable (suitable for the functional purpose) and what properties of the outside world are important for AI. In defining or assigning AI's purpose, we should describe AI in a kind of meta-language. The external purpose is thus "internalized" as AI's systemic goal. In this case, artificial intelligence should be provided with a mechanism that will translate the goal into actions to achieve it. It should be noted that AI, as part of the overall goal (purpose), will perceive certain sub-goals from man and even set for him certain goals (multi-purpose operating mode which allows to consider AI as an evolving, self-organizing system). Moreover, AI will be inevitably integrated into a context where it becomes intelligent (for simplicity's sake, we assume that man is a thinking being).

Let's distinguish two types of AI: silent and language-guided. The former is associated with weak AI (or AI systems). The latter relies on the concept of object-based language described in the aforementioned meta-language for AI to perceive an externally defined goal and describe its current sub-goals. An externally defined goal (external purpose) of weak AI equals its internal goal ("internal purpose"). On the contrary, a goal for strong AI could be set by everyone who knows the object-based language. Moreover, strong AI is able to formulate its own goal explainable in terms of the same object-based language.

A special-purpose processor (such as autopilot) is silent weak AI. Strong AI incorporating software for processing the statements expressed in a programming (object-based) language, as well as compiling, interpreting and other software, is language-guided AI (further referred to as LGAI). Weak

AI cannot be assigned a new goal, nor can it develop its own goals. Despite being able to interact with the pilot, the autopilot relies on a system of symbols and messages that is not exactly a language. The instructions to the pilot to change the direction will only change the goal's parameters (or the path to reach it). Meanwhile, LGAI already has the goals it is able to describe or can perceive a new external goal. The notion of "reason" is incompatible with weak AI while that of "intelligence", as was observed above, is rather a convention.

Let's separate the processes occurring in LGAI from the system in which they are organized to take place. Processes amount to the emergence of organized interaction in the interpreting system with a certain manifestation of intelligence. (In terms of mimicking man, the interpreting system itself should be associated with human mind).

Developing an AI theory essentially amounts to generalizing the process of interpretation across multiple models instead of the chosen one. Two theories thus emerge from what we have said: that of weak AI and of strong language-guided AI. While the former has long been known (automatic management theory in the field of technology, automata theory in mathematics), we are concerned here only with LGAI theory.

To identify and formulate an AI theory is to find a management pattern that explains the data and facts making AI operational. Therefore, developing a theory is to manage management (meta-management). The object-based language is used for LGAI as a "management managing" meta-language, with the language speaker (developer) to become actively involved in building the theory. It is thus obvious that no AI theory is possible without the involvement of legal profession.

To sum up this sketch of AI theory, let's note what makes AI so different from man. In the section "Developing an object-based legal language", we briefly mentioned a need in introspection which means LGAI's ability to look at itself. A self-developing nature prompts LGAI to review goals. In this case, goals depend on normative restrictions to be introduced by the lawyer. In this light (and in this light only) it is useful to discuss what makes LGAI different from man since man can (and knows he can) disobey. It would seem at first that this property should be ruled out for would-be LGAI to reduce the risk of harm to man. However, man can make mistakes that LGAI will strive to correct (let's recall Isaac Asimov's robotics laws [Asimov I., 2008]). Legal scholars are well aware of the principle of "waiver" which allows to waive someone's liability for harm to avoid bigger harm

(here we go back to the problem of risk discussed in the section "Elements of specialized language for driverless cars" in relation with the assessment of the extent of risk created by driverless vehicles). At the onset it would be probably unwise to allow LGAI to violate legal provisions.

Further point. Man realizes that the process he is part of evolves with time. Unlike AI, man will set goals (or formulate tasks) rather than choose them from a finite list as was noted in the section "Developing an object-based legal language".

The third difference is that man does not set goals in an absolute clear and consistent way due to the ambivalence of all natural languages. But it is precisely this characteristic that spurs up discussions and disputes as well as social, scientific and technological development. In other words, man will normally strive towards a fuzzy goal achievable in an unlimited number of ways. In its turn, an object-based language, more accurate and strict, forces AI to act with high certainty.

The fourth difference is that LGAI "knows" when the pursuit of a goal relies on preset algorithms or not whereas man is unaware of automatic action, that is, he is mainly aware of the goal-setting and problem-solving process. While man can be forced to realize his automatic actions (for example when asked to describe them), it will only result in slower execution and errors. Running down the stairs without thinking, you will slow down and even misstep, should you be asked to describe the successive movement of legs and feet or parameters of the staircase. Human consciousness turns on when at least two processes are performed at a time while regulation of automatic action is sub-conscious [Pask G., 1972: 19–20, 23].

Thus, the concerns that strong AI will surpass human intellect are not quite reasonable: AI and man "think" differently and are only comparable in terms of limited criteria such as problem solving speed, novelty of found solutions, inherent risks, legality, morality etc.

Conclusions

Like any object-based language, the specialized legal language has a certain history (that extends from the aforementioned GOSTs to this paper). To introduce new concepts, definitions of new terms should not contradict those of the earlier terms. Logical judgments will be thus restricted by ones previously used. As regards programming, such restriction is a prohibition to use an identifier (software assigned name for variables) until it has been

described. By introducing primary elements of the object-based language, the history will determine new elements to be invented by its developers.

An object-based language to feed fundamental provisions into AI is likely to embrace not only legal and technical elements but also those of other professional languages. There will probably be several object-based languages depending on AI application. Such languages could be more conveniently described as a family or *koine* of object-based languages (from Greek κοινὴ διάλεκτος or common dialect) which in social linguistics means a communication tool for a community of people (related language speakers) speaking in cognate tongues, "a non-native language to anyone of the communicants but quite "normal" from the perspective of structural complexity and therefore capable of serving an unlimited range of communication purposes" [Bagana Zh., Khalipina E.V., 2009: 19]. Importantly, there should also be a written form of such language.

The proposed way is not easy. But the conceptual, linguistic and practical problems to be faced by legal professionals along the way should not hold back the "juridification" of strong AI development. Lawyers and engineers will be able to understand each other and develop a specialized legal language (or more exactly, dialects for different AI applications). It will undoubtedly help AI to "understand" humans better. Legal profession should become a legitimate party to the process of AI design and development.

References

1. Asimov I. (2004) Laws of robotics. Moscow: Eksmo, pp. 781–784 (in Russ.)

2. Bagana Zh., Khalipina E.V. (2009) The role of language mixing in shaping global culture. *Nauchnye vedomosti*=Research Bulletin, no. 14, pp. 18–22 (in Russ.)

3. Baturin Yu.M., Polubinskaya S.V. (2022) Artificial intelligence: legal status or legal regime? *Gosudarstvo i pravo*=State and Law, no. 10, pp. 141–154 (in Russ.)

4. Descartes R. (1989) *Treatise on method. Rules for the direction of the natural intelligence.* In: Works in 2 vols. Moscow: Mysl Publishers, vol. I, pp. 256–262 (in Russ.)

5. Gilson H. (1992) Reason and revelation in the Middle Ages. Theology in the medieval culture. Kiev: University, p. 18 (in Russ.)

6. Korolev V.Yu. et al. (2007) The mathematical foundations of the risk theory. Moscow: Fizmatlit, p. 9 (in Russ.)

7. Krinitsky N.A. (1984) Algorithms are around us. Moscow: Nauka, pp. 76–80 (in Russ.)

8. Kudriavtsev V.N. (1970) Heuristic methods of crime qualification. In: Legal Cybernetics. Moscow: Nauka, p. 69 (in Russ.)

9. Leibniz G. (1984) *A history of the idea of universal characteristic.* Moscow: Mysl Publishers, p. 412 (in Russ.)

10. MacGarty J., Hayes P. (1972) Philosophical problems from the point of artificial intelligence. In: The cybernetic problems of bionics. Synthesis of models and engineering aspects. Moscow: Mir, pp. 40–88 (in Russ.)

11. Pask G. (1972) The meaning of cybernetics in the behavioral sciences (the cybernetics of behavior and cognition; Extending the meaning of "goal"). In: The cybernetic problems of bionics. Synthesis of models and engineering aspects, pp. 9–39 (n Russ.)

12. Shay L., Hartzog W., Nelson J., Larkin D., Conti G. (2016) Confronting automated law enforcement. In: Robot Law. Cheltenham: Edward Elgar Publishing, pp. 235–273.

13. Shay L., Hartzog W., Nelson J., Conti G. (2016) Do robots dream of electric laws? An experiment in the law as algorithm. In: Robot Law, pp. 274–305.

14. Stepulyonok D.O. (2010) The model and implementation methods of object-based languages. *Komp'yuternye instrumenty v obrazonanii*=Computer Means in Education Process, no. 4, pp. 21–29 (in Russ.)

15. Turing A. (1960) Do machine think? Moscow: Fizmatlit, 67 p. (in Russ.)

16. Wiener N. (1983) *Cybernetics, or Control and Communication in the Animal and Machine.* Moscow: Nauka, 340 p. (in Russ.)

Information about the author:

Yu.M. Baturin — Doctor of Sciences (Law), Professor, Corresponding Member, Russian Academy of Sciences.

"Artificial Intelligence": Problems of Civil Law Qualification

Liudmila Yurievna Vasilevskaya

Kutafin Moscow State Law University, 9 Sadovaya-Kudrinskaya Str., Moscow 125993, Russia, ljvasilevskaja@msal.ru

Abstract

Based on the civil law research methodology, the paper provides an insight into the concept of "artificial intelligence", its legal nature and peculiarities of transactions. The subject of research is the underlying doctrine, legal provisions on "artificial intelligence" and their enforcement practices. The research purports to lay a theoretical groundwork for the concept of "artificial intelligence" as a complex thing at law whose structure may comprise a variety of intellectual assets. The applicable civil law regime — specifics of formalization of the underlying relationships — is determined by the legal nature of "artificial intelligence" as an ideal phenomenon. In view of the complex nature of the object in question, the structural system and comparative law methods were used in the paper. In fact, the use of the structural system method allowed not only to analyze "artificial intelligence" as a complex structure but also to identify computer software as its core structural element. The comparative law method enabled to develop an idea of regulatory mechanism for "artificial intelligence" as well as to demonstrate the specifics of interpretation of the applicable provisions of intellectual property law. The paper also makes use of special methods of inquiry such as the logical and formal methods, with the latter allowing to define the concept of "artificial intelligence" and discuss its core legal features. As a methodological peculiarity, the study combines the theoretical and empirical levels of cognition. The use of the methods mentioned allowed to explore raised legal issues of "artificial intelligence" as they relate to the foundations of civil law. It is concluded that the main frequently used contractual arrangements to dispose of the exclusive right to "artificial intelligence" include the exclusive right transfer agreement and the licensing agreement. The paper provides an analysis why an exception from the general rules applicable to exclusive right transfer agreements and licensing agreements were made for "artificial intelligence".

Background

It is not accidental the problem of "artificial intelligence" has recently come under scrutiny in civil law studies. In the context of a large-scale economic digitization, "artificial intelligence" has become tradable, only to pose a number of practical questions to businessmen with answers expected primarily from civilian lawyers. What kind of agreements should apply to the creation of "artificial intelligence"? What transactions are allowed for its use? Is "artificial intelligence" pledgeable or heritable? Could it be transferred under the rules developed for physical objects? These questions cannot be answered unless the doctrine unambiguously defines the legal nature of this phenomenon, with the legislator enshrining the adopted stance in specific provisions. Unfortunately, despite a wealth of literature on the subject, there is yet no clear idea of "artificial intelligence" either in civil law or among researchers. A retrospective review at the problem allows to identify a number of methodological inaccuracies and failures behind wrong conclusions.

The first thing that calls attention is the methodological defect of most publications on "digital matters". Many authors consider the digital environment and digital data on technical devices and their systems to be ideal ones and essentially in opposition to the material world. This is wrong and contrary to the principles of the philosophical methodology on the difference between the material and the ideal. It is a vulgarization and a mistake to believe that the materiality boils down to the physical, corporeal world represented by specific things existing in time and space. The world outside physical things — including digital data and digital environment as a whole — is thus considered to be ideal. From the perspective of philosophical methodology, matter is not only a physically perceivable world but also a world of field structures, viruses and bacteria just as that of semiotic systems (signs, words, symbols, numbers). An ideal world (subjective reality, inner world) is comprised not only of human feelings and percep-

tions but also of thoughts, images and ideas. This ideal world will never be perceivable by a third party unless it is embodied in a material form. To be perceived by others, these forms may be visual (such as text and pictures), sonic, tactile or gustative. The forms allowing to perceive an ideal world (ideas, representations, concepts, intellectual assets, information) as a flow of numbers will differ for different devices and their systems (computers, gadgets, servers etc.). The digital (electronic) form is material just as any other (corporeal, sonic, gustative). Therefore, "artificial intelligence" visualized in a digital form on a device does not assume non-materiality (ideality). As an example of erroneous judgment, K.M. Mefodieva, in discussing the attributes of digital data, underlines their "non-material, digital form — source code in a numeric form" [Mefodieva K.M., 2019: 10, 12]. V.A. Laptev and P.I. Usenkov make a similar mistake when they suggest "digital things at law represented and expressed in an electronic form… can be embodied in a physical form" [Laptev V.A., Usenkov P.I., 2020: 76]. While an electronic form is essentially physical, the authors obviously ignore a principal difference between the ideal content and physical representation of a thing ideal in its essence.

There is no fundamental study on the subject. The AI publications by Russian researchers will often only echo the economic and technical literature in the English language. This negative trend results from the fact that, on the one hand, a large part of the foreign literature in the field is in English and, one the other hand, a review of foreign literature poses objective constraints: since researchers in Russia largely have the knowledge of English, they choose publications they can translate on their own without incurring significant costs of professional translation services. This means that a majority of authors become captivated by the approaches that dominate in countries of the Anglo-American legal doctrine. In particular, this is manifested in the use of economic analytical tools to study the legal aspects of digitization processes and "artificial intelligence". Without going into details of economic analysis of law, it is noteworthy that a negative impact of its propagation affecting the legal analysis of economy was primarily visible in the fact of using the concepts of economic and technical sciences to provide a legal description of many digital assets including AI. Such a metaphysical approach inappropriate in any branch of knowledge has brought about negative implications for legal studies of "artificial intelligence". Instead of competent civil law analysis of "artificial intelligence", the literature is crippled with numerous publications demonstrating a departure from the academic principles of civil law. For example, many authors have come

to discuss legal personality of "artificial intelligence" [Kuteinikov D.L. et al., 2019: 85–95]; [Dremliuga R.I., Mamychev A.Yu., Dremliuga O.A., Matyuk Yu.S., 2019: 127] while ignoring the absolute truth of civil law that only persons have a legal personality and that only the legislator can grant rights and not academics.

Moreover, AI's primitive imitation of human cognitive functions does not at all mean "artificial intelligence" can approach natural human intelligence in terms of functions, purpose and evaluation of *its* outcomes just by virtue of legal fiction allowed by the legislator. "Artificial intelligence" is a convention to be put, in our view, in quotation marks. Humankind is not anywhere close to recognizing AI a person at law[1]. For example, according to M. Kovalchuk, President of the National Research Center Kurchatov Institute, AI has "nothing to do with intelligence", "this just means a higher amount of computations, a cloud"[2]. The discussions to endow AI with a legal personality and delictual dispositive capacity follow from various reasons (economic, social, political, academic), the main (primary) reason being economic — financial, pecuniary interest of large businesses prompting major high-tech companies to search for new mechanisms of boosting consumption of innovative products for domination and control of production and sales markets, and to identify new ways and methods of minimizing liability vis-à-vis innovative product users.

Another wrong and no less extreme stance is assumed by the authors who identify "artificial intelligence" with its physical medium and thus extend to it the regime applicable to things at law. Their "definition of artificial intelligence as an AI-enabled thing" is puzzling [Somenkov S.A., 2019: 75]. The inevitable result of this mistake is confusion of civil law regimes applicable to things and items of intellectual property. As an intellectual asset, AI is essentially ideal and needs to be represented in an objective (physical) form to be perceived by others. As was noted above, ideal products existing on different devices and their systems have a digital (electronic) form, something that allows to represent any intangible asset not in a corporeal form (that of a printed text, scheme or image) but in another physical (digital) form such as signs, numbers, source codes existing in special data media as a modern way of storing the ideal outcomes of human activities (back up systems). For "artificial intelligence", computer is a physical medium that has

---

[1] Available at: URL: https://newizv.ru/comment/2/15-02-2017/252046-evgenij-kasperskij-nikakogo-iskusstvennogo-intellekta-poka-net (accessed: 18.07.2023)

[2] Available at: URL: https://nauka/17398623 (accessed: 19.07.2023)

all stored and executable programs in a binary format (digital form). Thus, a distinction should be made between the physical (digital) form of "artificial intelligence" and its physical medium. While existing on a physical medium, AI can be introduced into any physical object (such as spacecraft, drones, robotic devices), man (cardiac pacemakers), living creature (chipped animals) etc. This distinction between the digital form and physical medium is of practical importance: property right to a thing (physical medium) should be different from exclusive right to an intellectual asset such as "artificial intelligence". Therefore, in purchasing a thing (AI physical medium) such as AI-enabled computers or gadgets under a sale agreement, the buyer does not become the owner of the right to AI since the transfer of title to a physical medium does not assign the exclusive right to "artificial intelligence".

Thus, any legal analysis of "artificial intelligence" should avoid the defects outlined above.

1. The Legal Nature of "Artificial Intelligence"

It should be noted above all that no analysis of "artificial intelligence" as a transferrable property is possible unless we recognize that "a transfer of property assumes different *economic relations* to transfer things and other property between persons. Their *formalization under civil law in the form of different agreements for inheritance and other forms of assignment* (transfer of civil rights and obligations) is covered by the *concept of civil law transaction* (italics added.– L.V.)" [Sukhanov E.A., 2019: 59–60]. Formalization of economic relations with AI under the civil law will obviously pose a number of critically important questions, one being the legal nature of "artificial intelligence" since the answer will determine the kind of agreements applicable to the creation and use of AI. Thus, the identification of the legal nature of "artificial intelligence" will give an idea of the regime applicable under civil law. Moreover, it should be borne in mind that "such regime is actually established not for items themselves but for those who deal with them in legally binding transactions. Meanwhile, different things at law will differ in this capacity by the applicable legal regime and not by their physical and economic properties while the details of such regime will be represented by this or another variety of property (civil) rights" [Sukhanov E.A, 2017: 45].

As applied to "artificial intelligence", this principle of civil law means that before actual relations with "artificial intelligence" are formalized, it is important to determine the property in question (its nature) and the rights it gives rise to since the latter (property rights to an item) will determine the

applicable legal regime. It means that an analysis of "artificial intelligence" by a civil law specialist should be focused on the item's legal parameters — civil law regime (technical and economic ones are none of his business!).

"Artificial intelligence" as the outcome of complex programming is still not qualified as a thing at law. Clearly, it is an intellectual product is not part of intellectual assets to be protected (Articles 128, 1225 of the Civil Code of the Russian Federation; hereinafter: Civil Code). The list of intellectual assets is known to be exhaustive: they are those that the legislator has afforded legal protection to. Obviously, AI has been put outside the regulatory scope of the civil law, something that appears to be mind-boggling and inexplicable in the context of digitization of all life spheres across the board. Therefore, the civil law study of the problem of "artificial intelligence" should be focused, among other things, at identifying its place in the system of things at law.

Let us turn to the definition of "artificial intelligence" provided in paragraph 1(2), Article 2 of the Federal Law No. 123-FZ: "artificial intelligence is a set of technological solutions allowing in the performance of specific tasks to imitate human cognitive functions (such as self-learning and search for solutions beyond preset algorithms) and obtain outcomes at least comparable with those achieved by human intellect. The set of technological solutions comprises an information and communication infrastructure (including information systems, information and telecommunication networks, other information processing technologies), software (including those incorporating machine learning methods), processes and services to process data and search for solutions"[3].

In the introduction it was already stated a negative attitude to the imitation of human cognitive functions by "artificial intelligence". It has a sense now to look at the first part of the AI definition, in particular, a set of technological solutions. From the perspective of civil law, this definition has missed the point since it deals with a technical rather than legal qualification of "artificial intelligence" while the problem of legal qualification remains unsolved.

It has a sense to describe briefly the author's approach to the identification of legal essence of "artificial intelligence". To decide on AI's protect-

---

[3] Federal Law No. 123-FZ "On the Experiment to Introduce Special Regulation for Creating Necessary Conditions for the Development and Introduction of AI Technologies in a Constituent Territory of the Russian Federation — Federal City of Moscow — and on Amending Articles 6 and 10 of the Federal Law "On Personal Data" of 24 April 2020 // Collected Laws of Russia, 2020. No. 17, Article 2701.

ability by law (Article 1225, Civil Code) we should identify the intellectual property regime extendable to AI, be it copyright law or patent law. As for the former, it should be borne in mind that if "artificial intelligence" is recognized as a copyright-protected item, it becomes protected from the time it is materialized (in digital form on a physical medium) and at the same time recognized as a copyrighted property. An item protected by patent law is regulated otherwise: once AI is recognized as such, its protection by law depends on a patent to be issued. In the first case, copyright to an item is based on the fact of its materialization in any objective form while patent right will arise from the registration of items under patent law (Articles 1259, 1353 of the Civil Code).

Based on the definition of the Federal Law No. 123-FZ, author of article present believes it is possible to conclude that AI is a complex thing at law that can comprise a variety of protected intellectual assets. Items of varying legal nature which make part of a complex object can be used only as a whole which is a distinctive feature of any complex thing including "artificial intelligence". Let us formulate the necessary constitutional features of a complex object in line with provisions of paragraph 1 (1), Article 1240 of the Civil Code:

combining a number of protected intellectual assets;
representing an indivisible whole from the legal perspective;
making part of an exhaustive list of complex things by law.

The last feature of a complex object as envisaged by the Civil Code creates a legal uncertainty for qualifying many items (such as websites) as complex things at law [Vitryansky V.V., 2018: 60–65].

In our opinion, "artificial intelligence" can comprise the following structural components:

computer software (Article 1261 of the Civil Code);
algorithms as knowhow (paragraph 1, Article 1465);
technological solutions as inventions (paragraph 1, Article 1350)
databases as complex (paragraph 1, Article 1240) or composite things (paragraph 2, Article 1260).

We deal with a thing at law that has a complex structure, with computer software undoubtedly at its core. Each of the listed elements is protected either by copyright (computer software) or patent law (inventions). As regards algorithms (knowhow), the legislator, as we know, has provided a special regime that differs from those applicable to other items of exclusive rights are part of the AI structure.

The complex structure of AI as an essentially ideal phenomenon determines its qualitative certainty as different from its constitutive intellectual assets which, in our view, allows to qualify "artificial intelligence" as a complex object under to Article 1240. This interpretation is obviously contrary to the legislator's provisions since paragraph 1 of that Article provides, as was noted above, for an exhaustive list of complex objects, only to suggest, in our view, that the problem of the legal nature of "artificial intelligence" and of the civil law regime applicable to it as an intellectual asset needs to be addressed as a matter of priority.

AI is thus a complex object different from its constitutive intellectual assets in terms of function and representing a qualitatively new entity beyond a mechanical combination of its elements.

However, even if "artificial intelligence" is recognized by law as a stand-alone complex object under Article 1240, the question of its legal protection will be still open since the legislator has not identified complex things among those subject to copyright and patent law. Clearly, "artificial intelligence" as a complex object can be transferred via the exclusive (property) right which, unlike the thing in question, is transferrable. It is the exclusive right to "artificial intelligence" that determines the peculiarities of the applicable legal regime.

It is crucial for a legal definition of "artificial intelligence" to make it clear how a comprehensive exclusive right to a complex object will arise given that the exclusive rights to its elements are owned by different holders. Being a costly and readily tradable intellectual product, AI is created by several parties including the organizer and at least several holders of intellectual assets (structural elements of a complex object). This gives rise to the problem of multiple ownership of "artificial intelligence" as a complex thing at law. V. A. Dozortsev, a well-known researcher of civil law, has proposed the following indication of multiple ownership of a complex object: a complex object is "the outcome of a multi-layered process where the creative work of individuals results in elements used at the second stage by other individuals to make a composite thing as a whole" [Dozortsev V.A., 2005: 144].

According to V.O. Kaliatin and E.A. Pavlova, Article 1240 creates a set of rights, as it were, to a "framework" object with the organizer as the main party. The authors note that the organizer does not creatively contribute to a new thing while limiting himself to making the necessary arrangements to create it [Kaliatin V.O., Pavlova E.A., 2014: 156]. This stance is contrary to a widespread approach of many IT experts whereby the AI organizer is

directly involved in its development acting not only as a team leader and coordinator but also as a party to the creative process by making an intellectual contribution to the common business of creating new algorithms and new software and integrating the already existing intellectual assets as elements of the AI's complex structure.

One would be hard pressed to accept a view whereby the organizer's activity is essentially organizational and technical, otherwise the question would be why he should have the exclusive rights not only to specific intellectual assets (parts of AI such as computer software, inventions, knowhow etc.) but also the exclusive right to "artificial intelligence" as a whole.

2. Problems of Disposing of the Exclusive Right to AI

Once "artificial intelligence" is qualified by law as a complex intellectual object, this will give rise to a number of questions on disposing of exclusive rights.

Under paragraph 1, Article 1240 of the Civil Code, the right to a complex object ("artificial intelligence") is the right to use protected intellectual assets that are part of "artificial intelligence". According to E.A. Pavlova, "a person (individual or legal entity) who has organized for a complex object to be created has a kind of special right which, albeit not always exclusive, allows to perform the necessary actions to use the complex object" [Pavlova E.A., 2016: 152–157]. While there is no explanation why the organizer's "kind of special right" is not always exclusive, the author's approach becomes clear if one follows E.A. Pavlova's logic and adopts the above stance that the organizer does not creatively contribute to the new thing but limits himself to making the necessary arrangements. Once the person who has arranged for a complex thing to be created limits himself to making the necessary arrangements without creative contribution to a new thing, an exclusive right to the thing is unlikely to emerge. But a number of questions still need to be explained in this case. Who will have an indivisible exclusive right? How an indivisible right to a complex object is possible if it was created through creative teamwork? Here are some explanations before giving answers.

As was stated previously, we believe that the organizer's creative input into "artificial intelligence" is hard to be disputed or denied. Thus, no exclusive right to AI can arise without the organizer's involvement. Moreover, the question is why a person making arrangements to create AI should own the rights both to individual intellectual assets (AI components such as

computer software, knowhow etc.) and the right to "artificial intelligence" as a whole. How does it come around?

It is not accidental that the legislator has introduced special rules for the disposal of exclusive rights to protected intellectual assets within a complex object. To reduce the risks associated with the transferability of exclusive rights, opportunities for abuse by the organizer and also a need to avoid the problem of splitting the comprehensive right to a complex object, the legislator has provided in Article 1240 for an exception from the general rules applicable to the exclusive right transfer and licensing agreements. The legislator has treated this problem in a special way: as we said before, since no complex intellectual object is listed among those subject to copyright and patent law, the issue of protecting complex objects and underlying rights is not fully settled. Obviously, the legislator had to amend the general rules of disposal of exclusive rights in this context (Article 1233).

To enable the use of intellectual assets within a complex object, the main rule provided in paragraph 3, Article 1233 was amended as follows: an agreement for disposal of exclusive right is deemed licensing agreement by default. Secondly, paragraph 1 (2) of Article 1240 provides that an agreement to purchase the right to use an intellectual asset (to be) created specifically as part of a complex object is deemed an exclusive right transfer agreement. Under this agreement, the right to use an intellectual asset (to be) created as part of a complex object will go to the person who has arranged for the creation of the complex object: the exclusive right to the corresponding intellectual asset is thus fully transferred to the organizer who acquires the right to use the asset in any way not contrary to the law and the right to dispose of the acquired exclusive right. It is admissible to believe that contractual regulation of the relations between the organizer and holders of intellectual assets within a complex object which is envisaged by law removes to a large extent not only the problem of abuse of the right but also the problem of splitting the comprehensive exclusive right to a complex object.

It is noteworthy that the provision of paragraph 1 (2) of Article 1240 is dispositive: if so agreed, a licensing agreement may be entered between the organizer and holders of intellectual assets (to be) created as part of a complex object (parties to the agreement). It means that where the concluded agreement is not explicit as to its licensing nature (or else where the licensing nature does not explicitly follow from the contractually established limits for the use of respective assets), the agreement shall be deemed the one for transfer of the exclusive right. This is the first point.

The second point is that licensing agreements which provide for the use of intellectual assets integrated into a complex object cannot be restricted by any term (within the effective term of exclusive rights), unless otherwise provided for by the agreement (paragraph 1 (3), Article 1240), and cannot contain any provisions restricting the license holder's right to use the object of the agreement; such provisions are deemed void (paragraph 2 of that Article). In fact, the said rule serves to provide the organizer with all possible powers to use a complex object. Thus, the proprietor contracted by the organizer under such licensing agreement cannot refuse to assign the right to certain ways of using the intellectual asset as part of the complex object as a whole. While the parties may envisage other terms, it is the already mentioned rule established by paragraph 1 (3), Article 1240, that will apply by default and not the provisions of paragraphs 3 and 4 (2), Article 1235.

It is worth noting that neither paragraph 1 (3) nor paragraph 2, Article 1240, contain any provision for mandatory issuance of exclusive licenses under licensing agreements. Likewise, the legislator does not prohibit any sub-licenses to be issued under the said agreements. These questions are obviously something to be agreed between the parties.

It is important to distinguish between exclusive rights to intellectual assets integrated into a complex object and comprehensive exclusive right to a complex object as a whole, that is, one should not identify the exclusive right (the right to use an intellectual asset integrated into a complex object) acquired by the organizer of the team effort with the comprehensive right he has in respect of the complex object as a whole (the right to a complex object).

So, one should identify the right to "artificial intelligence" as a whole since, according to V. A. Dozortsev, "a complex creative product cannot be used as a whole if we protect only those elements which make up the complex product". In the opinion of this renowned researcher, the point is about the fiction of authorship which serves to provide a person who has arranged to create a complex object with the right to use it as a whole. The civilian lawyer believes that "conventional constructs not reflecting the reality" but serving to achieve the final outcome (creating a complex product) are possible in this case [Dozortsev V.A., 2005: 153, 154].

As applied to the creation of "artificial intelligence", Dozortsev's statement means, as may understand, the following. The organizer needs to obtain from each holder the right to use the intellectual asset integrated into a complex whole on the legal basis discussed above (agreement for full transfer of the exclusive right, licensing agreement). Thus, the person who has

arranged to create "artificial intelligence" should enter into corresponding agreements with authors of computer software, proprietors of knowhow (creators of algorithms), authors of inventions and other intellectual asset holders. In Professor Dozortsev's logic, the organizer should purchase the right to use the complex product as a whole (comprehensive right to a complex object).

As lawyers know, the Civil Code provides no regulation of the relationships of purchase in respect of the comprehensive right to a complex object. In our view, such purchase may involve an agreement between the organizer and each asset holder — either in the form of a specific agreement or a specific contractual term envisaged in the agreement for full transfer of exclusive right or in the licensing agreement entered with each holder for the use of specific intellectual asset within the complex whole. On the above legal grounds, the AI organizer will purchase not only the rights to specific protected intellectual assets within AI but also the right to "artificial intelligence" as a whole.

With regard to the rights available to the organizer of complex object, there is an approach described in literature that one would be hard pressed to accept. In particular, O. Yu. Shilokhvost has proposed the following: "Apart from the said rights — *the exclusive right and the right of use* (italics added. — L.V.) — the right to a complex object will comprise other powers not related to the use of the relevant assets as exclusive things at law" [Shilokhvost O.Yu., 2011: 167–168]. Obviously, the author opposes the exclusive right to that of use. It would be useful to remind that pursuant to paragraph 1 (1), Article 1229 of the Civil Code the holder of an exclusive right (individual or legal entity) to an intellectual asset has a discretion to use it in any way not contrary to the law.

In fact, an exclusive right is both the right to use an intellectual asset and the right to dispose of it — that is, two powers are traditionally identified. According to M.A. Rozhkova, three powers that constitute an exclusive right should be identified: 1) the right of possession of an exclusive right; 2) the right of use of an exclusive property, and 3) the right of disposal of an exclusive right [Rozhkova M.A., 2014: 28]. According to V.A. Dozortsev, the right of use essentially means "the holder's opportunity of sole action to operate a thing for a material benefit" [Dozortsev V.A., 2005: 48].

Therefore, an exclusive right to "artificial intelligence" essentially means the holder's opportunity to change the legal history of the right itself either by way of transfer to a third party or by limitation — giving the right of use

to another person, putting the exclusive right to pledge, trust, collective management. It is noteworthy how the legislator treats the disposal of an exclusive right in paragraph 1, Article 1233 of the Civil Code that provides that the holder may dispose of an exclusive right to intellectual assets by any way not contrary to the law and the essence of such exclusive right including by way of contractual transfer to another person (under an exclusive right transfer agreement) or by giving another person the right to use the intellectual asset within contractually established limits (under licensing agreement).

Thus, an analysis of powers to use intellectual property should distinguish: power to use a thing — physical medium of the intellectual asset to be exercised by the holder not only through a legal procedure (donation, purchase agreements etc.) but also practical action with such physical media; power to use the right (exclusive right) to an intellectual asset exercisable through a mechanism for the disposal of exclusive right by giving someone the right to use the intellectual asset.

Purchasing of a physical medium by way of entering into an agreement for donation or purchase (exercising the power to use a thing) should equally assume, according to A. Makovsky, *the use of intangible assets.* This well-known civilian lawyer wrote: "The *use* can apply not only to things but also to intangibles such as reading a book, enjoying music, contemplating a picture, *searching for information in a database* (italics added. — L.V.), etc. The use means consumption, familiarization, perception of the essence and properties of an intangible product, something which is sought by the user in the first place" [Makovsky A.L., 2010: 617].

While in the first case the power to use the physical medium of intellectual asset is exercised through a legal regime established by the legislator for things, it is done otherwise in the second case — by disposing of an exclusive right (not of a thing) through licensing agreements. While in the first case we deal with a physical medium of intellectual asset — a thing at law, the second case presents a different situation: the thing at law is the exclusive right to an intellectual asset. Obviously, the powers of use and disposal as the constitutive elements of the exclusive right show complex relationships of interdependence.

Thus, the main contractual constructs to dispose of the exclusive right to AI are:

right transfer agreement;
licensing agreement;

sublicensing agreement;
exclusive right pledge agreement;
exclusive right pledge management agreement;
exclusive right trust management agreement;
exclusive right collective management agreement;
etc.

The aforementioned agreements are not exhaustive of the ways to dispose of the exclusive right to "artificial intelligence".

The exclusive right to AI can be acquired by other persons also by general succession (inheritance, reorganization).

Both individuals and legal entities can act as the organizer. Under paragraph 4, Article 1240 of the Civil Code, the person who has arranged for the creation of a complex object has a non-property right to specify his name (designation) and also to require the same from users. This right is non-transferrable and inalienable but can hardly be qualified as a personal non-property right since it accrues only to individuals. Meanwhile, legal entities can also be organizers of complex objects. If the organizer is an individual, he will creatively contribute to the creation of AI as a team member in spite of the performance of his organizational and technical functions. Therefore, it is wrong from a legal standpoint to assert a personal non-property right afforded to the organizer.

According to A.L. Makovsky, Article 1240 contains a set of provisions on a "quasi right" to authorship of legal entities as creators of complex products [Makovsky A.L., 2008: 280]. From V.S. Tolstoy's standpoint, it would be wrong to assert that the authorship of legal entities has an impact on the discussed structure of complex products. The author believes that the emergence of a new intellectual asset at law has forced the legislator to address the issue of normative regulation of a special "exclusive" right of those arranging to create a complex object [Tolstoy V.S., 2009: 116]. M.A. Rozhkova has a different view: she believes "the solution to the qualification problem of the "quasi right" to authorship of legal entities is simple: …authorship is an inalienable right rather than a personal non-property right, and, therefore, it would be wrong to regard the right to authorship of either individuals or legal entities in relation to a specific variety of intellectual rights" [Rozhkova M.A., 2018: 98].

The problem of qualification of the non-property right of the AI organizer is obviously still a matter of discussion. It is not resolved in the doctrine and requires further theoretical elaboration.

Conclusion

At the current stage, the doctrine and Russian intellectual property law need to be undoubtedly improved with regard to such innovative product as "artificial intelligence". This need is prompted not only and not so much by the faster pace of digitization of all life spheres across the board and the transformation of the IT environment, as by conceptual defects of many provisions found in Part IV of the Civil Code. A considerable number of questions relating to AI's essence, legal nature and regime normally deal with legal gaps. The paper presents the author's civilian view on the problem of "artificial intelligence" and possible solutions to overcome drawbacks in civil law regulation of relationships in the field.

"Artificial intelligence" is an innovative intellectual product to be regarded as a thing at law. The views proposed by different authors on legal personality of "artificial intelligence" should be discarded as wrong: only individuals can assume legal capacity and delict dispositive capacity by law. In case of AI, we deal with a convention that the author believes appropriate to put in quotation marks.

AI is a complex object under intellectual property law, a complex entity that can incorporate a variety of protected intellectual assets such as computer software, inventions, knowhow etc. A combination of various structural elements in a new product determines its unique qualitative certainty beyond a mechanical sum of the protected intellectual assets.

Qualifying AI as a complex object under intellectual property law allows to conclude that provisions of Article 1240 of the Civil Code on an exhaustive list of complex things need to be refined since the said provisions fail to ensure proper rulemaking and regulation of the relevant relationships in the context of digital change and emergence of innovative intellectual products.

As an essentially non-material (ideal) phenomenon, AI exists in society only in objective digital form on various physical media (computers and other devices). The existence of "artificial intelligence" on a physical medium allows to build AI into any physical piece of technology (robotic device, spacecraft, drone etc.), man (cardiac pacemaker), living creatures (chipped animals) etc.

A distinction between AI's digital form and physical medium is important in practical terms: there should be a difference between the property

right to a thing (physical medium) and the exclusive right to AI as an intellectual asset. AI as an intellectual asset can be transferred only via the exclusive right that, unlike AI itself, is transferrable.

A distinction should be made between the legal regime applicable to AI's physical media and the one applicable to the exclusive rights to AI. With regard to contractual regulation of relationships, we deal with a thing at law in the first case and with exclusive rights in the second case. Agreements will differ accordingly: transfers of physical media will involve purchase and donation agreements in the first place while the disposal of the exclusive right — an exclusive rights transfer agreement, licensing agreement, pledge agreement, trust management agreement, etc.

A distinction should be made between the exclusive right to an intellectual asset within a complex object and the comprehensive exclusive law to a complex whole ("artificial intelligence"). It is not accidental that the legislator has provided for special rules for the disposal of exclusive rights to protected intellectual assets within a complex object. To reduce the risks associated with the transferability of exclusive rights, opportunities for abuse by the organizer and a need to avoid the problem of splitting the comprehensive right to a complex object, the legislator has provided in Article 1240 for an exception from the general rules applicable to the exclusive right transfer and licensing agreements to be taken into account when creating AI.

References

1. Dozortsev V.A. (2005) Intellectual Property Rights. Concept. System. Codification Objectives: selected papers. Moscow: Statut, 416 p. (in Russ.)

2. Dremliuga R.I., Mamychev A.Yu., Dremliuga O.A., Matyuk Yu. S. (2019) Artificial intelligence as a subject of law: pros and cons. *Contemporary Dilemmas: Education, Politics and Values,* no. 1, p. 127.

3. Kaliatin V.O., Pavlova E.A. et al. (2014) Comments to Resolution No. 5 of the Russian Supreme Court Plenum, Russian Commercial Court Plenum No. 29 "On issues arising in connection with entry into force of Part 4 of the Civil Code". In: Research Comments of Legal Practice in the Area of Intellectual Property Rights Protection. Moscow: Norma, 480 p. (in Russ.)

4. Kuteinikov D.L. et al. (2019) Regulating human interaction with autonomous devices: a discussion of legal regimes. *Zakony Rossii*=Lex Russica, no. 9, pp. 85–95 (in Russ.)

5. Laptev V.A., Usenkov P.I. et al. (2020) Subject Matter of Business Agreements Today. In: Digital Economy. Conceptual Basis of Legal

Regulation of Business in Russia: a research. Moscow: Prospekt, 488 p. (in Russ.)

6. Makovsky A.L. (2008) Comments to Article 1240. In: Comments to Part 4 of the Civil Code of Russia (by chapters). Moscow: Statut, 715 p. (in Russ.)

7. Mefodieva K.M. (2019) Civil Law Regulation of Digital Data in Germany, U.S and Russia. Candidate of Juridical Sciences Summary. Moscow, 28 p. (in Russ.)

8. Pavlova E.A. (2016) Comments to Article 1240 of the Civil Code. In: Comments to the Civil Code of Russia (Part 4). L.A. Trakhtengerts (ed.). Moscow: NORMA, 485 p. (in Russ.)

9. Rozhkova M.A. (2014) Intellectual Property: Main Aspects of Protection. A manual. Moscow: Prospekt, 248 p. (in Russ.)

10. Rozhkova M.A. et al. (2018) Qualification of Intellectual Property Items. In: Civil Concept of Intellectual Property in the Russian Legal System: a collection of papers. Moscow: Statut, 272 p. (in Russ.)

11. Shilokhvost O.Yu. (2011) Comments to Article 1240. In: Article-by-Article Comments to the Civil Code (Part 4). P.V. Krasheninnikov (ed.). Moscow: Statut, 926 p. (in Russ.)

12. Somenkov S.A. (2019) Artificial Intelligence: A Thing or a Person? *Vestnik* Moskovskogo *Gosudarstvennogo Juridicheskogo Universiteta imeni Kutafina*=Moscow State Kutafin Law University Bulletin, no. 2, pp. 75–85 (in Russ.)

13. Sukhanov E.A. (2019) The Sources of Civil Law. In: The Civil Law: a manual. Moscow: Statut, 576 p. (in Russ.)

14. Sukhanov E.A. (2017) Law of Things: manual. Moscow: Statut, 560 p. (in Russ.)

15. Tolstoy V.S. (2009) *Information Civil Law*. Moscow: Academia, 293 p. (in Russ.)

16. Vitriansky V.V. (2018) *Reform of Civil Law in Russia: Interim Results.* Moscow: Statut, 528 p. (in Russ.)

Information about the author:

L.Yu. Vasilevskaya — Doctor of Sciences (Law), Professor.

Technologies of "Artificial Intelligence": Problems of Qualification and Legal Regime

⛉ Ekaterina Borisovna Poduzova

Kutafin Moscow State Law University, 9 Sadovaya-Kudrinskaya Str., Moscow, Russia 125993, ekaterinak7785@yandex.ru, https: // orcid.org/ 0000-0002-1945-4761

🗏 Abstract

Based on the civil law methodology, the paper analyzes the concept, legal nature and legal regime of the technology of artificial intelligence. The subjects discussed include: the civil law doctrine; provisions of the national and international law; enforcement practices. The study purports to theoretically justify the legal nature and legal regime of "artificial intelligence". In view of the above subject and purpose, the paper has relied on the comparative legal and structural system methods as well as simulation method. In particular, the comparative legal method has allowed to explore specific elements of "artificial intelligence" in the context of the national and international regulation of relevant relationships. The use of this method has also allowed to justify the legal regime applicable to technologies of "artificial intelligence" as intellectual assets. The use of the structural system method has allowed to substantiate the legal nature of AI and to identify its structural elements and the scope of application (information and health care services; development and use of robotic devices in the oil and gas sector; apartment buildings design etc.). The simulation method has provided an insight into the correlation between the concepts of "artificial intelligence" and "AI implementation form", and helped to propose a solution to the issue of legitimacy of sharing the exclusive right to AI. As a methodological peculiarity, the study combines the theoretical and empirical levels of cognition. The said combination of methods has allowed to look at the problems of qualification and legal regime in the context of controversial doctrinal and practical approaches. The study has found that the "artificial intelligence" technology, being a complex technical product, is essentially an intellectual asset (property). It was established that the doctrinal approaches qualifying AI technologies as a ("digital") person at law or a physical thing were unfounded. The paper argues in favor of recognizing the exclusive right to the AI technology as an intellectual asset.

Background

The "artificial intelligence" technologies are known to be widely used in a variety of fields [Vavilin E.V., 2021: 13-35]. Under the National Strategy for the Development of Artificial Intelligence[1] (sub-paragraphs a, b, c, l, paragraph 5), the AI technology as a specific digital product is designed to generate technological solutions including new intellectual assets. It is noteworthy that the Strategy makes no mention of other areas of AI application. For example, nothing is said about AI being used in technologically complex processes such as surgery (surgical medical interventions) in health care[2]. Using the area of application as a criterion, one could distinguish a number of AI technologies used for provision of various services (in particular, telemedicine technologies in health care); industrial production; national defense and security; apartment building utility system management; job matching services etc.

Of special interest is the specific use of AI technologies for the provision of health services. This area involves not only telemedicine solutions but also other AI technologies (such as those embodied in robotic devices or RD). Telemedicine consultations are currently in high demand at the federal health care centers. For example, a free telemedicine service is available at the Kulakov Research Center of Obstetrics, Gynecology and Perinatology as part of the Health Cloud project[3]. That consultations are also

---

[1] Presidential Decree No. 490 of 10 October 2019 "On the Development of Artificial Intelligence in the Russian Federation" (attached to the National Strategy for the Development of Artificial Intelligence for the period until 2030). Available at: URL: http://www. pravo.gov.ru 11.10.2019 (accessed: 10.11.2022)

[2] For surgery (surgical interventions), robotic AI devices are used. For example, da Vinci robotic surgeon (da Vinci Surgical System) comprises two units: a surgeon's console and a controller with four robotic arms. Available at: URL: ttps://ru.wikipedia.org/wiki/ Da_Vinci_(surgical robotic system) (accessed: 25.07.2023)

[3] See: Free telemedicine consultations. Available at: URL: https://ncagp.ru/index. php?_t8=200&pr_p_razdel=0&pr_p_object=383& (accessed: 25.07.2023)

provided by the staff of the Morozov Children's Municipal Hospital under Moscow's Health Department[4]. In fact, AI provides an example of combining different physical forms (AI external manifestations) with "artificial intelligence" as a technology. In using AI technologies in digital medicine, both patients and their proxies will pass health information to databases (data systems). Health care institutions will use AI technologies embodied in robotic devices for provision of relevant services.

While being important for the development of this country's technological potential, AI technologies are not on the list of protected intellectual assets. The Civil Code of Russia (hereinafter CCR) does not specifically regulate the development of "artificial intelligence" or the disposal of the relevant exclusive rights.

The civil doctrine and enforcement practice have developed superficial approaches to legal qualification and definition of the regime applicable to AI technologies. Thus, AI technologies have been wrongly identified with "artificial intelligence", with the latter assumed to be a "person at law", "digital personality", "digital person" or a thing "educable like a child".

Based on the above, it has a sense there is a special urgency to discuss the problems of legal qualification and legal regime of AI technologies.

1. "Artificial Intelligence" Technologies: Legal Qualification Discussion

Before identifying a legal regime applicable to AI technologies, it is necessary to determine their legal nature in the first place. As was already mentioned above, the civil doctrine has proposed a variety of approaches to qualify AI technologies. It is useful to discuss them in detail.

One of the doctrinal approaches recognizes AI as a person at law, with its advocates identifying "artificial intelligence" with AI technologies.

Thus, R. Dremliuga, O. Dremliuga, A. Mamychev and Yu. Matyuk endow AI with a legal personality but fail to identify any universally applicable objective criteria to recognize AI as a person at law [Dremliuga R.I., Mamychev A.Yu., Dremliuga O.A., Matyuk Yu.S., 2019: 127]. They just argue that AI and related technologies mimic human cognitive functions.

---

[4] Available at: URL: // https://xn--90adclrioar.xn--p1ai/telemedicine/ (accessed: 25.07.2023)

We believe this position to be contrary to the provisions of the CCR on persons. A natural person as a person at law is characterized by legal capacity and competence (Article 21 of CCR); a legal person has legal capacity (Article 49). A necessary qualifying attribute of a person at law is will (the presence of will and the ability to manifest it). Specific operating capabilities of AI are determined by its developers. These digital products do not possess a will of their own or an ability to manifest it. No amendment to the civil law can make up for the lack of will in AI technologies.

As an argument in favor of endowing "artificial intelligence" with a legal personality, S. Afanasiev notes its "cognitive features and skills" of a person at law (natural person) which AI will possess in the future [Afanasiev S.F., 2022: 13–31]. In contrast to the above statement, the author concludes that a combination of these features and skills "does not make AI personable". In our view, this stance is wrong. No mimicking of human properties and skills will make AI a person at law endowed with a will and volition.

In analyzing the definition of "will", it is possible to conclude that the presence of will is mainly manifested by the freedom of behavior[5]. The operational algorithm of AI technologies is determined by developer. These digital products are designed by their creators (natural persons or a team) as part of their job or under a statutory contract (commissioning contract or R&D contract for the performance of research and development). AI technologies embody the will of their authors or those who commissioned them. In our view, AI technologies are not free in their behavior, its variability being pre-determined at the time of development. The variability of AI behavior depends on the creator's or customer's will (persons at law).

A view advanced by some researchers [Kuteinikov D.L. et al., 2019: 85–95] whereby autonomous devices have "full freedom of will" and "become persons" with "a special legal status similar to that of a natural person" is also contrary to the generally acknowledged understanding of will and volition.

Persons at law exercise their civil rights "by their will and in their interest" (paragraph 2, Article 1 of CCR). AI technologies do not have either will or interest of their own and are thus unable to participate in legal relationships as persons at law. The term "similar to the natural person's legal status" proposed by these authors. It fails to clarify what combination of

---

[5] Available at: URL: http://slovardalja.net/word.php?wordid=3639 (accessed: 25.07.2023); URL: http://slovarozhegova.ru/word.php?wordid=3741 (accessed: 25.07.2023)

rights, duties and penalties applicable in the event of default is meant. The authors' opinion is contrary both to provisions of civil law and approaches developed by its enforcement practices.

A number of scholars recognize AI's legal personality on the grounds that it is embodied in a "digital (electronic) person". In support for AI's recognition as a "digital person" with the relevant status, M. Aksenova refers to "the use of a legal fiction similar to that of a legal person" [Aksenova M.A., 2020: 18–24]. One would be hard pressed to accept such argument. A legal fiction is only possible once there is a person possessing other properties and qualities than those shared by natural and legal persons. Meanwhile, we know nothing of such persons at the moment.

The author believes doctrinal concept of a "digital (electronic) person" or "digital (electronic) personality" does not have any constitutive features, being a product of a mechanical combination of the concepts "electronic" and "person" (or "personality"). The unjustified accentuation of these categories does not hold as an argument in favor of "digital" persons being parties to "digital relationships". The category of "digital person" as an "embodiment" of "artificial intelligence" fails to provide any conclusive proof that this digital entity has a legal personality.

No activity by natural or legal persons in a digital environment will give rise to new "digital" persons at law. The parties to digital interactions are natural and legal persons who use their "digital projections". As a reflection of natural or legal persons at law in a digital environment, the digital projection cannot make up an independent "digital" person at law. Any recognition of new "digital" persons at law ("electronic person" or "digital personality") endowed with a special "digital" status is, in our view, baseless.

A doctrinal link between AI technologies and "digital persons" is wrong for a number of reasons. Firstly, an absolute identification of a specific person at law with his or her "digital projection" involves certain practical difficulties [Puchkov V.O., 2020: 143–158]. Secondly, a failure to recognize someone's digital projection by law does not allow to establish a legal link between a person and his or her digital projection. In this connection, it is equally impossible to assert a legal link between a person's "digital projection" and AI technology.

E. Lungu and A. Kartskhia propose to enshrine "personal digital status" in the Civil Code as well as recognize "digital personality" as a special person at law [Lungu E.V., 2020: 61–63]; [Kartskhia A.A., 2017: 17–26]. We be-

lieve that this step, if implemented, will undermine the actual link between a person at law and his or her "digital projection" and will give rise to artificial "persons" at law devoid of delict dispositive capacity. It is noteworthy that E. Lungu and A. Kartskhia do not take into account a controversial approach developed by enforcement practices in this regard.

Thus, only an identifiable person at law will act in a digital environment[6]. Meanwhile, the interested party has to prove that the person and his or her digital image coincide[7]. In our view, this approach is fraught with a number of practical problems. It is applicable where wrongdoers come into unauthorized possession of someone's account or identity. Where a wrongdoer violates the third party rights through dishonest use of his "digital projection", the affected party will have no remedy. Operators and proprietors of information systems do not normally disclose information to identify and authenticate users and are not liable for an incorrect procedure of personal identification or actions committed under an anonymous, somebody else's or false "digital projection".

Here is another example of wrong qualification of AI as a person at law. A. Serova endows an information system with the attributes of a person at law [Serova A.V., 2019: 65–71], with a data system, robotic device and a chatbot (application for computers, smartphones etc.) put on par.

A. Serova's argument is not acceptable. An information system is a form of container for big data. The database and "artificial intelligence" are elements of big data. A data system cannot be confined to "artificial intelligence". Equating AI and data system is wrong from the legal standpoint. In our opinion, a distinction should be made between data systems (transferrable assets), their developers, proprietors and operators. The operator and proprietor (holder of the exclusive right) are liable for correct operation of the data system and observance of the rights of users and third parties.

While the operator (and/or proprietor) is a party to civil law relationships, the data system cannot be recognized as a person at law. The law on information and information technologies formally allows a natural person

---

[6] See, for example, Arbitration Court of Moscow District Resolution of 03.07.2018 on case No. A40-73666/2017 // SPS Consultant Plus; Arbitration Court of Sakha Republic (Yakutia). Decision of 27.12. 2018 on case No. A58-5241/2014 // Ibid.

[7] See, for example, 5th Arbitration Court of Appeal Resolution of 19.02. 2020 on case No. A51-330/2019 // SPS Consultant Plus: Judicial practice; 18th Arbitration Court of Appeal Resolution of 27.01. 2020 on case No. A76-40662/2018 // Ibid.

to act as data system operator[8]. Meanwhile, under the special law a data system operator — for example, an investment platform — can only be a legal entity (business company)[9].

Private and public information systems[10] are transferrable. Endowing these transferrable assets with the properties, "status" and "powers" of a person at law does not have any legal basis of its own and is a way for unfair proprietors to evade legal liability. In our view, there is no argument to support the qualification of AI technologies (data systems) as a special person at law.

According to D. Ponomareva and A. Barabashev, "artificial intelligence" may be of two types: "autonomous and subordinated AI", the former endowed with the attributes of a person at law since, as these authors believe, they can produce intellectual assets [Ponomareva D.V., Barabashev A.G., 2020: 36–43].

According to Article 1228 of the Civil Code, an intellectual asset is created by "an individual whose work has produced it". As was noted above, AI "creative" activities are enabled by the developer(s) of these technologies, first of all through the algorithms to "perform" specific actions. Any AI technology devoid of these algorithms will have no "ability" to create. Intellectual assets resulting from the "activities" of AI technologies will be authored by the creator(s) of these digital entities or anyone holding the relevant rights. If created by a legal person, they will be authored by the employee(s) of such legal person as part of their job.

Probably the issue of recognizing "artificial intelligence" as a legal person is best resolved by L. Yu. Vasilevskaya who wrote that insisting on a legal personality of "artificial intelligence" and endowing it (similar to natural and legal persons) with social and legal attributes of legal capacity, personality and delict dispositive capacity is a departure from the classical

---

[8] See paragraph 12, Article 2 of Federal Law No. 149-FZ "On Information, Information Technologies and Data Protection" of 27.07. 2006. Collected Laws of Russia (hereinafter CLR), 31.07.2006, No. 31 (part 1), art. 3448.

[9] See subparagraph 7, paragraph 1, Article 2 of Federal Law No. 259-FZ "On Raising Investments Through the Use of Investment Platforms and on Amending Specific Regulation of the Russian Federation" of 02.08. 2019. Available at: URL: http://pravo.gov.ru (accessed: 24.12.2022)

[10] On the status of a data system operator see, for example, Article 8, Federal Law No. 572 FZ "On the Identification or Authentication of Individuals Through the Use of Biometric Personal Data, Amending and Voiding Specific Regulations of the Russian Federation". Available at: URL: http://pravo.gov.ru (accessed: 29.12.2022)

principles of civil law since it is the legislator and not academics who endows persons with these legal attributes [Vasilevskaya L. Yu., 2021: 3–16]. For the same reason, no AI technology can be treated as a person at law.

Under another doctrinal approach AI is associated with a physical object (an item of property). Thus, S. Somenkov believes "artificial intelligence" to be an "AI-enabled thing". According to this author, "artificial intelligence is similar to a child — a product of upbringing and education" [Somenkov S.A., 2019: 75].

Somenkov's position appears to be wrong since defining the legal nature of "artificial intelligence" via its functional capabilities falls short of identifying the legal essence of AI technologies. A majority of modern devices are AI-enabled. Reducing their legal regime to that of things is to undo the legal role of inventors and programmers behind AI technologies. It is noteworthy that under paragraph 1, Article 1227 of CCR "intellectual property rights are independent from property and other material rights to physical media (things)". In this connection, it is wrong from the legal standpoint to identify AI (or AI technologies) with physical media.

According to E. Sukhanov, a renowned scholar, "different things at law will differ in this respect by legal regimes rather than physical or economic properties" [Sukhanov E.A., 2017: 45]. Recognizing AI technologies as "AI-enabled things" is tantamount to erasing the boundaries between the legal regime of things and that of intellectual property. This will leave the rights of the developer, organizer or any person possessing the exclusive right to AI technology without legal recognition and protection.

An analysis of the classical definition of "education" suggests that its main feature is "consistent influence on mental and physical development of children"[11]. In using AI technologies, persons at law provide specific information processed by the software that makes part of these digital systems. Meanwhile, the provision of information is not an equivalent of "consistent influence on mental and physical development" of AI technologies.

Paragraph 1 (2) and (3), Article 2 of Federal Law No. 123-FZ[12] of 24.04.2020 provides definitions of "artificial intelligence" and "AI technol-

---

[11] Available at: URL: // https://slovardalja.net/?ysclid=lda3sf1rep508916409 (accessed: 25.07.2023)

[12] Federal Law No. 123-FZ "On the Experiment to Introduce Special Regulation for Creating Necessary Conditions for the Development and Introduction of AI Technologies in a Constituent Territory of the Russian Federation — Federal City of Moscow — and

ogy". An analysis of these provisions reveals certain key attributes in the definition of each of these concepts. Both "artificial intelligence" and AI technologies are qualified as a structural systemic combination of complex technical objects (intellectual assets embodied in physical media (computers, smartphones, other devices etc.)). It is noteworthy that paragraph 3.5 of the Explanatory Memorandum to the EU draft artificial intelligence act explicitly emphasizes complexity as a feature of "artificial intelligence"[13].

"Activities" by AI technologies can result in new things at law including intellectual assets. For instance, AI technologies can perform medical treatment (such as surgical intervention) or develop a new product or invention prototype commissioned by a customer. The roadmap for the development of cross-cutting digital technologies "Neurotechnology and artificial intelligence" specifically notes a possibility of using artificial intelligence to design complex objects[14]. We believe that AI technologies are inherently capable of designing complex objects because they have the complexity required to solve complex technical issues. Meanwhile, no AI technology can be deemed the author of the intellectual assets it "creates".

The author is invariably a person at law possessing the exclusive right to the given AI technology or the right to use the intellectual asset under a licensing agreement. Such licensing agreement should presumably specify that the exclusive right to an intellectual asset created by the AI technology is attached to the license holder. Since only a natural person can own intellectual property (Article 1228, CCR), an intellectual asset will be authored by the proprietor (license holder) or by an employee thereof, where the former is a legal entity.

---

on Amending Articles 6 and 10 of the Federal Law "On Personal Data". Available at: URL: http://www.pravo.gov.ru (accessed: 01.09.2023)

[13] EU draft act on "artificial intelligence" (Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonized Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts). 21.04.2021 COM (2021) 206 final 2021/0106 (COD). Available at: URL: // https://artificialintelligence-act.eu/the-act/ (accessed: 01.09.2023). On 11.05. 2023 two committees of the European Parliament (on civil liberties and on single market) supported current wording of the draft. Available at: URL: https://www.forbes.ru/mneniya/489701-vzalis-za-um-kakie-pravila-primenenia-iskusstvennogo-intellekta-vvodit-es?ysclid=ljid6lsiy0427221219 (accessed: 25.07.2023)

[14] See Table 8, paragraph 3.11.1, roadmap for the development of cross-cutting digital technologies "Neurotechnology and artificial intelligence". Available at: URL: https://digital.gov.ru/ (accessed: 14.10.2019)

2. The Legal Regime of AI Technologies: Theory and Practice

AI technologies are intellectual assets which, as we know, are not subject to legal remedy (paragraph 1, Article 1225 of CCR). It is noteworthy that, compared to "artificial intelligence", AI technologies can comprise several types of artificial intelligence for different applications. They can also comprise previously created intellectual assets. AI technologies include an invention that enables them to mimic human cognitive functions. According to S. Sinitsyn, "items of patent law could potentially cover all technologies across the board". The author proves that "while computer software can be copyright protected, devices that use computer software or related inventions can be patent protected" [Sinitsyn S.A., 2022: 263, 311, 312].

Under paragraph 1, Article 27 of the TRIPS Agreement, "patents shall be available for any invention" in "all fields of technology"[15]. Meanwhile, paragraph 1, Article 52 of the European Patent Convention exhibits a questionable approach. While Article 52 reads that European patents shall be granted for any invention in all fields of technology[16], paragraph 1 (2) of this Article does not regard "programs for computers" as inventions. Computer software is not recognized as an invention subject to patent law by virtue of EU Directive 2009/24/EU as well (paragraph 1, Article 1)[17].

While sharing the position of researchers on the need in comprehensive protection of computer software, we believe it cannot be subject to patent law as an independent item. As part of AI technologies, computer software is technologically linked to other intellectual assets.

---

[15] Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS) 15.04.1994. In Russia document entered into force on 22.08.2012 (SZ RF. 10.09.2012, № 37); International Investment Instruments: A compendium. Vol. I. New York–Geneva, 1996. P. 337–371.

[16] European Patent Convention of 5.10.1973. Available at: URL: //https://www.patika. ru/Skachat_PDF/Evropeyskaya_patentnaya_konvencia.pdf?ysclid=ljjypasohe606433405 (accessed: 25.07.2023)

[17] See Directive 2009/24/EU of The European Parliament And of The Council of 23 April 2009 on the legal protection of computer programs (Codified version) (Text with EEA relevance). Available at: URL: //https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009L0024 (accessed: 25.07.2023). Article 1 Object of protection 1. In accordance with the provisions of this Directive, Member States shall protect computer programs, by copyright, as literary works within the meaning of the Berne Convention for the Protection of Literary and Artistic Works.

As was already noted, AI technologies are intellectual assets that under Article 1226 of the Civil Code "give rise to intellectual property rights including exclusive rights". According to Article 128 of CCR, things at law include property rights such as exclusive rights[18]. Qualifying exclusive rights as a thing at law will give rise to the question of legal regime.

Things (physical objects) traditionally involve proprietary rights (Articles 209, 216 of CCR) as well as classical powers of ownership, use and disposal. The question is whether the regime applicable to things should be extended to exclusive rights. Article 250 of CCR provides for a sale of shares in a jointly owned property. Under paragraphs 2, 4, Article 454 of CCR, the general purchase and sale provisions apply to securities and property rights unless otherwise followed from the content or nature of these rights. By virtue of paragraph 1, Article 454, the seller is the one who owns a property. Meanwhile, Article 28 of the Federal Law "On the Securities Market"[19] provides for a (property) right to paperless tradable securities. We believe that the legislator wrongly identifies the regime of paperless tradable securities with that of things because paragraph 1 (2), Article 142 of CCR associates paperless securities with "liabilities and other rights".

The Supreme Court of Russia recognizes the assignment of a claim under a sale or donation agreement as legitimate (paragraph 4 of Article 454 and paragraph 1 of Article 572, respectively)[20]. Meanwhile, according to V. Vitriansky, the provisions of Articles 382–390 of CCR "apply on a priority basis (with respect to the general provisions on purchase and sale of goods)" [Vitriansky V.V., 2005:17].

Thus, the ownership of a property right is recognized by law in a number of cases. However, the stance adopted by the legislator cannot be qualified as correct from the perspective of the principles of the law of pandects

---

[18] In this regard, see, for example, Constitutional Court of Russia Resolution No. 10-P of 24.03. 2023 "On constitutionality check of paragraph 4, Article 1370 of CCR and paragraph 3 of the Rules for remuneration for service inventions, service utility models, service commercial prototypes in connection with a complaint by Gidrobur-Service, a limited liability company. Available at: URL: pravo.gov.ru, 27.03.2023; 9th Arbitration Court of Appeal Resolution of 15.05. 2018 on case No. A40-124668/2017. Available at: https://kad.arbitr.ru/PdfDocument/ (accessed: 25.07.2023)

[19] Federal Law No. 39-FZ "On the Securities Market" of 22.04.1996. CLR, No. 17, 22.04.1996, art. 1918.

[20] See the Supreme Court Plenary Resolution No. 54 "On certain issues of application of Chapter 24 CCR regarding the change of liable persons based on a transaction" of 21.12.2017 // Supreme Court of Russian Federation Bulletin, No. 3, March 2018.

and understanding of the property right and property items generally acknowledged by academic science in Russia. The question is whether the legal regime of things and that of property could be extended to exclusive rights. Before addressing this question, we need to identify the constitutive features of the exclusive right to an intellectual asset.

It is noteworthy that, according to E. Sukhanov, an attempt to extend the regime of proprietary interests to those qualified as "intangibles" has naturally failed [Sukhanov E.A., 2017: 45]. This conclusion also holds for the identification of the constitutive features of exclusive rights.

We believe that qualifying the exclusive right as an "intangible thing" and recognizing the right of ownership (or other proprietary right) to it is wrong. As we know, the property and other proprietary rights under the Russian law are related to triple powers (ownership, use and disposal). Article 1233 of CCR provides only for the disposal of an exclusive right. To prevent the extension of the regime of things to exclusive rights, the legislator prohibits to apply the provisions of Part II of CCR to intellectual property rights (paragraph 3, Article 1227). As we know, independent contractual arrangements — such as exclusive rights transfer agreements, licensing agreements and franchising agreements, not identical to purchase-sale and rental agreements are envisaged to dispose of the exclusive rights (Articles 1233–1237 of CCR). According to I. Zenin, "the disposal of exclusive rights exhibits principally important differences. By granting a non-exclusive license, license holder (grantor) does not lose the possibility of further using the property" [Zenin I.A., 2023]. We believe that the regime of things and, therefore, property right cannot apply to the exclusive right due to its special legal nature not identical to that of things.

Recognizing inventions as part of AI technology makes it fairly relevant to discuss and address the problem of dividing the exclusive right and apportioning shares.

According to Yu. Kharitonova, the judicial practice "absolutely rejects that exclusive rights are shareable" [Kharitonova Yu. S., 2018: 65–72]. The Supreme Court of Russia determination discussed by this scholar prohibits any division of exclusive rights into shares "because the provisions on shared ownership (Chapter 16 of CCR) cannot apply to intellectual property rights in principle"[21]. Later on the Supreme Court provided another

---

[21] Supreme Court Decision No. 305-KG18-2488 of 03.12.2018. In Decision No. 287-PEK18 of 03.12.2018, the Supreme Court has dismissed a supervisory appeal to be considered at a session of the Supreme Court Presidium // SPS Consultant Plus.

justification of the prohibition to share exclusive rights. Thus, it was explained that while exclusive right to intellectual assets "can be held jointly by several persons", it "does not mean that the said persons are entitled to divide the exclusive right they own and apportion shares"[22].

In addressing the issue, one has to bear in mind that the exclusive right will arise in respect of IA[23] as a whole rather than its part. Meanwhile, O. Gutnikov and S. Sinitsyn conclude that "the current wording of paragraphs 2 and 3, Article 1229 CCR does not contain any explicit prohibition to divide the exclusive right owned by several holders into shares" [Gutnikov O.V., Sinitsyn S.A., 2019: 67–73].

It is noteworthy that the draft of the Federal Law "On Amending Part Four of the Civil Code of Russia"[24] envisaged the regulation of relationships between multiple parties on the intellectual property holder's side including those related to division of the exclusive right into transferrable shares. If interpreted literally, provisions of paragraphs 2 and 3, Article 1229 of CCR do not explicitly prohibit any division of the exclusive right into shares.

As was already mentioned, AI technologies are complex intellectual assets created with possible involvement of a considerable number of people. Since the extent of personal involvement may vary, the size of shares of the exclusive right to AI technology will depend on the complexity and amount of tasks performed by each team member. The size of ideal shares to be apportioned will be determined by an agreement between the developers.

We believe that the applicability of the terms of such agreement is beyond doubt. Meanwhile, their interpretation by courts in the event of a dis-

---

[22] Paragraph 35, Supreme Court Plenum Resolution No. 10 of 23.04. 2019 "On Application of Part Four of the Civil Code of Russia" // Supreme Court Bulletin No. 7, July 2019 ("SCP Resolution No. 10"). The Intellectual Property Court which previously admitted the divisibility of the jointly owned exclusive right into shares expressed in percent (see Intellectual Property Court Resolution No. S01-1002/2017 of 15.12. 2017 on case No. A40-210165/2016 , now refers to paragraph 35 of SCP Resolution No. 10 of 23.04.2019. See, for example, Intellectual Property Court Resolution No. S01-592/2018 of 28.10.2019 on case No. SIP-540/2017 // Intellectual Property Court Resolution No. S01-811/2021 of 27.04.2021 on case No. SIP-955/2020 // Consultant Plus.

[23] IA — intellectual asset.

[24] See the draft Federal Law "On Amending Part Four of the Civil Code of Russia" developed by the Center of Competences of the Skolkovo Foundation pursuant to paragraph 01.01.004.011.001 "Drafting a Federal Law for improving the relationships between exclusive right co-owners" of the action plan of the federal project "Statutory Regulation of the Digital Environment". Program Digital Economy of the Russian Federation. Available at: URL: https://sk.ru/legal/ (accessed: 25.07.2023)

pute may pose problems. Since the law does not provide for a possibility to divide an intellectual asset into shares, the terms of the agreement on sharing the exclusive right to AI technology will be interpreted by courts with reference to the provisions of paragraph 35, SCP Resolution No. 10 where the Supreme Court explains: "The above does not mean the said persons are entitled to divide the exclusive right they own and apportion shares".

It is believed the court is highly likely to void the agreement's section on the division of the exclusive right to inventions that make up the technology. Where the said agreement provides the basis for other agreements to be concluded (such as for disposal of the shares of the exclusive right), they will be likewise voided (Article 168 of CCR). To justify their decision, courts may refer only to the lack of legal provisions on the divisibility of the exclusive right.

In view of the principles enshrined in Articles 1 (2) and 421 (2) of CCR, author of the paper believes that the contractual terms to divide the exclusive right to the AI technology into shares are implementing the rights of natural and legal persons to establish any terms not contrary to the law. These terms are not subject to restrictions related to protection of the fundamental principles of constitutional system, morals, health, rights and legitimate interests of others, as well as national defense and security (subparagraph 2, Article 1 (2) of CCR. The stance to void such agreements follows only from the fact that lower courts are required to abide by the Supreme Court's interpretation of law.

In addressing the issue of divisibility of the exclusive right, we have to take into account the doctrinal understanding of exclusivity. Thus, V. Dozortsev, in characterizing the exclusive right to intellectual assets, believes it is exclusive not because of a single person's ownership but because it is attached by federal law exclusively to a particular person(s) on the bases established by law [Dozortsev V.A., 2008: 120]. Thus, the researcher allows for multiple party ownership of the exclusive right to an intellectual asset. According to N. Scherbak, it is legitimate "to apportion ideal shares of the exclusive right co-owned by several persons". The author believes that "a relevant agreement entered between the holders" will be qualified as a typical basis for apportioning ideal shares of the exclusive right [Scherbak N.V., 2021: 166–192].

Since a right is exclusive as long as the legislator recognizes it as vested in a particular person(s), this exclusivity will not be lost where the right is shared. In view of the ideal nature of shares, their apportioning will not re-

sult in the division of the intellectual asset as in the case of the mechanism envisaged by Article 133 (1) of CCR for property (things). Therefore, the division of the exclusive right to AI technology into shares is possible from the legal perspective.

Conclusion

The analysis of doctrinal and practical approaches to legal nature and legal regime of AI technologies suggests the following. The advocates of a doctrinal approach to recognize "artificial intelligence" as a person at law do not distinguish between "artificial intelligence" and AI technologies. Meanwhile, such "person" at law has no will. The analysis of the classical definition of "will" suggests it is mainly characterized by the freedom of behavior. AI technologies do not exhibit such freedom. The operational algorithm is built into AI technologies by the developer. These digital products are designed by their authors (natural persons or a team) as part of their job or under a statutory contract (commissioning contract or R&D contract). We believe that in "performing" certain actions, the AI technology follows the will of its developer/customer. In our view, AI technologies are not free to act since the variability of their behavior is determined at the time of development and depends on the will of their developers or customers (persons at law).

In view of author of the paper, the concept of "digital/electronic person" or "digital/electronic personality" proposed by the doctrine lacks constitutive features and results from a mechanical combination of the terms "electronic", "persons" and "personality". These concepts devoid of convincing justification do not support the recognition of special "digital" entities, parties to "digital relationships". The category of "digital person" as the "embodiment" of AI provides no convincing argument in favor of legal personality of AI technologies. In this connection, we believe that intellectual assets are authored by the individuals possessing either the exclusive right to AI technologies or the right to use the said technologies under a licensing agreement to create intellectual assets.

A person at law is the operator of a data system, not the system itself which is a transferrable property and cannot be recognized as a person at law. Formally, the law on information and information technologies allows natural persons to act as data system operators. It is a legal person (business company) that operates a data system such as a digital financial platform

under the law. In this regard, we believe that the doctrinal approach that identifies the AI technology with the data system does not hold.

The approach identifying AI technologies with things is not acceptable. Recognizing AI technologies as "AI-enabled things" is tantamount to erasing the boundaries between the legal regime of things and that of intellectual property. This will leave the rights of the developer, organizer or any person possessing the exclusive right to AI technology without legal recognition and protection.

 AI technologies to be a complex of technically sophisticated objects (intellectual assets) embodied in physical things (computers, smartphones or other devices). Complexity as a feature of "artificial intelligence" is explicitly states in paragraph 3.5 of the Explanatory Memorandum to the draft EU Artificial Intelligence Act.

Qualifying AI technologies as new intellectual assets is the only reasonable solution to the problem of their legal nature. AI technologies do not boil down to a sum of components that constitute them. AI technologies should be recognized as independent intellectual assets whose legal regime is not identical to that of their components.

AI technologies are not among intellectual assets subject to legal remedy (paragraph 1, Article 1225 of CCR). Meanwhile, they include a protectable invention. While sharing the position of researchers on comprehensive protection of computer software, we believe they cannot be patented as an independent object. Computer software is functionally related to other intellectual assets that make part of AI technologies. It has a sense to believe that computer software is patentable only in combination with other elements of AI technologies.

Under Article 1226 of CCR, AI technologies as intellectual assets "involve only those intellectual property rights that include the exclusive right". This raises the question of applicability of the exclusive right regime to intellectual assets such as AI technologies. Pursuant to Article 128, things at law include property rights, in particular, exclusive rights but qualifying the exclusive right as a thing at law raises the question of the applicable legal regime. We believe that qualifying the exclusive right as an "intangible thing" and recognizing the right of ownership to it is wrong. As we know, independent contractual arrangements — for example, exclusive rights transfer agreements, licensing agreements, franchising agreements not identical to purchase-sale and rental agreements — are envisaged to dispose of the exclusive rights.

AI technologies are complex intellectual assets characterized by multiple ownership since a considerable number of persons may be involved in the process of their development. The extent of personal involvement in the development of AI technologies may vary, with the size of apportioned shares of the exclusive right to be determined by an agreement between the developers.

A right is exclusive because the legislator recognizes that it is attached to a particular person(s). One should distinguish the exclusive right to an intellectual asset from a property right to a physical medium (thing). We thus believe that the exclusivity will not be lost where an exclusive right is shared. In view of the ideal nature of shares, their apportioning will not result in the division of the intellectual asset as in the case of the mechanism envisaged for division of property (Article 133 (1) of CCR). Provisions of paragraphs 2 and 3, Article 1229, if interpreted literally, do not explicitly prohibit any division of the exclusive right into shares or apportioning of ideal shares of the exclusive right to intellectual assets. Therefore, the division of the exclusive right to AI technology into shares is possible from the legal perspective. Since the Supreme Court of Russia in its explanations prohibits to divide the exclusive right into shares and apportion these shares, we believe that paragraphs 2 and 3 of Article 1229 of CCR need to be amended through adoption of the provisions allowing to divide the exclusive right into shares and to apportion these shares.

References

1. Afanasiev S.F. (2022) Material and procedural legal personality of artificial intelligence. *Vestnik grazhdanskogo protsessa*=Bulletin of Civil Process, no. 3, pp. 12–31 (in Russ.)

2. Aksenova M.A. (2020) The concept of "legal person" in the legal environment. *Yurist*=Lawyer, no. 7, pp. 18–24 (in Russ.)

3. Braginsky M.I., Vitriansky V.V. (2005) The contractual law. A textbook. Vol. 2: Property Transfer Agreements. Moscow: Statut, 840 p. (in Russ.)

4. Dozortsev V.A. (2005) Intellectual property rights. Concept. System. Codification objectives. Collected articles. Moscow: Statut, 416 p. (in Russ.)

5. Dremliuga R.I., Mamychev A. Yu., Dremliuga O.A., Matyuk Yu. S. (2019) Artificial intelligence as a subject of law: pros and cons. *Contemporary Dilemmas*: *Education*, *Politics and Values*, no. 1, p. 12.

6. Gutnikov O.V., Sinitsyn S.A. (2019) Multiplicity of ownership of the exclusive right: aspects of theory and practice. *Zakony Rossii: opyt, analiz, praktika*=Russian Laws: Experience, Analysis, Practice, no. 12, pp. 67–73 (in Russ.)

7. Kartskhia A.A. (2017) Digital imperative: new technologies create a new reality. *Avtorskoye pravo i smezhnye prava*=Copyright and Adjacent Rights, no. 8, pp. 17–26 (in Russ.)

8. Kharitonova Yu. S. (2018) Joint exercise of exclusive rights. *Pravo i ekonomika*=Law and Economy, no. 11, pp. 65–72.

9. Kuteinikov D.L. et al. (2019) Regulating human interaction with autonomous devices: legal regime discussion. *Zakony Rossii*=Lex Russica, no. 9, pp. 85– 95 (in Russ.)

10. Lungu E.V. (2020) Constitutional relationships in light of transition to digital state: regulatory trends. *Rossiyskaya justitsiya*=Russian Justice, no. 2, pp. 61–63 (in Russ.)

11. Ponomarev D.V., Barabashev A.G. (2020) Patent protection of research outcomes and artificial intelligence: issues and challenges. *Pravo i tsifrovaya ekonomika*=Law and Digital Economy, no. 3, pp. 36–43 (in Russ.)

12. Puchkov V.O. (2020) Main aspects of the digital image of person at law in the civil doctrine and practice. *Arbitrazhniye spory*=Arbitarion Disputes, no. 3, pp. 143–158 (in Russ.)

13. Scherbak N.V. (2021) An ideal share of the exclusive copyright: utopia or reality? *Vestnik grazhdanskogo prava*=Bulletin of Civil Law, no. 5, pp. 166–192 (in Russ.)

14. Serova A.V.(2019) Introducing digital technologies into relationships between persons working at a given employer. *Rossiyskiy yuridicheskiy journal*=Russian Law Journal, no. 5, pp. 65–71 (in Russ.)

15. Sinitsyn S.A. (2022) *Comparative patent law: current issues.* Moscow: Infotropik, 416 p. (in Russ.)

16. Somenkov S.A. (2019) Artificial intelligence: a thing or a person? *Vestnik Moskovskogo Juridicheskogo Universiteta imeni O.E. Kutafina*= Bulletin of Kutafin University, no. 2, pp.75–85 (in Russ.)

17. Sukhanov E.A. (2017) Law of things: manual. Moscow: Statut, 560 p. (in Russ.)

18. Vavilin E.V. (2021) How artificial intelligence is transforming civil law and procedural relationships: shaping new legal regimes. *Vestnik grazhdanskogo protsessa*=Bulletin of Civil Process, no. 6, pp. 13–35 (in Russ.)

19. Vasilevskaya L. Yu. (2021) "Artificial intelligence": problems of civil regulation. *Ekonomika i pravo*=Economy and Law, no. 10, pp. 3–16 (in Russ.)

20. Zenin I.A. (2023) Intellectual Property Law. Available at: Urait e-library. https://urait.ru/library?utm_=&ysclid=ljcuajhygi184594193 (accessed: 01.08.2023) (in Russ.)

Information about the author:

E.B. Poduzova — Candidate of Sciences (Law), Associate Professor.

Artificial Intelligence vs. Judicial Discretion: Prospects and Risks of Judicial Practice Automation

Valentina Aleksandrovna Rodikova

Institute of Management, Russian Presidential Academy of National Economy and Public Administration, 66 Eighth March Street, Yekaterinburg 620144, Russia, rodikovajus@gmail.com, SCIENCE INDEX: 4890-5030

Abstract

The paper fits into a wide public discussion of the introduction of artificial intelligence into the national judicial system, with the underlying risks and legal vulnerabilities analyzed on specific examples of enforcement practices. The effective national legislation on the use of artificial intelligence and the latest international laws on the subject including the EC's AI Act Compromise Text were studied. The paper looks into the risk factors characteristic of judicial discretion and judicial AI both on a comparative and standalone basis. Controversial opinions by law enforcement agencies, national and international researchers, AI developers on the prospects of using AI in the justice system were explored. The paper provides conclusive arguments that the introduction of AI is not feasible in the short and medium term in view of the current risks and geopolitical environment, state of the legal framework and judicial principles effective in the Russian Federation.

Keywords

artificial intelligence; neuronet; judicial discretion; strong and weak artificial intelligence; risk-based approach; justice system; forecasting; biometric identification.

<div align="right">

"If you can run faster than that you'll be fine"[1].

E. Musk

</div>

Background

In launching the digital transformation of Russia in December 2020, President Vladimir Putin pointed out to a need "to ensure broad introduction of AI technologies and big data analysis" including "experimental legal regimes to use AI in specific economic sectors and social services"[2].

However, self-learning neural networks, technologically in vogue worldwide as the core of generative AI technology, have failed to be enshrined in the national legislation, only to become the apple of discord among law enforcement agencies. The disputes concerning the risks of neural networks used in judicial practice have been especially violent, in particular, because of a lack of the doctrinal definition of artificial intelligence and provisions to regulate and address the likely negative scenarios. Likewise, no final risks of convergence of AI and data science (big data processing) in the process of deep learning of neural networks were defined.

When a majority of researchers date the introduction of the term "artificial intelligence" back to 1956 (G. McCarthy's presentation) [Smith C., 2006], they forget about three robotic technology laws proposed by Isaac Asimov in 1942 that essentially constrain in relative terms the emerging neural network-based products. Despite more than well-established history of the phenomenon under discussion, its conventional definition is not there yet, largely due to the fact that the term is too common [Kok J. et al., 2009: 2] and that there are legitimate doubts whether AI is a hoax launched in the interest of international corporations such as Intel (along the lines of "greenhouse effect" embedded into the public mind in the interest of Dupont Corporation), still more so since the so-called neural network itself, unlike the proposals to introduce it into social and economic sectors at large, did not evolve much from ABBYY Finereader, a text recognition software marketed in 1993.

---

[1] You can run away from it: Elon Musk jokes about his «friendly» robot // URL.: https://www.thesouthafrican.com/lifestyle/elon-musk-tesla-bot-humanoid-form-ai-robot-watch/ (last accessed on 22.05.2023).

[2] Putin has announced a need in digital transformation of Russia // Available at: URL: https://tass.ru/ekonomika/10172635 (accessed: 22.05.2023)

While defining AI as a field of science, P. Morkhat believes that "the main problem why an exact and common definition is not yet developed" is "a lack of clarity what is exactly intelligence as such" [Morkhat P.M., 2017: 26]. The author proposes to describe AI via its "key features: learning/self-learning ability, ability to understand and reflect, self-control" [Morkhat P.M., 2017: 31].

S. Russel and P. Norvig identify four main approaches to define "artificial intelligence": those based on human thinking and behavior and on rational thinking and behavior [Russell S., Norvig P., 2010: 1–2]. Some international researchers believe AI (what appears to be an optimal point of view) to be a platform for a number of promising technologies used in automated logic and big data processing [Haskins A., Arora S., Nilawar U., 2017: 4] or a branch of science [Rissland E.L., 1990: 1958–1959].

1. Enshrining AI Regulation in the National Law

Some authors [Laptev V.A., 2021] wrongly assume that the term "artificial intelligence" first appeared in the Action Plan (Roadmap) of the National Technological Initiative Autonet[3] in 2018.

Thus, the term was mentioned in paragraph 20, Presidential Decree No. 642 of 01 December 2016 "On the R&D Strategy of the Russian Federation" and in paragraph 12, Presidential Decree of 09 May 2017 "On the Information Society Development Strategy in Russia for 2017–2030". Further on, AI was repeatedly in the legislative focus: under paragraph 9, Presidential Decree No. 490 of 10 October 2019 "On the Development of Artificial Intelligence in Russia"[4], the use of neural networks was actually restricted to the so-called "weak AI" capable of solving only narrow tasks. AI was described in the document as "a set of technological solutions allowing to mimic human cognitive functions (such as self-learning and search for solutions outside a preset algorithm) and address specific tasks with results at least comparable with those of human intellect". This definition was later reproduced in Article 2 of Federal Law No. 123-FZ "On the Experiment to Introduce Special

---

[3] Action Plan (Roadmap) of the National Technological Initiative Autonet. Annex No.2 to Minutes No. 1 of 24 April 2018. Moscow, Presidium of the Council for Economic Upgrading and Innovative Development of Russia under the President of Russia, p. 21.

[4] Presidential Decree No. 490 of 10 October 2019 "On the Development of Artificial Intelligence in Russia" // Available at: URL.: http://www.kremlin.ru/acts/bank/44731 (accessed: 20.05.2023)

Regulation for Creating Necessary Conditions for the Development and Introduction of AI Technologies in a Constituent Territory of the Russian Federation — Federal City of Moscow — and on Amending Articles 6 and 10 of the Federal Law "On Personal Data" dated 24 April 2020.

The experimental regime introduced by this law has not been extended to other constituent territories, in particular, because of numerous problems regarding the implementation of the embedded substantive imperatives (copyright to neural network's outcomes, personal data processing, security, confidentiality etc.).

The legislation has not defined to what extent AI could be used to process specific categories of anonymized personal data (such as medical data) for more efficient public and municipal governance — data which, according to A. Saveliev, have "a special legal status due to potentiality of highly negative implications for the person if the processing terms were violated"[5]. There is no legal basis for AI to assume liability for the harm to human life and health as well as no understanding whether neural networks have a legal personality.

At the same time, experts in the military use of AI note that the three reasons for choosing the incoming data as the principal target are, by the order of priority, "complete dependence of insights on the amount and quality of inputs; difficulty to establish the fact of data diddling or editing; opportunity to gain a major advantage over a party in a dispute/conflict if decisions were made on the basis of analysis of misleading information" [Galkin D.V., Stepanov A.V., 2021: 73]. This is also true where AI is embedded into the system of justice.

Domestic technical regulations present artificial intelligence as a simulatable (artificially mimicable) intellectual activity of human mind (paragraph 3.17, GOST R 43.0.5-2009 "Information support of technologies and operator activities. Data exchange processes in technologies. General provisions)[6].

Starting from Presidential Decree No. 490 of 10 October 2019 "On the Development of Artificial Intelligence in Russia", AI has been qualified as

---

[5] Artificial intelligence and law: a link between the two? // Available at: URL.: https://www.garant.ru/news/1401154/ (accessed: 20.05.2023)

[6] National standard of the Russian Federation. Information support of technologies and operator activities. Data exchange processes in technologies. General provisions // Available at: URL.: https://docs.cntd.ru/document/1200079262 (accessed: 17.05.2023).

either "strong" or "weak" in the wake of the Western approach contained in the latest IBM research[7] of 2023. Thus, the Russian legislator and international developers believe "weak" ("narrow") AI to be the one focused on specific practical problems (Apple's Siri, Amazon's Alexa, IBM's Watson, autonomous vehicles, systems for voice recognition, virtual agents, computer vision, advisory mechanisms etc.). In contrast, "strong" AI is a combination of artificial general intelligence (AGI) and artificial super intelligence (ASI), the latter being a theoretical form that provides a device with an intellect superior to that of man (self-consciousness capable of solving problems, learning and planning for the future). For the national legislator it is associated with high risks since end results are not predictable and decision-making algorithms unclear.

There is currently no single document in Russia to regulate the development, implementation and use of AI, and to define the acceptable level of risks, legal personality of the parties involved, etc.

2. Latest European Law on Artificial Intelligence

Since AI systems, along with huge potential to boost economic growth, innovative development and global competitiveness, obviously carry major risks for security and protection of the core human rights and liberties, the European Commission published back in February 2020 the so-called "White Book" on artificial intelligence with a proposal to set up the European framework on AI and the limits of its use.

In October 2020 the European Parliament adopted 3 AI-related legislative resolutions on ethics, civil liability and intellectual property; in April 2021, the European Commission made proposals on the so-called AI Act which contained a technologically neutral definition of AI systems and also four risk categories for AI applications: unacceptable (contrary to EU values), high-risk (negatively affecting the security and core values of individuals), limited risk (those that meet specific transparency obligations) and minimal risk (those without obligations except those of the effective law).

On 6 December 2022 the European Council approved the general approach to the AI Act[8] explaining the requirements to high-risk AI systems

---

[7] What is Artificial Intelligence // Available at: URL.: https://www.ibm.com/topics/artificial-intelligence (last accessed on 23.05.2023).

[8] Council, Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights // Available at: URL.: https://www.consilium.europa.eu/en/press/press-

which identifies the general purpose AI systems, defines the regulatory scope (for example, national security, defense and related areas including law enforcement are ruled out) and proposes to create "regulatory sandboxes" to support AI-enabled innovations and open-code AI components[9]. As a result of discussions in the European Parliament, the world's first Transparency and Risk Management Rules for AI[10] were approved on 11 May 2023.

The drafters follow a risk-oriented approach to establish obligations for both AI suppliers and users depending on the aforementioned risk levels generated by artificial intelligence. However, before engaging in negotiations with the European Council to finalize the AI Act, the European Parliament will have to approve the draft "negotiating mandate" at its session scheduled for 12–15 June 2023.

The Rules completely prohibit the following AI practices:

real-time remote biometric identification systems in publicly accessible spaces;

remote biometric categorization using sensitive characteristics (such as sex, race, ethnicity, citizenship, religion, political orientation);

enforcement forecasting systems (based on profiling, location or past criminal behavior);

emotion detection systems at enforcement and judicial bodies, workplaces and education institutions;

indiscriminate deletion of biometric data from social networks, using video footage generated by surveillance cameras to create face recognition databases (in violation of human rights such as the right to privacy).

The Rules provide for obligations (individual legal regimes) shouldered by suppliers of basic models such as GPT, and extra "transparency" requirements, in particular, disclosure of the fact that the content was generated by AI. It is noted that the amendments are designed to establish human control over AI, with neural networks to be "safe, transparent, traceable, non-discriminatory, environmentally friendly"[11]. High-risk categories were expanded to include harm to people's health, safety, core rights and envi-

---

releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/ (accessed: 23.05.2023)

[9] Ibid.

[10] Al Act: a step closer to the first rules on Artifical Intelligence // Available at: URL.: https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence (accessed: 23.05.2023)

[11] Ibid.

ronment, as well as impact on electors during a political campaign and on trading platforms of social networks with more than 45 million users. It is assumed that both natural and legal persons have the right to make complaints about AI systems and receive explanations of decisions they generate[12].

Of special interest for the issue being discussed is paragraph 38, Chapter 1, Title III of the AI Act Compromise Text of 16 May 2023 which provides key risk scenarios — equally applicable to the Russian regulatory system — of AI use in law enforcement and judiciary activities: "Actions by law enforcement authorities involving certain uses of AI systems are characterized by a significant degree of power imbalance and may lead to surveillance, arrest or deprivation of a natural person's liberty as well as other adverse impacts on fundamental rights guaranteed in the Charter of 07 December 2000. In particular, if the AI system is not trained with high quality data, does not meet adequate requirements in terms of its performance, its accuracy or robustness, or is not properly designed and tested before being put on the market or otherwise put into service, it may single out people in a discriminatory or otherwise incorrect or unjust manner.

Furthermore, the exercise of important procedural fundamental rights, such as the right to an effective remedy and to a fair trial as well as the right of defence and the presumption of innocence, could be hampered, in particular, where such AI systems are not sufficiently transparent, explainable and documented. It is therefore appropriate to classify as high-risk a number of AI systems intended to be used in the law enforcement context where accuracy, reliability and transparency is particularly important to avoid adverse impacts, retain public trust and ensure accountability and effective redress.

In view of the nature of the activities in question and the risks relating thereto, those high-risk AI systems should include in particular AI systems intended to be used by or on behalf of law enforcement authorities or by Union agencies, offices or bodies in support of law enforcement authorities, as polygraphs and similar tools insofar as their use is permitted under relevant Union and national law, for the evaluation of the reliability of evidence in criminal proceedings, for profiling in the course of detection, investigation or prosecution of criminal offences, as well as for crime analytics regarding natural persons. AI systems specifically intended to be used for administrative proceedings by tax and customs authorities should not

---

[12] Al Act: a step closer to the first rules on Artifical Intelligence // Available at: URL.: https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence (accessed: 23.05.2023)

be classified as high-risk AI systems used by law enforcement authorities for the purposes of prevention, detection, investigation and prosecution of criminal offences.

The use of AI tools by law enforcement and judicial authorities should not become a factor of inequality, social fracture or exclusion. The impact of the use of AI tools on the defence rights of suspects should not be ignored, notably the difficulty in obtaining meaningful information on their functioning and the consequent difficulty in challenging their results in court, in particular by individuals under investigation". [13]

In light of the above reasonably restrictive approach to the use of AI systems in the judicial system, it appears expedient to discuss the risks resulting from AI integration into the national system of justice compared to the established factive presupposition (axiomatic modality which supports the sense and presumption of a statement and, while not being part of the presumption, makes sure that it is true) [Strawson P. , 1952: 113] of judicial discretion.

3. Judicial Discretion: Risks, Limits, Algorithmization

According to some researchers, the existence of judicial discretion as an institution is explained by the existence of objective regulatory peculiarities of some relationships and legislative gaps, where the weight of subjective criteria is minimal or absent [Tretyakova T.N., Karamanukian D.T., 2020: 6].

A number of authors believe judicial discretion to be "a specific type of law enforcement activities based on reasoning as a way to find the best solution in a given situation" [Makarikhina O.A., 2014: 15], something that actually identifies this institution with AI which could be hypothetically used in legal proceedings to algorithmize the process of searching for an optimal solution in a certain context.

I.A. Pokrovsky understood judicial discretion as "the right to interpret the law more freely, complement and even rectify it as may be required by the sense of justice and fairness" [Pokrovsky I.A., 1998: 90]. On the contrary, other researchers perceive the judge's personal conviction as "an outright opportunity for arbitrary judgment" [Morkhat P.M., 2018: 9].

---

[13] DRAFT Compromise Amendments on the Draft Report Proposal for a regulation of the European Parliament and of the Council on harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts // Available at: URL.: https://www.europarl.europa.eu/resources/library/media/20230516RES90302/2023 0516RES90302.pdf (accessed: 23.05.2023)

Judicial discretion is apparently an exercise of court powers to solve the case on the legitimate, fair and well-founded basis while using an opportunity to impose sanctions/render a judgment under a number of legitimate options and limits in the context of conflicts of law and legislative gaps.

Judicial discretion actually differs from AI as much as the judge's personal conviction from the standard algorithm of rendering a judgment (since classical AI operates outside a preset algorithm, its decisions are not predictable).

At the same time, the national law does not enshrine a number of key factive presuppositions (presumed judgments) such as definitions of the key concepts that make up constituent elements of a crime, only to multiply a possible risk of unjust, unfounded and illegal decisions in the context of judicial discretion understood as certain freedom of opinion.

Thus, a vast majority of "reputational" disputes involving legal persons as a claimant will be resolved by court on the basis of subjective principles of judicial discretion.

There is no factive presupposition in the national civil law to allow for an objective and impartial assessment of circumstances in cases concerning business reputation of legal persons. Meanwhile, the decision-making powers of a judge are based on a syllogism where one of the components, apart from a legal provision (if any), is the actual circumstances of the case [Barak, A. 1999] whose unambiguous, implicit parameters and matching criteria enable a fair judgment.

One example is where a court has (or does not have) reasonable grounds supported by law to qualify the defendant's statement as an asserted fact/ personal opinion, and reasonable grounds to qualify as irrelevant the defendant's statements addressed to the claimant. Different legal regimes applied by courts in considering business defamation cases where they ratify the defendant's subjective opinion or statement of fact, as well as the criteria to refer them to a given presupposition, are not enshrined in law. In the Defamation Review of Practice of 16 March 2016, the delineation of these regimes given a lack of clear reference criteria was claimed to be "the hardest decisions for courts"[14], only to "cause judicial errors"[15].

---

[14] Defamation Review of Practice (approved by the Presidium of the Supreme Court of Russia on 16 March 2016) // Available at: URL.: https://www.vsrf.ru/documents/thematics/15165/ (laccessed: 04.11.2021)

[15] Ibid.

Meanwhile, the initiatives to build neural networks into the national judicial system such as proposed by V.A. Laptev [Laptev V.A., 2021] appear to be fraught with even greater risks than judicial activism and multiplicity of imperfect opinions and personal convictions, and regulatory gaps.

4. Prospects and Risks of Judicial Practice Automation in Russia

Under the most likely scenario of the phased introduction of AI systems into judicial practice envisaged, in particular, by V. Laptev, AI will be implemented consecutively as an assistant judge as part of legal proceedings and in considering cases on their merits (short-term prospects); for evaluation of evidence and expert assessments (medium-term prospects); and as a possible replacement of judges to perform specific functions (long-term prospects) [Laptev V.A., 2021].

Meanwhile, recognizing and translating audio minutes of court sessions and documents provided by the parties into a machine-readable format will predictably involve putting sensitive information (including personal data) within reach of an unlimited range of people, both developers of an interface and other individuals, at the risk of unauthorized access and theft of "big data".

A "restricted" approach to storing and processing personal data — including those generated and processed in the judicial system — seems to be more effective, including in the context of current geopolitical threats, since it rules out sporadic and other mechanic "failures" related to the use of controversial technologies such as AI.

One example is China where relevant resolutions were adopted for National Data Administration[16] for control of the privacy and security of data of this newly created agency established, among other things, to secure full state control over all sensitive data of both individuals and economic institutions which are not designed to be processed, transmitted or used outside the government system. The Judicial AI (AI for legal assistance) actually handle in China more than 200 thousand cases a month [Stepanov, O.A., 2022: 229-237], with the neural network integrated into cloud-based national Big Data systems controlled by a number of public agencies. Mean-

---

[16] Available at: URL.: https://www.technologyreview.com/2023/03/15/1069814/china-new-bureau-data-economiy/ (accessed: 23.05.2023)

while, on 11 April 2023 the Cyberspace Administration of China published for public discussion the draft of Administrative Policies for Managing Generative AI-Enabled Services[17] which was innovative in imposing the responsibility for the content created by generative AI (chat bots etc.) on "suppliers" — developers and/or distributors of software solutions, with operators assuming an additional obligation to protect personal data where their processing is envisaged by the product.

However, the concept of AI learning on open judicial data (AI-Ready Open Juridical Data) does not stand up to criticism as there is neither a national strategy to publish such data for machine and deep learning nor quality (maturity) criteria of such data.

The government's involvement as a customer, regulator and contributor to AI implementation including in the judicial system makes it principally impossible to use an open code in this process (decentralized model of software development and management), something also contrary to the requirements of FSS Order No. 97 of 16 March 2022[18]and FSS Order No. 171 of 01 May 2021[19]. Not surprisingly, Russia is not in the index of responsible AI users created by Canada's IDRC CRDI including in the judicial system because participation requires a large amount of strategically important data to be provided.

Using AI for legal assessment of evidence including to analyze handwriting and forgery is unlikely because AI has so far failed to pass even CAPTCHA test (Completely Automated Public Turing Test to Tell Computers and Humans Apart), to say nothing about its ability to reliably determine the ownership of texts, their context, language nuances or abstract concepts.

Making judicial Big Data available to judicial AI, just like a need to digitize the judicial system across the board advocated by some of those who

---

[17] Notice of the State Internet Agency to seek public opinion on the Administrative Policies for Managing Generative AI-Enabled Services (draft for comments) // Available at: URL.: http://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm (accessed: 26.05.2023).

[18] On providing FSS officers with powers to send requests to credit institutions, tax authorities, agencies for state registration of real estate transactions and data system operators for digital financial assets: Order No. 97 of 16.03.2022 // Available at: URL: https://www.garant.ru/products/ipo/prime/doc/404342484/ (accessed: 11.05.2022)

[19] On approving organizational and technical data security requirements to authorized officers of certification centers of the federal executive agency authorized to register legal persons: FSS Order No. 171 of 01.05.2021 // Available at: URL: https://ppt.ru/docs/prikaz/fsb/n-171-250719 (accessed: 11.05.2022)

represent the community of judges [Laptev V.A., 2021] with all incoming documents to be put into digital form, will not only fail to remove the technological inequality (because of variable IT literacy of the population, disparity of documents filed with courts, lack of adequate technical support etc.) but will also require to assign an operator responsible for the integrity of all sensitive data of the parties to legal proceedings. No system of those currently available seems to be able to ensure either security of this process or a definitively objective outcome of automated rendering of justice through the use of AI.

In this regard, the proposed development of cloud-based AI administered via Internet is so much more risky that the access to the potential software's interface will be left actually unprotected from third-party hacking since the golden rule "an offline computer cannot be hacked" will not be observed.

Pursuant to Article 1, Federal Law No. 3132-1 FZ "On the Status of Judges" dated 26 June 1992, the judicial power is autonomous, independent and exercised by judges as natural persons, not by a neural network. A number of institutions of law are not objectively liable to be assessed by artificial intelligence:

evaluating a need in compensation for moral harm and relevant amount;

measuring the extent of influence of the controlling parties on the situation of a corporate debtor in a bankruptcy case;

identifying the nature of complicity pursuant to Article 33 of the Criminal Code;

choosing a sanction among several alternatives;

assessing whether a compensation for damage and reparation for wrong are adequate in imposing a fine to waive criminal liability or whether imprisonment (a term of sentence) should be chosen etc.

Thus, it is not quite clear how AI will assess Supreme Court Determination of 16 February 2023 on case No. 67-UD22-30-K8 that "a crime against the public order, interests of public and local government service does not prevent the case to be closed and a criminal sanction imposed"[20] — as a provision or its interpretation by a judicial authority.

Major conflicts of law in legal instruments of any branch of law, subjective glossaries and comments interpreting a particular disputable situation

---

[20] Cassation Court Determination on case No. 7-UD22-30-K8 of 16.02.2023 // Available atr: http://vsrf.ru/stor_pdf.php?id=2215490 (accessed:24.05.2023)

challenge the validity of a neural network's decision to identify the prevailing provision and to make the only right choice.

The criminal justice experience of algorithmic forecasting (profiling) in the United States (PSA (Public Safety Assessment), COMPAS (Courcional Offender Management Profiling for Alternative Sanctions)) has shown the highly random nature of the resulting assessment, only to undermine the constitutional rights of individuals to "fair trial and individualized sanction, once algorithmic assessment becomes the only basis for a court ruling" [Talapina E.V., 2022: 4–27], the more so since the responsibility for wrong predictive decisions made by AI is not assigned to anyone.

Neither the national nor international legal doctrine can answer the question who will select judicial practices for machine and subsequent deep learning and on what criteria and basis, given that AI is trainable only on Big Data. There could be hundreds of thousands or even millions of such "impeccable solutions" for each branch of law, each type of cases and issues handled by courts of all instances and panels of all competences.

It is unclear whether this should cover the cases reviewed by higher instance courts, largely insufficient for neural network learning, who will finally determine the "proper" decision-making algorithm, and whether AI will rely exclusively on statistical information generated by the analysis of absolutely all court decisions.

Thus, in late 2021 the Supreme Court of Russia adopted Determination No. 305-EC21-14231 to formulate a critically important stance whereby an enterprise or organization did not have to prove the fact of established reputation and adverse effect of defamation if no claim was made to make up for reputational loss[21]. A year later, on 8 November 2022, the same authority indicated in Determination No. 78-KG22-44-K3 that pursuant to Article 56 of the Code of Civil Procedure the claimant had to prove the circumstances underlying his claims, that is, to prove the fact of established reputation in the given field of business relationships (industry, services, education etc.) which was not at all presumed. One can only guess which of the two opposed decisions by the same authority will be assessed by AI as the only right choice.

Likewise, the proponents of judicial AI never explain whether deep learning will involve exclusively the cases which stood up on appeal or

---

[21] Determination No. 305-EC21-14231 on case No. A41-54681/2020 //Available at: URL: https://kad.arbitr.ru/Kad/Card?number=A41-54681%2F2020 (accessed: 09.12.2021)

whether the list will include those never appealed against irrespective of the impulsive cause. An appeal to a higher court against the trial court's decision involving AI will algorithmically mean the decision will be upheld if considered by the appeal and cassation courts using same interface. Thus, a "traditional" way of considering such cases has to be envisaged starting with the court of appeal.

The variable approach of courts at different levels to consider even standard cases; legal paradoxes and lack of consensus between the doctrine and enforcement practices; courts addressing certain cases on an exceptional basis for lack of clear definitions of major concepts in the national law and given a considerable number of value judgments (good faith, materiality of harm, insignificance, permanent disfigurement, mitigating circumstances, generality, custom etc.) — these things are contrary to the algorithmic nature of machine learning, only to result in probable errors both in AI-enabled analysis of certain facts and "unbiased" decision making.

One example of likely fallacies can be a hypothetically broad interpretation by AI of the provisions of the Supreme Court Plenum Resolution of 18 April 2023 on "the relatives of a police officer, military serviceman or public official"[22]. Thus, according to D. Veretennikov, "such wording… can result in neighbors, doormen (watchmen), utility workers, postmen etc. wrongly considered as relatives on the sole basis that the victim gave such evidence"[23]. Since deep learning makes AI operate on the basis of formal mathematical logic outside any preset algorithm, there is no telling what a neural network will be guided by in associating a person with "other relatives" in light of the Supreme Court's explanations.

Even basic ("narrow") AI interfaces for technical judicial functions such as recognizing and digitizing handwritten and audio documents; referring cases to the courts of relevant jurisdiction; collecting legal statistics, searching for party contact details for service of process; or performing expert functions are not feasible in the short and medium term for the said reasons such as vulnerabilities in sensitive data processing and a lack of functional operator; possible threats to national security; extensive value

---

[22] Supreme Court Plenum Resolution No. 11 "On handling criminal cases of martial offenses" dated 18 April 2023 // Available at: URL.: https://www.vsrf.ru/documents/own/32440/ (accessed: 25.05.2023)

[23] SC to protect relatives of police officers. Advokatskaya Gazeta.18.04.2023 // Available at: URL.: https://www.advgazeta.ru/novosti/vs-predlagaet-zashchitit-blizkikh-pravo-okhranitelyam-lits/ (accessed: 25.05.2023)

judgments and conflicts of law; lack of universal criteria to select "impeccable" verdicts for deep learning of neural networks; lack of multi-layered neural networks capable of evaluating actual circumstances of the case; low digital literacy of legal profession etc.

Contrary to the opinion of judicial AI proponents such as P. Morkhat and V. Momotov, Supreme Court Presidium member, there are grounds to believe that AI will not only fail to ensure "barrier-free access to justice for population" [Morkhat P.M., 2018: 6–11] and "an space for legal proceedings" [Momotov V.V., 2020] but can result in new obstacles to proper implementation of Article 46 of the Russian Constitution. The Machine-Readable Law Concept Note drafted by the Skolkovo Center in 2021 and submitted for approval to the Ministry of Economic Development has likewise failed to be implemented due to the emerging risks and despite the belief that its planned introduction would allow to reduce legal costs of individuals and to ensure transformative change of the regulatory and supervisory domains and those of administrative and legal proceedings[24].

Unlike judges, the developers of AI interfaces and of relevant roadmaps for AI implementation in the system of justice are not subject to higher reputation and qualification requirements, only to increase the likelihood of legal and reputational risk scenarios in the course of third-party development of judicial AI despite all declarations of openness and independence.

Meanwhile, neither the Supreme Court of Russia or the Constitutional Court of Russia proposed to "administer" judicial AI (GosTech (Federal Government) could not a priori administer the digitization process at courts since the judicial branch is separated from the legislative and executive branches under Article 10 of the Constitution) have adequate skills and knowledge to analyze machine learning algorithms and assess AI decision-making methodologies in a given case.

As we mentioned above, non-legislated ways to protect the personal data of litigants are a separate category of risk factors realized in using AI technologies in court practice. Utilization of AI systems in legal proceedings leads to an exponential growth in the probability of data array hacking through API-technology (Application Programming Interface, a set of tools and functions describing the interaction between the interface user (e.g., the Pravosudiye (Justice) State Automated System portal etc.) and

---

[24] Machine-readable law: a likely future? // Available at: URL: https://www.garant.ru/news/1464143/ (accessed: 19.05.2023)

the personal data operator). Russian laws do not list information, including personal data, which AI is entitled to access through API technology; there is no actual state supervision over the transfer of information containing personal data of individuals within the subsystems of government agencies, including courts, etc. on the basis of the person's consent to each such operation, which is expressly stipulated by the requirements of Law No. 152-FZ. At the same time, utilization of cloud-based distributed registry technologies with one-way encryption to collect, process and store personal data without the involvement of a single operator (Proton.mail, a webmail service with encryption, Mega file-sharing service, use a similar approach) fails to meet the provisions of Federal Security Service Order No. 97 of 16 March 2022[25] , FSS Order No. 171 of 01 May 2021[26].

If we look at blockchain technologies as a method of secure storage and decomposition of personal data in order to anonymize them and further endow them with negotiability, which a range of researchers propose as a secure alternative to steganographic and cryptographic methods of personal data protection, including data potentially processed by Judicial AI in judicial practice[27] , we see that they are not supported by the necessary legal basis required for their implementation as a protection method. Their mechanical introduction in the civil law regulation of relevant metadata circulation poses fundamental risks, both reputational and legal, for personal data operators (here, the judiciary system) and the state.

Courts lawfully process personal data (People Data), namely Volunteered Data, except for Observed & Inserved Data relating to an indirectly identifiable person, in accordance with the provisions of Russian law (Articles 6, 10, 11 of Federal Law of 27.07.2006 "On Protection of Personal Data" No. 152-FZ ;hereinafter: Law No. 152-FZ.) Other essential require-

---

[25] On Authorizing Federal Security Service Officials to Send Requests to Credit Institutions, RF Tax Bodies, Bodies Responsible for the State Registration of Rights to Immovable Property and Transactions Therewith and to Operators of Information Systems in Which Digital Financial Assets are Issued: FSS Order No. 97 of 16 March 2022 // URL: https://www.garant.ru/products/ipo/prime/doc/404342484/ (Last accessed: 11 May 2022).

[26] On Approval of Organizational and Technical Requirements in the Field of Information Security for Authorized Persons of the Certification Centre of the Federal Executive Body Authorized to Perform State Registration of Legal Entities: FSS Order No. 171 of 01.05.2021 // Available at: URL: https://ppt.ru/docs/prikaz/fsb/n-171-250719 (Last accessed: 11 May 2022).

[27] Kozin I.S. A Method of Ensuring Secure Personal Data Processing on the Basis of Blockchain Technology. Scientific and Technical Bulletin of Information Technologies, Mechanics and Optics. 2019. No. 5. Pp. 892–899.

ments include compliance with general principles of processing (Article 5, Federal Law No. 152-FZ)[28] ; the operator should perform data localization, notify Roskomnadzor (Federal Service for Supervision of Communications, Information Technology, and Mass Media), and undertake organizational and technical measures for personal data protection (Articles 22, 18, 18.1–19 of Federal Law No. 152-FZ)[29] , and the person in question should give a specific, conscious and informed consent.

In its decision on Case No. A40-5250/17-144-51, Roskomnadzor expressed a more stringent position stating: "It is not possible to assert without the written consent of the user that the data was provided by the person in question and that the applicant's actions violate Paragraph 1, Part 1, Article 6 of Federal Law No. 152-FZ (processing of data without the person's consent)."[30]

Hence, blockchain technology as a method for AI to ensure secure storage and use of People Data arrays in course of judicial activities contradicts the very idea of both the law on personal data and the regulator's position because it implies decentralization and public availability of information, where personal data can be provided to all participants of the distributed registry. While preventing direct data leaks at any given moment, the distributed registry (e.g., containing data of litigants in concrete proceedings) violates the basic principle of law: one purpose—one consent—one recipient.

If AI is utilized to administer a distributed registry (appoint a person responsible for inclusion/exclusion from the register; for completeness, reliability and procedure of information use), then such register ceases to be a distributed register. If it is a classical peer-to-peer blockchain with free data flow, it is impossible to prevent misuse of data, including personal data.

The above entails clear legal risks for the corporate data operator (registry holder) (in this case, a specific judicial authority) arising from the use of personal data by AI systems. The risks, in particular, include penalties under criminal liability (Art. 137 of the RF Criminal Code "Violation of Personal Privacy"), administrative liability (Art. 13.11 of the RF Code of

---

[28] Available at: URL: http://www.consultant.ru/document/cons_doc_LAW_61801/ (accessed: 08.05. 2022)

[29] Ibid.

[30] Judgment of Arbitration Court of Moscow on Case No. A40-5250/17-144-51. 05 May 2017 // Available at: URL: https://sudact.ru/arbitral/doc/YLVZ7F3cAwU0/ (accessed: 04.05. 2022)

Administrative Offences "Violation of the Law of the Russian Federation in the Field of Personal Data"), civil liability (Art. 15 "Recovery of Losses due to Violation of Personal Data Processing Rules", Art. 151 of RF Civil Code "Compensation of Moral Damages", Art. 24 of Federal Law No. 152-FZ "Infliction of Emotional Distress (Moral Suffering) to a Citizen due to Violation of Personal Data Processing Rules"), and disciplinary liability (Art. 90, 192 "Other Violations in the Field of Personal Data Processing", subparagraph "C", Paragraph 6, Subparagraph 1, Art. 81, "Disclosure of Personal Data by an Employee" of the RF Labor Code.)

If the above risk scenarios are realized through the fault of the data operator, this will not only entail quite material losses for the judicial system, but also a heavy blow to the business reputation of the judiciary bodies.

The law does not regulate the issue of civil and criminal liability arising from the use of AI systems in the administration of justice: e.g., the person responsible for making a likely inappropriate, unlawful, erroneous decision in a particular AI technology case has not been identified.

In the case of the "Chinese" scenario of introducing AI into judicial practice, the software developer is likely to become the responsible person. In this case, the courts will additionally have the responsibility to protect data processed by AI.

It is worth noting that, unlike the Peoples' Republic of China, the Russia does not have a structure, which is similar to the National Data Administration of China (so there is a reason why NDA, the acronym name of this body, coincides with the common designation of a confidentiality agreement) and which should be responsible, among other things, for the circulation of personal data processed by AI.

Several foreign researchers note that as a result of implementation of AI technologies in the judicial practice, e.g., in China, "one should get ready for the undermining of the judiciary by technology companies and the capital ."[31] The intelligent judicial SoS (system of systems) is now connected to the desktop of each judge in the PRC. Based on machine learning technology, it automatically searches for similar cases, "suggests" laws and regulations, drafts legal documents, and modifies alleged human errors in the verdict, if any. According to Xu Jianfeng, director of the Information Centre of the

---

[31] China's court AI reaches every corner of justice system, advising judges and streamlining punishment // URL.: https://www.scmp.com/news/china/science/article/3185140/chinas-court-ai-reaches-every-corner-justice-system-advising (accessed: 27.05. 2023)

Supreme People's Court of China[32], this raises a number of questions about determining who is liable for a judgement that is made using robotic systems. While Artificial Intelligence advises judges and optimizes punishment, including through the Smart Court interface that allows the system to access police, prosecution and government databases and integrate with China's social credit system, it is not an entity responsible for the decisions it actually makes as a result of certain generative actions.

The Supreme People's Court of the People's Republic of China requires that a judge must consult with artificial intelligence on every case; if the judge rejects the AI's recommendation, the system requires a written explanation for subsequent audit. As a result, judges predictably strive to follow recommendations to avoid having to "challenge the system," even if the artificial intelligence chooses a less appropriate reference or law in a particular case. The result is a decision that is not always optimal and lawful.

In this regard, it has a sense to believe that even before the introduction of Judicial AGI elements into the domestic judicial practice, it would be justified and necessary to legally assign the responsibility for errors or other legal consequences arising in the process of judicial proceedings involving AI to the institution (authority) that licensed a particular AI interface to participate in judicial practice. An institution (authority), which actually owns a certain software product and recognizes its "legal integrity" for participation in the judicial process (in Russia it is the Judicial Department at the Russian Supreme Court), thereby assumes the burden of responsibility for judgments made using this product. A similar conclusion can be drawn in cases where an unmanned aerial system involving artificial intelligence makes decisions on the elimination of a person in a combat, and the responsibility rests not with the developer of the UAV interface or the executor of the order (the serviceman), but with the agency that is the balance holder of the unmanned system equipped with a particular software program.

The legislator and law enforcement agencies advocating the use of AI in the system of justice have equally failed to fully appreciate the risk of discrimination against the parties to legal proceedings created or reproduced by AI as a result of algorithmic bias [Kharitonova Yu. S., Savina V.S., Panyini F., 2021: 488–515], as well as risk scenarios related to vulnerabilities, automation errors, network failures. For example, network gateways (traf-

---

[32] Ibid.

fic control servers between the local network of the national justice system and the Internet) are fully produced by international companies such as Cisco, Huawei, Panasonic etc. Thus, the system cannot be safe from possible attacks either now or in the near future. While a sensitive data theft in a cellular network is a dangerous invasion of privacy, a potential hacking of the judicial system relying on the infrastructure of unfriendly countries is fraught with violation of human rights and liberties envisaged by Chapter 2 of the Constitution, particularly in AI-enabled decision-making on criminal cases.

"We first create the core of the model and teach it to operate with words, remember their combinations, make logical chains… Next comes a superstructure to carry certain meanings… It will later manage all processes. If we adjust (the superstructure) to handle regulations, it will produce — just as a lawyer –specific answers to specific questions without any offhand interpretations… The software will develop an understanding of what is expected from it. The question is who sets the selection criteria as an expert and for what purpose"[33] — this is how a domestic developer describes the creation of a next judicial AI interface.

Conclusion

The disparity of learning sources and their selection criteria, uncertainty of input meanings currently appear to be a key problem that cannot be resolved in the current regulatory and enforcement context in introducing AI and its derivatives into the national system of justice. The lack of legislative recognition of subjects of responsibility for decisions made using AGI makes the corresponding initiatives for its implementation in judicial practice not only careless, but dangerous.

In this regard, the reference of AI proponents to Argentine where Prometea, an AI-enabled interface, has been used since 2018 for independent analysis of circumstances on standard lawsuits, with the decisions 100 percent ratified by judges [Atazhanov A., Ismailov B., 2020: 269-284] appears to be misplaced. The Laser Program to generate well-founded decisions "on the basis of in-depth analysis of case circumstances and similar decisions" [Stepanov O.A., 2022: 229–237] has failed to be implemented in

---

[33] Russia to actively develop substitutes for ChatGPT // Available at: URL.: https://therussiannews.ru/news/technologies/v-rossii-aktivno-razrabatyvayut-analogi-chatgpt/ (accessed: 25.05.2023).

the national justice system largely because of the emerging risk scenarios. V. Shananin noted in addition that "artificial intellect should be implemented exclusively on the principles of human control, selection and priority" [Shananin V.A., 2022: 143–146].

At the same time, an optimal combination of the national justice system with AI as a key digitization technology without drifting towards regulatory arbitration; proactive compliance policy of development companies and enforcement agencies coupled with active adoption of new regulations can provide an adequate basis for supporting a global trend to make AI a major competitive factor in both domestic and international markets and an additional driver of economic growth of the Russian business.

References

1. Atazhanov A., Ismailov B. (2020) International experience of introducing modern technologies into the justice system. *Obschestvo i innovatsii*=Society and Innovations, no. 2, pp. 269–284 (in Russ.)

2. Barak A. (1999) *Judicial discretion*. Moscow: Norma, 376 p. (in Russ.)

3. Galkin D.V., Stepanov A.V. (2021) Security aspects of military AI applications. *Voennaya mysl*=Military Thought, no. 4, pp. 72–79 (in Russ.)

4. Haskins A., Arora S., Nilawar U. (2017) Impact of artificial intelligence on Indian real estate: transformation ahead. Madras: Colliers Radar Property Research, 13 p.

5. Kharitonova Yu.S., Savina V.S., Panyini F. (2021) A bias of AI algorithms: issues of ethics and law. *Vestnik Permskogo* gosudarstvennogo *universiteta*=Perm State University Bulletin, issue 53, pp. 488–515 (in Russ.)

6. Kok J., Boers E., Kosters W. et al. (2009) Artificial intelligence: definition, trends, techniques, and cases. In: Encyclopedia of life support systems. Artificial intelligence. J.N. Kok (ed.). Paris: Eolss Publishers, 401 p.

7. Laptev V.A. (2019) The concept of artificial intelligence and liability. *Pravo. Zhurnal Vysshey shkoly ekonomiki*=Law. Journal of the Higher School of Economics, vol. 12, no. 2, pp. 79–102 (in Russ.)

8. Makarikhina O.A. (2014) On judicial discretion in civil and arbitration proceedings. *Arbitrazhniy i grazhdanskiy protsess*=Arbitration and Civil Process, no. 6, pp. 14–17 (in Russ.)

9. Morkhat P.M. (2018) Judicial AI as a way to overcome judicial discretion. *Teoriya i istoriya prava i gosudarstva*=Theory and History of Law and State, no. 5, pp. 6–11 (in Russ.)

10. Morkhat P.M. (2017) On defining the concept of artificial intelligence. *Teoriya i istoriya prava i gosudarstva*=Theory and History of Law and State, no. 12, pp. 25 –32 (in Russ.)

11. Pokrovskiy I.A. (1998) Main problems of civil law. Moscow: Statut, 349 p. (in Russ.)

12. Rissland E. (1990) Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning. *Yale Law Journal,* vol. 99, no. 8, pp. 1957–1981.

13. Russell S., Norvig P. (2010) Artificial intelligence: a modern approach. Boston: Prentice Hall, 1132 p.

14. Shananin V.A. (2022) Using AI system in judicial practice. *Yuridicheskaya nauka*=Legal Science, no. 11, pp. 143–146 (in Russ.)

15. Stepanov O.A., Basangov D.A. (2022) On the prospects of AI impact on legal proceedings. *Vestnik Tomskogo* gosudarstvennogo *universiteta*=Bulletin of Tomsk State University, no. 5, pp. 229–237 (in Russ.)

16. Strawson P. (1952) *Introduction to logical theory.* London: Macmillan, 266 p.

17. Talapina E.B. (2022) AI-aided data processing and discrimination risks. *Pravo. Zhurnal Vysshey shkoly ekonomiki*=Law. Journal of the Higher School of Economics, vol.15, no. 1, pp. 4–27 (in Russ.)

18. Tretyakova T.N., Karamanukyan D.T. (2020) The concept of judicial discretion. *International journal of professional science,* no. 2, pp. 5–8 (in Russ.)

Information about the author:

V.A. Rodikova — Postgraduate Student.

Artificial Intelligence Governance and China's Experience under the Community of Common Destiny for Mankind Concept

Jia Shaoxue

Center for International Legal Training and Cooperation for the SCO, Shanghai University of Political Science and Law, 7989 Weiqingsong Ave., Qingpu District, Shanghai 201701, China

Abstract

In recent years artificial intelligence (AI), backed by big data and the Internet, has been rapidly developing and determining the future direction of the world's science and technology development. Although artificial intelligence is beneficial to the scientific and technological revolution and industrial modernization of mankind, it has also brought new risks. People pay more and more attention to the potential risks of artificial intelligence that should be effectively managed. Artificial intelligence risks are characterized by the diversity of technological threats, the similarity of AI risks faced by different countries and the high complexity of governance, something that requires concerted efforts by all countries. It is necessary to carry out the development of artificial intelligence in the country from the perspective of the common interests of mankind, ensure the safety and manageability of artificial intelligence, and strengthen international cooperation. At present Western countries advocate the concept of technological hegemony and technological monopoly, and developing countries have little opportunity to express their opinions on the governance of artificial intelligence, and China's Community of the Common Destiny for Mankind Concept is necessary for the governance of artificial intelligence. Based on that concept, the paper explores China's new practices and proposals for the domestic and international AI governance. In response to the problem of overuse and misuse of new technologies, China proposes to establish an artificial intelligence governance system that includes joint management by various actors, open and transparent regulation, comprehensive consultations, and the development of effective evidence-

based laws, so as to promote the beneficial development of artificial intelligence in the future and contribute to the deepening of AI governance based on the Chinese proposal.

Keywords

global governance; artificial intelligence; risks; community of common destiny for mankind; Chinese experience.

Background

Humanity has embraced the age of artificial intelligence. A major driving force of the fourth industrial revolution, AI technology is giving a new lease of life to such important sectors as military science, finance, education, science and technology, culture etc., while providing enormous capabilities for the historical evolution of humankind and creating a new model of global development [Shen X., Shi B., 2018:15]. AI is shaping the future of human society in an unprecedented way. While countries take the inherent challenges of AI technologies seriously, uncertainty of the risks is a major social concern. In the context of already existing or likely threats in the course of AI evolution, countries need to manage and regulate these risks as a matter of priority. At present, both domestic and international academic circles lack an analysis of the Chinese concept and approach to AI governance. For this reason, this paper dwells on China's Community of Common Destiny for Mankind concept to discuss the peculiarities of AI governance and Chinese proposal to manage AI with the global development prospects in view.

1. Specifics of AI Threats

An enormous commercial and social value of AI technologies is now propagating them across different spheres of life. As a new generation of information technologies, AI normally exists in the form of software and hardware to include a host of applications responding to vision, hearing and different sensory stimuli such as imitation of human games, language

translation, automated driving, face recognition etc. Depending on the use, the following three AI categories could be distinguished: weak artificial intelligence, artificial general intelligence, artificial superintelligence.

Weak artificial intelligence covers AI technologies endowed with some cognitive capability and widely used in everyday life, such as voice recognition, translation, face recognition etc. This type of AI has enjoyed the most large-scale development and marketing success.

Artificial general intelligence has cognitive ability matching that of man, with a single AI system able to perform a multitude of cognitive activities and behave intelligently, such as managing unmanned combat aircraft for an autonomous analysis of terrain and assessment of threats, functioning as generative AI, etc. [Zhang L., 2023: 126–128].

Artificial superintelligence has a cognitive ability beyond that of man, only to surpass man in such spheres as scientific innovations and autonomous production of knowledge.

Thanks to a breakthrough in data science, computing capabilities and algorithms, AI has entered a new age of explosive development. Some researchers believe that AI will evolve exponentially, once the singularity limit is overcome [Han Y., Zhang F., Peng J., 2023: 122]. The greatest peculiarity of artificial intelligence is the likelihood of becoming self-conscious in the future [Yu N., 2017: 95–96]. If AI is not guided by human standards and not restrained in its growth, the risk will become unmanageable. AI technologies have inherent risks and threats which for practical governance translate into the following aspects:

First, AI-related threats are diverse. Technological threats are largely concentrated in the following spheres: in the military sphere, AI is able to make independent decisions while its ability to collect and analyze huge amounts of data can undermine the traditional methods of warfare such as the use of unmanned aircraft and other types of arms, only to increase the gap in military power between countries. As regards the economy, AI will replace man and change the future of work to inevitably generate deeply rooted conflicts in the global social structure resulting in segregation and inequality [Ma C., 2018: 48–55]. AI also affects the industrial development at the national level which can create financial risks, sectoral monopolies and other negative implications. In the social sector, AI technologies are subject to algorithmic discrimination and biases to give rise to legal and moral dilemmas fraught with considerably more violations of privacy and ethical risks. In short, AI technologies gradually affect human behavior and

result in risks not predicted by system developers, only to engender multiple threats for human society with regard to employment, law, privacy, ethics and security [Wu S., Luo J., 2018: 112–114].

Second, countries face similar AI risks. In the age of globalization, many AI-related issues of political governance are of global importance. As AI technologies spread out, the underlying risks grow and progress across the world, with characteristically cross-border dissemination from the national to international level. No country, organization or person can independently handle AI technological threats. The reliance of artificial intelligence on big data for algorithmic operation results in security risks such as personal data and state secret leakages — for example, widespread theft of personal data, intrusion into public networks and loss of control over national data, something that no sovereign state with a traditional closed governance system can be safe from. Moreover, the global use of AI technologies is faced with general problems. For example, there is research across the world to develop self-driving vehicles, only to result in numerous legal problems. Who will be held responsible in the event of an accident between a self-driving vehicle and a human driver, if neither party is at fault? There is a need to assign and accept the relevant legal obligations.

Third, AI-related risk management is complex one. AI governance has moved beyond the scope of relationships between individuals up to the intergovernmental level and, not confined to the protection of privacy, data leakage etc., extends to the level of human consciousness, operation of state and society. Countries are faced with the choice of methods to govern AI. Disputes between countries on who has the right to formulate and interpret the AI governance rules have made the global cooperation in this area problematic. With the importance of AI governance recognized worldwide over the last few years, the trend for an independent way has become relatively obvious in practice. Governing AI requires not only a generally accepted concept but also technical implementation in the form of rules. Since AI technologies have a global scope in different countries, cultures and spheres, the success will depend on cognitive understanding of each country and at the same time on the extent of concerted action by the international community.

2. Pressing Problems of AI Governance

To explore AI governance, one need to first have basic theoretical understanding. The core AI governance elements are several and include subjects,

objects and methods of governance. The first largely include governments, international organizations, public institutions etc. Objects of governance include AI technologies themselves and related problems. Methods — that is, specific means and policies to govern AI — largely cover ethical limitations, technological innovations, regulatory and legal provisions reflecting the rules, concepts and the underlying values to be observed in using AI technologies. The joint efforts of all subjects result in control over an object of governance to provide a basis for addressing global challenges and transnational threats which emerge in the process of technological change. A current controversy over who should direct AI governance, what is efficient as regulation, what values should be upheld and what methods adopted prevents collective action, with AI governance becoming a global social problem affecting the interests of the population at large, problem of competition and difficulties to conceptualize values.

2.1. Diverging Interests of AI Governance Subjects

AI governance subjects include public authorities, non-governmental organizations (NGOs), enterprises, research centers, private individuals etc. Governments assume the leading role in AI application, research and development; high-tech companies act as developers and suppliers while NGOs, research institutions and individuals are important parties in terms of relevant assessment and opinion. Depending on the governance scope, the activities of numerous parties involved in this process normally take place on two different levels: national and international.

AI governance at the national level is a very complex task involving conflicts of interest between different subjects. While new AI should be regulated, overly rigid regulation will obstruct technological process, with businesses interested in minimal provisions for more profits and room for independent decision-making; meanwhile, the public sector will opt for stability and security including to avoid non-ethical and illegitimate use of the technology. NGOs, research centers and individuals are watchdogs to guard against moral prejudice, discrimination, racism, human rights and other AI-related problems, and perform the monitoring function with regard to public opinion by producing societal moral judgments in the process of governance. In the age of AI, too strict or relaxed regimes will concern the interests of each subject resulting in a chaotic AI governance pattern from the perspective of law, ethics and economy.

AI governance at the international level is not only an intrinsically technical problem but also that of international development standards. The economic foundations and technological resources of countries are imbalanced on the international scale, with evident disproportions and deviations observed in the technological development. The international community has realized that common standards contribute to global development of AI but the harmonization process is not simple. Moreover, technological AI standards vary across countries and regions. Policy development will normally fall behind the speed of technological change: decision-makers cannot fully understand AI for lack of adequate experience, only to make wrong decisions, with cooperation mechanism between civil servants and technology researchers often absent. Moreover, despite the adoption of certain national technological standards by the international community, international organizations at the sectoral level cannot engage in a technical dialogue for lack of the relevant practical experience.

2.2. Aggravating Competition for AI Governance

Thanks to an overall breakthrough in three core components — data, algorithms and arithmetic capability — AI has demonstrated a capacity matching or even surpassing that of man in spheres such as education and technologies, traffic management, financial investment, legal proceedings to become a field for competition between countries [Li C., 2021: 127–128]. The progress in AI technologies is related with the increase of competitiveness. The international technological rules and coordination mechanisms applicable to AI are currently dominated by the governments of developed countries such as the United States. Over the last few years the United States, European Union, OECD and other large countries and organizations worldwide have been following each other in launching AI policy plans to resolve pressing issues.

The European AI Strategy builds on trust as a prerequisite of the human-centered approach to AI. In April 2019, the European Commission published the Building Trust in Human-Centric Artificial Intelligence[1], a document describing the key requirements and concept of trustworthy AI presented by the High-Level Expert Group on AI in the Ethics Guidelines for Trustworthy AI. According to the Guidelines, a trustworthy AI should be: lawful — respecting all applicable laws and regulations; ethical — re-

---

[1] Available at: https://ec.europa.eu/digital-single-market/en/news/communication-building-trust-human-centric-artificial-intelligence (accessed: 01.09.2023)

specting with ethical principles and values; robust — both from a technical perspective and taking into account its social environment[2].

On 22 May 2019 the OECD countries officially approved the first package of intergovernmental AI principles by approving international standards of robustness, security, sustainability, fairness and safety of AI systems[3].

In May 2023 the US Administration published a new National AI R&D Strategic Plan defining key priorities and purposes of the Federal Government's investments into AI research and development[4]. As part of the international efforts to ensure responsible use of AI, the G7 initiated the same month the AI Hiroshima Process which promotes an open and constructive dialogue on the implications of AI tools such as ChatGPT, an AI model supported by Microsoft OpenAI. Moreover, at the Hiroshima summit the G7 leaders stressed the need in developing and adopting the relevant technical standards to support AI "robustness". They also noted the importance of ensuring the compliance of AI advances with common democratic values[5].

On 8 June 2023 the United States and United Kingdom approved the Atlantic Declaration for economic partnership underlined the need in further strengthening of cooperation in such fields as artificial intelligence to ensure the American and British leadership in the key and novel technologies[6]. This Declaration reaffirms the fact that Western countries will be fully involved in the global governance of the emerging technologies to make it a major field for further discussions and global leadership.

Governments currently regard AI technologies as a key to the future of their countries, thus manifesting a clearly national interest. For lack of a major coordinating body vested with absolute powers, many countries have set for a dominant position in AI governance rules on the argument of technological gap and technological inequality. Striving to secure the maximum domination, Western countries headed by the United States are tak-

---

[2] Available at: https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai (accessed: 31.08.2023)

[3] Available at: https://globalcentre.hse.ru/news/276245330.html?ysclid=lmbc9egy9e 516596400 (accessed: 02.09.2023)

[4] Available at: https://d-russia.ru/administracija-ssha-opublikovala-novyj-strategiche-skij-plan-issledovanij-i-razrabotok-v-oblasti-iskusstvennogo-intellekta.html (accessed: 03.09.2023)

[5] Available at: https://www.fullrio.com/economy-70466 (accessed: 03.09.2023)

6 Available at: https://baijiahao.baidu.com/s?id=1771350683397528978&wfr=spider& for=pc (accessed: 03.09.2023)

ing steps to hold back the developing countries and break away from them in the technological development, only to further undermine the cooperative nature of the global AI governance. In the age of artificial intelligence, the workforce from developing countries is involved in the international division of labour on much looser terms, with the governing power of sovereign countries in decline [Han Y., Zhang F., Peng J., 2023: 138-139]. The problem of technological inequality is obstructing technological progress in developing countries while an enormous potential of the leading nations may finally result in a technological hegemony [Mei L., 2023: 53]. Today developing countries do not have much to say on AI governance as the projects they are involved in are relatively few. For this reason, developing countries need to constantly improve their technological potential in this field and promote a reform of the existing global system of AI governance in the interest of their own development.

2.3. Lack of Value-based Consensus in AI Governance

The process of technological development of AI is closely related with the world's global development path, civilizational concepts and ideologies. Due to the specifics of political systems, national contexts and cultural traditions the AI-related technological policies worldwide largely differ and have different values in view. Many countries are attempting to impose AI values to promote their own development needs and interests while sticking to a technological model underpinned by their core values. While on 25 November 2021 the United Nations Educational, Scientific and Cultural Organization (UNESCO) adopted the Recommendation on the Ethics of Artificial Intelligence[7], the first ever global standard on AI ethics, such global awareness standards are few and not binding.

AI should be underpinned by right values with fairness and equity as the main value-based principles, otherwise it can depart from its original purpose to become a tool for those in power to keep their privileges [Sun W., 2017:120-126]. There is currently no universally applicable regulatory system with international AI governance rules [Zhu M., Xu C., 2023: 1037-1049]. From the global perspective, the diverging governance concepts in different countries are a major problem for AI governance. AI systems worldwide are influenced by different values, only to prevent the effective cooperation for global AI governance. Western countries headed by the

---

[7] Available at https://d-russia.ru/junesko-prinjala-rekomendaciju-ob-jeticheskih-aspektah-iskusstvennogo-intellekta.html?ysclid=lmbgz5ysye848477271 (accessed: 02.09.2023)

United States are attempting to dominate by combining values with technological monopoly to keep their leadership and the established world order. Apart from the technology as such, the United States clearly showcase the values of the Western world — freedom, democracy and the rule of law — while underlining that AI technologies should comply with American values and interests. Imposing "unilateral" values on other countries aggravates the existing conflicts in the global AI governance.

3. China's AI Governance Proposals

The extent of reducing AI-related risks depends on evidence-based governance mechanism to be created. Based on the regulatory framework, countries should ensure ethical support for the development of relevant technologies and interpretation of algorithms, as well as a proper balance between technical responsibility and ethics, so that AI systems could be used in a fair, transparent and safe environment. In a deeper sense, AI challenges the subjective status of "man", only to pose questions such as how man and machine can co-exist; how the legal liability between man and machine can be defined; and how AI's legal status and liability can be determined. The primary purpose of AI governance is to ensure safety of man, so that machines would comply with the existing moral and value-based human attitudes. China's concept of the Community of Common Destiny for Mankind does not only follow the logic of common human development but also paves a realistic way to address the AI development dilemma.

3.1. Promoting collective governance of multiple subjects

AI development involves multiple stakeholders to require collective participation of many subjects in AI governance. Domestically, each state has to establish linkages between companies and government agencies to shape a cooperation model in a competitive context; companies should be made to comply with their social obligations and follow the principles of safety in developing and applying new technologies; civil society should play a monitoring role in achieving social consensus, promoting common goals and improving the efficiency of AI governance. Globally, there is an evidently one-sided trend in AI governance, with developing countries less involved in these efforts and unable to make their voice heard. The AI governance community should involve not only developed economies but also developing nations. From the global perspective, AI governance concerns common interests of all mankind, something that requires to fully account for a balanced devel-

opment of global AI technologies and give more voting power to developing countries to promote a favourable impact of AI technologies worldwide by bridging the digital divide. AI governance should promote openness and cooperation, fully mobilize the enthusiasm of multiple stakeholders, shape a truly multi-principle model of governance involving national public authorities, R&D companies, international organizations and civil society, bring together the existing governance platforms and institutions on the global scale, create and improve a wider platform for international cooperation.

The Community of Common Destiny for Mankind concept assumes joint consultations and cooperation, and respect for common interests of all mankind. Being part of the overall structure of the Community, the global multi-level synergetic cooperation is a major element of China's involvement and promotion of the global AI governance. AI development is hinged on the synergetic governance system shaped at the global level by cooperation between all parties. China adheres to the Community of Common Destiny for Mankind, opposes the technological monopoly of a few countries in AI and focuses on joint cooperation of all countries and especially on technological exchanges between developing countries. In 2017, China created the Agency for Promotion of Development Planning of New Generation AI to organize and implement development planning of new generation AI and major R&D projects[8]. Universities, research centers and companies have established AI committees for advice on AI-related implementation issues. The government has created a system for promoting AI advances for better social governance and guidance. An important goal is to develop sectoral guidelines for self-regulation of the AI sector and for sharing of the best AI development practices with those in need.

3.2. Creating an open and transparent regulatory mechanism

A majority of AI-related innovations are implemented by the global technological powers such as Western countries, hence the importance of the policies for responsible technological regulation for these countries. In the globalized world, the question of how to minimize the negative implications of technologies through regulation is the key for governing AI. Europe and the United States are reinforcing the legal regulation of AI. On 14 June 2023 the European Parliament has voted for the approval of the AI Act which became the world's first comprehensive regulation on AI to pass the parliamentary pro-

---

[8] Available at: https://www.sohu.com/a/646248011_121106842 (accessed: 02.09.2023)

cess. The Act purports to supervise AI systems by classifying them across four risk categories ranging from "minimum risk" to "unacceptable risk"[9].

In terms of normative regulation and innovations, countries largely differ from each other. To address this issue, governments should strike an optimal balance between regulation and innovation without going to extremes like over-regulation or whateverism. Moreover, since a common regulatory system serving the interests of a few technologically advanced countries will inevitably fail, it is necessary to take into account common regulatory interests of countries as part of international cooperation and set up a multi-party AI regulatory network. With each country being a "regulator" and "competitor" at the same time, they will jointly work to establish a common, transparent and interpretable regulatory regime for AI.

While attaching much importance to the development and use of AI technology, China proposed to make AI part of the national development priorities back in 2015 to promote deep integration of AI into political and social life and to use the national leadership for regulatory guidance to ensure sound and robust development of the AI industry. Domestically, China strives to improve the open and transparent AI regulatory system and to develop a system of rapid response to technological risks. At the international level, the focus is made on the joint regulatory involvement of the global community. China proposes to step up the research on the global issues of common interest, upholds the creation of international organizations on AI and joint development of the relevant international standards. With reliance on the Community of the Common Destiny for Mankind concept, China is hoping to promote common and transparent regulatory standards at the global level for safe and widespread use of AI technologies.

3.3. Towards the Principle of Comprehensive Consultations

The social impact of artificial intelligence is largely about human values. The strife towards universal human values embraces peace, equity, development, justice, democracy, freedom etc. In view of the complexity and diversity of human societies and abstract expression of human culture, the global AI governance consensus should be upheld by the principal question — common destiny of mankind. Since AI affects all mankind, AI governance should be human-centered and provide for human interests and human agency. Governments and societies should collectively work to ensure

---

[9] Available at: http://news.sohu.com/a/687493860_121124603 (accessed: 02.09.2023)

human autonomy in governance practices [Gao Q., 2020:101]. At present, several international organizations have proposed reference frameworks for regulating AI governance but the global AI governance mechanism is yet to be improved, with no common reference framework in place.

In June 2019 China has published the Principles of New Generation AI Governance for Responsible AI[10]. It differs from AI guidance issued by other countries in more focus on the importance of jointly building the Community of Common Destiny for Mankind for sustainable economic, social and environmental development based on the cooperative model rather than the one dominated by any single country. In particular, China has put forward eight principles including harmony and friendship, integrity and equity, inclusion and joint use, respect for privacy, safety and control, common responsibility, openness and cooperation, flexible governance. AI should be developed to preserve social stability, with responsible AI to be implemented on the basis of a comprehensive review of risk management initiatives. Thus, the Principles of New Generation AI Governance for Responsible AI encourage coordination and cooperation between global organizations, public authorities, research centers, education institutions, companies, civil society and population for promoting AI development and governance, as well as underline the need in a broad consensus with regard to the international AI governance system, standards and norms established with the help of international organizations.

In June 2020 China's research centers published the White Book for Sustainable AI Development putting forward for the first time ever the principle of sustainable AI development based o"respect for consultations and study of the engagement culture"as well as the solution to future AI governance problem by"promoting sustainable development of the AI industry and creating the Community of Common Destiny for Mankind[11]. In promoting the synergetic cooperation between countries, China is striving to set up a cooperative platform on AI, with relevant issues proposed for the agenda of the G20, APEC and BRICS workshops or those held on a bilateral basis. China advocates a global AI governance mechanism based on the Community of Common Destiny for Mankind concept to make sure that AI governance serves to achieve common good, remove the digital divide, ensure social equity and justice, observe moral and ethical standards, contribute to the progress of human civilization.

---

[10] Available at: https://m.gmw.cn/baijia/2021-06/29/34959031.html (accessed: 05.09. 2023)

[11] Available at: https://baijiahao.baidu.com/s?id=1670258881368998719&wfr=spider &for=pc (accessed: 05.09.2023)

3.4. Developing and Approving Effective and Evidence-based Laws

The spectacular development of AI technologies since the early XXI century had a considerable impact on the existing legal system and public governance methods, with the disruption of law and order being a major challenge faced by mankind. The regulatory failure and disruption of law and order are manifested at the central level as "governance deficiency" [Zhang W., 2021: 18–23]. The AI challenge theoretically means that certain traditional legal concepts or views no longer compatible with AI are to be amended accordingly [Chen J., 2018: 137–138]. The inadequacy of laws and regulations to identify persons at law and assign liability for AI products can impact the development of related sectors. China actively advocates "human-centered" AI serving "a good cause". China's AI governance system is now evolving towards comprehensive and delicate governance based on exploring a possibility to have a governance system combining "soft ethics" and "tough law". Academic lawyers have conducted profound studies in the area of data rights, confidentiality in the Internet, personal data rights, core human rights and other aspects of different subjects, with positive results being achieved [Chen P., 2018: 71–72]. Following the idea of security and parallel development, China has adopted a number of underlying laws and policies to regulate and transform the new generation of AI technologies.

The improvement of laws and regulations on AI-related data security comes first. Public and regulatory authorities have adopted the relevant regulations to respond to regulatory needs in their respective domains in a positive way. The Provision for the Development of New Generation AI (published on 20 July 2017) is a policy document to develop AI in China before 2030 with a focus on the goals, key objectives and guarantees of the new generation AI. Based on this document, China has adopted and made effective a number of regulations such as the Law on Personal Data Protection (in force since 1 November 2021) which provides that no organization or individual can illegally gather, use, process or transmit personal data of other individuals, illegally offer for sale, provide or disclose personal data of others; they should not engage in personal data processing operations that pose a threat to national security or public interest. The Law on Data Security (in force since 1 September 2021) expands the importance and scope of data application with a special focus on the data security regime.

Defining the development limits of AI technologies comes second. AI is an emerging technology to be regulated with an adequate account for

innovative developments and applications while providing for tighter regulation of the legal liability of developers, suppliers and users, and defining their core obligations by formulating relevant laws. In 2020, the legislative plan of the Standing Committee of the National People's Congress (SC NPC) mentioned AI-related legislation and regulation by explicitly noting a need to focus on legal issues related to new technologies and fields such as artificial intelligence, blockchain and gene editing. To implement the legislative plan, China has adopted the following regulations: Ethical Code for the New Generation AI (published on 25 September 2021) to guide natural and legal persons involved in AI-related activities on ethical standards; Provisions on Promoting the AI Industry Development in the Shenzhen Special Economic Zone (in force since 1 November 2022), China's first bylaw to promote the sector's development; Provision on Managing Algorithmic Recommendations for Web-Based Information Services (in force since 1 March 2022) to impose the main responsibility for algorithmic security on platform companies and to provide users with the right to chose recommendations and delete data labels; it also contains a clear requirement to algorithmic recommendation services to observe public morals [Xu K., 2022: 125–130]; Provision on Governing the Deep Synthesis of Web-Based Information Services (in force since 10 January 2023) which provides that deep synthesis technologies cannot be used for any activity prohibited by laws and regulations, with suppliers to assume the main responsibility for security. All this reflects the value-based focus on disseminating algorithms for common good at the level of algorithmic governance in China. Finally, the Time-Bound Policies for Governing the Generative AI Services (in force since 15 August 2023) contribute to managing the relevant risks as a bylaw applicable to generative AI [Zhang X., 2023: 43–48].

Conclusions

The development of laws and regulations in different countries lags behind and is not adequate to the explosive growth of AI technologies, with the issues of how to determine the vector of technological progress, set up a platform for cooperation, formulate governance standards and assign risks and responsibilities yet to be properly resolved. As the world becomes globalized, a major factor is the development of common and coordinated AI governance rules, something that requires a unanimous consensus between countries on global governance vectors and rules to be achieved with the help of international mechanisms. To address common problems faced

by mankind, China has put forward the Community of Common Destiny for Mankind concept as a clear reference for promoting global AI governance, and argues for stronger international cooperation based on equality and mutual assistance, with all countries to achieve the shared use of information and collectively establish the AI governance system.

References

1. Chen P. (2018) On the principle of subjectivity in the construction of network legal rights. *Zhong guo fa xue*=Chinese Law, no. 3, pp. 71–88 (in Chinese)

2. Chen P. (2019) The power of algorithms: application and regulation. *Zhe jiang she hui ke xue*=Zhejiang Social Science, no. 4, pp. 52–58 (in Chinese)

3. Chen P. (2019) Government in the era of smart governance: risk prevention and capacity enhancement. *Ning xia she hui ke xue*=Ningxia Social Science, no.1, pp. 95–104 (in Chinese)

4. Chen J. (2018) Legal challenges of artificial intelligence: Where should we start? *Bi jiao fa yan jiu*=Comparative Law Studies, no. 5, pp. 136–148 (in Chinese)

5.Gao Q. (2020) A primer on intelligent revolution and modernization of national governance. *Zhong guo she hui ke xue*=Chinese Social Science, no.7, pp. 81–102 (in Chinese)

6. Han Y., Zhang F., Peng J. (2023) Order reconstruction: global economic governance under the impact of artificial intelligence. *Shi jie jing ji yu zheng zhi*=World Economy and Politics, no.1, pp.121–149 (in Chinese)

7. Jiang K. (2019) Law as algorithm. *Qing hua fa xue*=Tsinghua Law, no.1, pp. 64–75 (in Chinese)

8. Li C. (2021) Legal governance of artificial intelligence discrimination. *Zhong guo fa xue*=Chinese Law, no. 2, pp.127–147 (in Chinese)

9. Liu X. (2003) Dilemmas and directions of cognitive science research programs. *Zhong guo she hui ke xue*=Chinese Social Sciences, no. 1, pp. 99–108 (in Chinese)

10. Ma C. (2018) Social risks of artificial intelligence and its legal governance. *Fa lv ke xue (xi bei zheng fa da xue xue bao)*=Legal Science. Journal of Northwest University of Politics and Law, no. 6, pp. 47–55 (in Chinese)

11. Mei L. (2023) Technology displacing power: changing power structure of national governance in the age of artificial intelligence. *Wu han da xue xue bao (zhe xue she hui ke xue ban)*=Journal of Wuhan University. Philosophy and Social Science Edition, no. 1, pp. 44–54 (in Chinese)

12. Sun W. (2017) Rethinking value on artificial intelligence. *Zhe xue yan jiu*=Philosophical Research, no. 10, pp. 120–126 (in Chinese)

13. Shen X., Shi B. (2018) *The future computed: artificial intelligence and its role in society*. Beijing: Beijing University Press, 275 p. (in Chinese)

14. Wu S., Luo J. (2018) Legal governance of artificial intelligence safety: a review around system safety. *Xin jiang shi fan da xue xue bao(zhe xue she hui ke xue ban)*=Journal of Xinjiang Normal University. Philosophy and Social Science Edition, no. 4, pp. 109–117 (in Chinese)

15. Xu K. (2022) China's construction and theoretical reflection on the system of accounting laws. *Fa lv ke xue(xi bei zheng fa da xue xue bao)* =Legal Science. Journal of Northwestern University of Politics and Law, no. 1, pp. 124–132 (in Chinese)

16. Yu N. (2017) Self-consciousness and object-consciousness: the class nature of artificial intelligence. *Xue shu jie*=Academia, no. 9, pp. 93–101 (in Chinese)

17. Zhang X. (2023) Data risks and governance paths of generative artificial intelligence. *Fa lv ke xue(xi bei zheng fa da xue xue bao)* =Legal Science. Journal of Northwest University of Politics and Law, no. 5, pp. 42–54 (in Chinese)

18. Zhang A., Sun Y. (2021) The subjective perspective of algorithmic power and its state capacity shaping. *Xue shu yue kan*=Academic Monthly, no. 12, pp. 96–105 (in Chinese)

19. Zhang W. (2021) Building legal order of intelligent society. *Xin hua wen zhai*=Xinhua Digest, no. 3, pp. 18–23 (in Chinese)

20. Zhang L. (2023) Legal positioning and hierarchical governance of generative artificial intelligence. *Xian dai fa xue*=Modern Law, no. 4, pp. 126–141 (in Chinese)

21. Zhu M., Xu C. (2023) International soft law regulation of artificial intelligence ethics: current situation, challenges and countermeasures. *Zhong guo ke xue yuan yuan kan*=Bulletin of Chinese Academy of Sciences, no. 7, pp. 1037–1049 (in Chinese)

Information about the author:

Jia Shaoxue — Doctor of Sciences (Law), Associate Professor.

Analysing Risk-Based Approach in the Draft EU Artificial Intelligence Act

Dmitryi Leonidovich Kuteynikov[1], Osman Alikovich Izhaev[2]

[1, 2] Social and Legal Approaches to Applying Artificial Intelligence and Robotics Laboratory, State and Law Institute, Tyumen State University, 6 Volodarskogo Str., Tyumen, Russia, 625003

[1] kuteynikov@me.com; https://orcid.org/0000-0003-1448-3085 https://elibrary.ru/ author_profile.asp?id=776358, izhaev.osman@gmail.com, https://orcid.org/0000-0003-3777-8927, https://elibrary.ru/author_profile.asp?id=827391 SPIN-код: 3197-8857, AuthorID: 827391

[2] izhaev.osman@gmail.com, https://orcid.org/0000-0003-3777-8927, https://elibrary.ru/author_profile.asp?id=827391

Abstract

The article delves into the risk-based approach underpinning the draft EU Artificial Intelligence Act. Anticipated to be approved by the end of 2023, this regulation is poised to serve as a cornerstone in the European Union's legal framework for governing the development and deployment of artificial intelligence systems (AI systems). However, the ever-evolving technological landscape continues to present novel challenges to legislators, necessitating ongoing solutions that will span years to come. Moreover, the widespread proliferation of foundation models and general purpose AI systems over the past year underscores the need to refine the initial risk-based approach concept. The study comprehensively examines the inherent issues within the risk-based approach, including the delineation of AI system categories, their classification according to the degree of risk to human rights, and the establishment of optimal legal requirements for each subset of these systems. The research concludes that the construction of a more adaptable normative legal framework mandates differentiation of requirements based on risk levels, as well as across all stages of an AI system's lifecycle and levels of autonomy. The paper also delves into the challenges associated with extending the risk-oriented approach to encompass foundation models and general purpose AI systems, offering distinct analyses for each.

Keywords

artificial intelligence; AI systems; large language models; generative AI systems; foundation models; general purpose AI systems; Draft Artificial Intelligence Act; risk-based approach; conformity assessment procedure; audit of AI systems.

Introduction

The draft EU Artificial Intelligence Act[1] (AIA Draft) is a comprehensive act intended to regulate interactions in most of the areas related to the development and application of AI systems [Veale M. et. al., 2021: 112]. The EU initiated its development in 2018 involving a wide range of experts and the business community. As part of this work, a number of conceptual papers were presented that gradually formalised the key principles on which the future act was based. [2] The first text of the Draft was published in April 2021. In June 2023, the European Parliament approved the document with its amendments. This was followed by the trilogue stage, which involves agreeing on a unified text of the document on the basis of the positions worked out by the agencies. According to Euro MPs, the Draft will be ap-

---

[1] 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. Available at: URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206 (accessed: 30.08.2023)

[2] The most important of them are: Ethics guidelines for trustworthy AI. Available at URL: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai; Policy and investment recommendations for trustworthy Artificial Intelligence. Available at URL: https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence; High-Level Expert Group on AI: Final assessment list on trustworthy AI (ALTAI). Available at URL: https://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment; White Paper On Artificial Intelligence — A European approach to excellence and trust. Available at URL: https://commission.europa.eu/system/files/2020-02/commission-white-paper-artificial-intelligence-feb2020_en.pdf (accessed: 08.10.2023).

proved by the end of 2023. According to the latest version of the text, the Draft will be in force twenty four months after its approval.

The AIA Draft is risk based, that involves differentiating the requirements for bringing AI systems to market depending on their potential risk to human rights. In one form or another, this approach is the basis of regulatory concepts in many countries, including the USA[3], China[4], and Russia.[5] However, it is in the EU that it is closest to legislative implementation. Legislators in other countries and regions are either closely studying the European experience or directly declare their desire to adopt it [Gstrein O., 2022: 755].

The broad substantive and extraterritorial scope and the depth of detail make the Draft an extremely important document on a global scale, with the potential to have a major impact on the regulation across many countries [Greenleaf G., 2021: 9]. This trend has previously characterised other acts of the European Union and has been referred to in the academia as the Brussels Effect[6] [Balford A., 2012: 19].

It is also worth noting that technology companies planning to place their AI products on the EU market are looking for a policy on the development

---

[3] See ideas on different groups of legal requirements for AI systems depending on the potential risk of their application, which are contained in the most important documents characterising the US approach: Blueprint for an AI Bill of Rights. Available at: URL: https://www.whitehouse.gov/ostp/ai-bill-of-rights/ (accessed: 08.10.2023) и NIST AI Risk Management Framework. Available at: URL: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf (accessed: 08.10.2023). The need for a risk-based approach has also been repeatedly expressed at US Congressional hearings on new legislative initiatives. The same approach is also reflected in the bill introduced in September by Senators R. Blumenthal and J. Hawley's Bipartisan Framework for U.S. AI Act. Available at: URL: https://www.blumenthal.senate.gov/imo/media/doc/09072023bipartisanaiframework.pdf (accessed: 08.10.2023)

[4] See Artificial Intelligent White Paper 2022 describing China's regulatory approach and including a provision combining a risk-based approach with the level of autonomy (the proposal is to establish three groups of AI systems according to their level of autonomy and three groups according to the risk of their use in relation to human rights). Available at: URL: https://cset.georgetown.edu/publication/artificial-intelligence-white-paper-2022/ (accessed: 08.10.2023)

[5] The Concept for the Development of Regulation in Artificial Intelligence and Robotics Technologies until 2024 explicitly states that it is premised on a risk-based and human-centred approach. The Code of Ethics for Artificial Intelligence contains similar provisions. Available at: URL: http://government.ru/docs/all/129505/ (accessed: 08.10.2023)

[6] The Brussels effect refers to the unilateral influence of acts and standards adopted at the EU level on the legal systems of other countries. A similar phenomenon has previously been observed, e.g., in laws on data circulation, antitrust regulation, environmental protection and food safety.

and use of AI systems that will take into account most of the provisions of the Draft to facilitate future compliance. Moreover, developers are already partly taking these requirements into account. For example, a recent study by a group of scholars from the Stanford Institute for Human-Centered Artificial Intelligence (HAI) evaluated, using twelve criteria, how well the most advanced foundation models currently meet the requirements of the Draft. The authors of the study concluded that the degree of compliance with the act varies widely from 25% to 75%. However, meeting all or most of the legal requirements is quite feasible, which will help to improve the quality of functioning and product safety[7].

Thus, in view of the fact that the AIA Draft is the most comprehensive initiative to date, a study of its approaches is essential for balanced regulation, including regulation in the Russian Federation, because Russia, like most other states, has not yet moved from the stage of approving concepts to the development and adoption of laws and regulations. At the same time, the key principles underlying the Concept for the Development of Regulation in Artificial Intelligence and Robotics Technologies until 2024 approved by the Decree of the Government of Russia[8], are, for the most part, similar to those contained in the Draft mentioned. Also, technology businesses planning to participate in the international market in the future should understand the development of global regulatory trends.

The authors of the paper aimed to explore the risk-based approach contained in the Draft, identify the main regulatory legal requirements imposed on entities placing AI systems on the market, and analyse the key challenges facing the legislator at this stage. The results of this research can be used by government agencies in the development of concepts and regulations, as well as by businesses in preparing to meet the requirements for placing AI systems on the markets.

A series of general and specific scholarly methods were applied in the course of the work. The analysis method was used to divide the Draft and other statutory acts into separate parts, which allowed for a detailed examination of their structure and internal elements. The synthesis method was used to combine the internal elements of the reviewed documents into single semantic blocks, which contributed to obtaining comprehensive knowledge

---

[7] See: Do Foundation Model Providers Comply with the Draft EU AI Act? Available at: URL: https://crfm.stanford.edu/2023/06/15/eu-ai-act.html (accessed: 08.10.2023)

[8] Available at: URL: http://government.ru/docs/all/129505/ (accessed: 08.10.2023)

about the subject matter under study. The induction and deduction methods helped to identify common features and differences characteristic of the way the risk-based approach is applied in various countries and regions. The systematic approach helped to systematise and structure the knowledge about the subject matter under study. The formal legal method was used to study the provisions of individual legislative acts, which helped to determine the features of legal regulation of public relations in the area under consideration. The comparative legal method was used to identify the advantages and disadvantages of the risk-based approach stipulated in the Draft.

The Risk-Based Approach in the Draft: Features and Key Challenges

1. The Concept of AI Systems and Their Classification by Risk Levels

1.1. AI Systems Definition in the Draft

To begin consideration of the AI systems and the way they are classified by the risk level, we have studied their definition given in the Draft. This is essential for understanding what particular products potentially fall within its scope. The latest version of the document[9] offers the following definition: "Artificial intelligence system (AI system) means a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments."[10]

---

[9] All the three versions of the Draft contain definitions of the term 'AI system' that slightly differ from each other. The European Commission text (2021): "...software that is developed with one or more of the techniques and approaches (these approaches are listed separately in an annex to the document) 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." The EU text: "...a system that is designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge based approaches, and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts."

[10] 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. Available at: URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206 (accessed: 30.08.2023)

As this definition is quite broad, it allows including into AI systems a large number of software products developed on the basis of various methods and techniques, and not only those based on neural networks or machine learning techniques. Technology neutrality is another important feature. AI systems are defined through essential attributes that are inherent to them rather than by listing relevant technologies and methods. It should also be noted that the definition under review was an intentional move by European legislators towards terminology unification at the international level. For example, the Recommendations of the Council on Artificial Intelligence of the Organization for Economic Co-operation and Development (OECD) contain a similar definition.[11] Currently, this version is the most widespread and has become the basis for regulatory concepts in many OECD countries (including such leaders in the field of AI technologies as the USA[12]).

The approach to AI systems definition that aims to identify their main attributes is the most flexible of all and is justified for a legislative document. The attributes in question include: tasks performed, human role in tasking, operating environment, autonomy, and self-learning. More concrete recommendations on AI systems classification that are not technologically neutral may be in the future included in technical standards and in enactments issued by executive authorities [Schuett J., 2023: 3].

At the same time there is a variety of AI systems that can be used in completely different scenarios, from recommendation generation and content creation to critical infrastructure management and national security. Consequently, a specific set of means and methods of legal impact should be applied to different groups of such systems.

1.2. Classification of AI Systems by Risk Levels

The Draft under review uses a risk-based approach to classify AI systems into groups. The higher the risk of human rights violations from the

---

[11] Note: the OECD document contains the following definition: "An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy." AI system: An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy. Available at: URL: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (accessed: 30.08.2023)

[12] NIST AI Risk Management Framework (AI RMF 1.0). Available at: URL: https://www.nist.gov/itl/ai-risk-management-framework (accessed: 30.08.2023)

use of individual AI systems, the more stringent the requirements placed on them. The Draft provides for a total of four such groups: prohibited AI systems, high-risk AI ones, limited-risk AI ones, and low-risk AI ones. Each group of AI systems has its individual legal requirements.

The Draft applies to entities operating AI systems in the EU. 'Providers' who deploy such systems in the EU market are among such entities, and it does not matter where they are domiciled or actually located. The decisive factor is whether the results of these system operation are intended for use within the EU. Even if the provider is in a third country but uses output data in the EU, it will fall under provisions of the Draft. The document then uses the term "deployer" of an AI system; however, what it means is not the end user but entities using an AI system at other levels (downstream usage). This is supported by the provision that deployers are individuals who do not use such systems for personal (non-professional) purposes. In addition, the original version of the document used the term 'user', and the current version uses the term 'deployer.' In this way, lawmakers sought to stress that they meant specifically entities using AI systems in their products. The Draft also applies to importers, distributors, authorised representatives of providers and manufacturers of products. Such entities — unlike providers and users — must be located or registered in the EU.

One disadvantage of the risk-based approach is its inflexibility: as technology evolves, the classification of AI systems will have to be revised frequently.[13] Experts suggest that this problem could be somewhat mitigated, in particular, by using a more flexible approach to categorising AI systems into groups based on the risk. Their risk assessment system consists of two steps: the development of risk scenarios and the application of a proportionality test. Such an approach may improve the application of the Draft AIA [Novelli C. et. al., 2023: 4–5].

At the same time, it is possible that dividing regulatory requirements for AI systems only into risk groups will not address all of the challenges facing lawmakers. For example, the text of the Draft proposed by the European Commission did not allow regulating the market entry of foundation models[14] and general purpose AI systems[15] whose wide-scale use began in a large

---

[13] The regulation of artificial intelligence. Available at: URL: https://link.springer.com/article/10.1007/s00146-023-01650-z (accessed: 08.10.2023)

[14] The Draft gives the following definition for the foundation model: 'foundation model' means an AI system model that is trained on broad data at scale, is designed for generality of output, and can be adapted to a wide range of distinctive tasks.

[15] The Draft AIA gives the following definition for the general purpose AI system:

number of fields only at the end of 2022. A big part of the problem is that the Draft focuses on establishing responsibilities for various entities that are going to place systems on the market. At the same time, other distribution channels are usually typical of foundation models. For example, the most powerful and popular channels are privately owned. Companies provide access to their use and customisation for commercial purposes for a fee through software interfaces (APIs). That means some companies build and deploy these systems, while others apply them to solve a wide range of tasks. However, the latter group do not have access to the full source code of the model, the training data, or the infrastructure (sometimes this can be third-party cloud computing power); nor can they improve or adjust the model. Hence, it is not possible to use an approach that focuses all attention only on the actors that actually place AI systems on the market. Thus, it is necessary to establish regulatory requirements for all stages of the life cycle of AI systems, such as development, deployment, and application.

The present level of foundation models opens up a broad range of opportunities for the creation of autonomous agents on their basis in the coming years, and such agents would be capable of undertaking individual activities, including legally significant ones, on behalf of a human. So, it has a sense to look at the level of AI system autonomy as one of the areas that requires legal regulation.

Thus, to work out a more flexible regulatory approach, we need to differentiate requirements both by risk levels and by all stages of the life cycle of AI systems and the degree of their autonomy. This classification will make AI systems more flexible, that will allow to apply a wider range of legislative requirements. For example, it will become mandatory to test some systems mentioned and foundation models in regulatory sandboxes before placing them on market; some such systems will have to undergo external independent audits; others will have to undergo internal compliance assessments. Additionally, the law may require that to place some AI systems on the market, internal ethical and corporate standards and risk management frameworks must be established.

2. Regulatory Requirements for Certain Groups of AI Systems

2.1. General Principles Applicable to All AI Systems

'general purpose AI system' means an AI system that can be used in and adapted to a wide range of applications for which it was not intentionally and specifically designed.

The Draft establishes a list of general principles to guide operators (providers etc.) at all stages of development and operation of AI systems and foundation models. These principles include: human agency and oversight; technical robustness and safety; privacy and data governance; transparency, diversity, non-discrimination and fairness; social and environmental well-being. These guidelines do not, however, directly impose additional legal obligations on operators. The meeting of the Draft specific requirements that relate to different AI system types and foundation models will for the most part serve as evidence of their compliance. These principles should then be incorporated into technical and corporate standards. Moreover, the Draft explicitly places the obligation to include them in technical standards on the European Commission and the future AI Office[16]. These documents will help to develop rather abstract principles into technical requirements.

2.2. Prohibited AI Systems

The risk-based approach stipulates a separate group of AI systems that, by virtue of their functional characteristics, pose an unacceptable risk to human rights and freedoms. For this reason, their use is illegal in the EU. The Draft identifies several groups of prohibited uses of AI systems.

It is prohibited to use these systems that (in a covert manner) manipulate a person's behaviour so that this results in material harm to her/him or another person. This prohibition will apply to AI systems, which simultaneously meet the following criteria: the system influences the person in question at the subliminal level or performs deliberate manipulation; the person makes an uninformed decision; the system causes substantial harm. The initial version of the Draft stipulated that this prohibition applies to all cases where physical or psychological harm is caused. This understanding was too narrow because AI systems can also cause social, cultural, financial and other harm. [Neuwirth R., 2023: 6–7].

The Draft also prohibits AI systems to make use of vulnerable human attributes (age, disability, etc.) resulting in behavioural change and substantial harm. In other words, it is illegal to use AI systems to classify individuals by using legally protected sensitive attributes.

Social scoring of individuals (groups of individuals) is placed in an independent group of prohibited practices. It is not permitted to assess a

---

[16] A new European Union body to be established under the current text of the Draft. The document defines its intended competence and structure.

person on the basis of their social behaviour or known or predicted personality characteristics. Such an assessment must result in discriminatory treatment of certain individuals (groups): (a) in a social context unrelated to the context in which data about them were originally generated or accumulated; or (b) that is unjustified or disproportionate to their social behaviour or its severity.

The list of prohibited scenarios for the use of AI systems also includes: use of remote real-time biometric identification systems in public places; use of predictive analytics to determine the likelihood of an individual committing an offence; creation of databases based on untargeted collection of facial images from the Internet or CCTV footage; use of emotion recognition software in law enforcement, border control, educational institutions, and at the place of work.

And, finally, video footage from publicly accessible locations may not be analysed using remote biometric identification systems unless such use is subject to judicial authorisation under EU law for the purposes of a search (of persons) related to a criminal offence.

From the point of view of applying above prohibitions, the provisions that do not allow the use of subliminal influence techniques are a challenge [Neuwirth R., 2023: 3]. It is clear that subliminal techniques can significantly influence decision making and lead to undesirable consequences for the individual. At the same time, the term "subliminal" is difficult to define, and the Draft gives no explanation of its meaning. AI systems can often influence human behaviour using both conscious and subliminal techniques at the same time. For example, smart glasses can influence the human psyche in an overt way by showing pictures, videos, playing music, and, at the same time, in a covert way, read the person's emotions through eye movement recognition, electrical activity in the brain, heartbeat and heart rhythms, muscle activity, blood density in the brain, blood pressure, and skin conductivity.

As a result, it would be difficult to establish whether subliminal techniques have been used, and that these techniques have caused a significant distortion of a person's behaviour. The Draft or other acts should clearly define the term "subliminal techniques" and clarify the legality of their use.

The issue of classifying certain systems as prohibited has been a matter of debate among political forces due to the difficulty in balancing human rights and the public interest. Not everyone who participated in the discussions

were satisfied with the results of the consensus reached after the text was approved by the European Parliament. In particular, human rights organisations asked the EU bodies to be more diligent in protecting human rights during the trilogue. For example, one of the proposals was to involve civil society actors in assessing the impact of AI systems on fundamental human rights, provide for the possibility to appeal decisions taken by AI systems, including through human rights defenders, and establish flexible compensation for victims. It was also proposed to introduce restrictions on the use of AI systems in law enforcement, migration control, and national security.

Thus, legislators should formulate clear criteria for classifying AI systems as prohibited. It will allow developers to better understand the permissible boundaries when creating products, on the one hand, and avoid arbitrary classification of systems as prohibited by law enforcement authorities, on the other.

2.3. High-Risk AI Systems

Title III of the Draft lists the requirements to high-risk AI systems. According to Article 6, AI systems listed in Annexes II and III belong high-risk AI systems, independently or as a component of the safety system of another product.

Annex II contains two lists of acts of the harmonised EU laws, those based on the New Legislative Framework, and others. The acts categorised under this Annex define products and areas of the economy in which the application of AI systems is associated with increased risk. Annex III sets out eight groups that categorise AI systems as high-risk systems by the areas of their application. These include, among others: biometric and biometrics-based systems; AI systems for the management and operation of critical digital infrastructure; AI systems for education and vocational training. Together, these Annexes are intended to provide an exhaustive list of high-risk AI systems by allowing for the inclusion of large areas of the economy as well as more specific usage scenarios.

The EC will develop updated requirements for categorising such systems after consultation with the AI Office at least six months before the Draft enters into force. Law-enforcement agencies in the EU have enough time to make final and balanced decisions so as not to impose excessive requirements and in this manner stifle entrepreneurial activity.

There is a new layer of regulation in the current version of the document that significantly reduces the list of systems, which can be categorised among high-risk systems. For instance, high-risk AI systems identified on the basis of the areas of their application (Annex III) will now only be recognised as such if they significantly threaten life, safety and fundamental human rights. AI systems for managing and operating critical digital infrastructure must additionally pose a significant risk of harm to the environment. Introducing this layer of requirements was a major step towards liberalising business requirements. This has significantly reduced the list of AI systems that will be classified as high-risk systems.

The Draft stipulates a number of requirements that must be met for high-risk AI systems to be placed on the market. A risk management system must be established and implemented, and then needs to be updated in a timely manner throughout the life cycle of the AI system; data sets (training, validation and testing data sets) for the AI systems that are based on such systems should be quality tested; all necessary documents about the system must be created and updated in a timely manner before the system is placed on the market; the system should be able to record all activities during its operation in a special logbook; the operation of the system should, as far as possible, be understandable and transparent to different levels of providers and end users; systems should be designed to be controllable by a human being; systems should be designed from the outset to meet the requirements of safety, reliability, accuracy, resilience and cybersecurity. Alongside the above provisions, additional requirements are placed on individual high-risk AI systems. For example, these must be registered in a single database and must undergo the fundamental rights impact assessment for high-risk AI systems.

Conformity assessment, as envisaged in the Draft, is an integral part of high-risk AI systems' safety and reliability. Providers of high-risk systems must undergo this procedure before releasing their product to the market. There are two types of conformity assessment procedures: (a) the conformity assessment procedure based on internal control referred to in Annex VI; (b) the conformity assessment procedure based on assessment of the quality management system and assessment of the technical documentation, with the involvement of a notified body[17], referred to in Annex VII.

---

[17] Notified body means a conformity assessment body notified in accordance with the Draft and other relevant EU harmonisation legislation.

This second type of procedure will be used in a relatively limited number of cases where either technical standards and common specifications developed by the European Commission are not applicable, or the supplier voluntarily decides to undergo an external conformity assessment regardless of the categorisation of the AI system under a particular risk level. A voluntary conformity assessment by a notified body can be a competitive advantage, as it will mean that the public agency has guaranteed product safety to consumers. Such an incentive will help improving the overall quality of AI systems without introducing additional stringent regulatory measures.

The Draft has been repeatedly criticised, and it has become the subject of scholarly discussions in the context of conformity assessment procedure. In particular, a group of experts noted that the Draft did not provide detailed explanations on how such an assessment should be undertaken [Mökander J. et al., 2022: 251]. The guidelines developed to date in academia can significantly help businesses overcome this shortcoming. Examples of such documents include: capAI — a guide to going through this procedure, which documents in detail all the measures that high-risk AI system providers need to take[18]: Guidelines for assessing the ethics and reliability of AI systems at different stages of their life cycle in determining the intended use, design, and development [Vetter D. et al., 2023: 5].

 Another point of debate is that effective verification of AI systems requires an external independent audit based on ethical principles and standards [Mökander J. et al., 2021: 21–22]. Scholars note that not only lawyers, engineers and philosophers, but also specialists in the field of management should be involved in the development of audit procedures. This conclusion was based on the experience of auditing AstraZeneca's AI systems for ethical compliance. The authors of the study showed that the main difficulties organisations face in auditing AI systems are related to usual management problems. They also touched upon questions of the audit structure. For instance, the authors proposed a 'three-layer' audit for large language models: audit of management, audit of the model, and audit of its application [Mökander J. et al., 2023: 5, 464].

Thus, classifying a small group of systems as high-risk AI systems is a positive measure aimed at creating favourable conditions for business and innovation development. The same applies to the conformity assessment procedure,

---

[18] CapAI — A Procedure for Conducting Conformity Assessment of AI Systems in Line with the EU Artificial Intelligence Act. Available at: URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4064091 (08.10.2023)

which in the vast majority of cases will be conducted on the basis of internal control. It seems, however, that some of the most powerful AI systems and foundation models may eventually require more stringent requirements, such as external independent auditing and licensing, to place on the market.

2.4. Limited-Risk AI Systems and Low-Risk AI Systems

This group of AI systems should meet additional requirements for operational transparency (Title IV). For example, providers should ensure that all necessary measures are in place to make it clear to users that they are interacting with AI systems. They should also provide information on the permissible functions of the AI system, human control over it, the entity making the final decisions, and the procedures for challenging these decisions in accordance with the law. Providers of authorised systems that recognise human emotion should seek consent to process biometric information of the individuals in question. It is also stipulated that 'deepfakes' must be labelled — unless the content is obviously generated for artistic, humorous or other purposes.

The main idea behind these provisions is that individuals should be informed about their interactions with AI systems. For example, they need to know that their emotions or other characteristics are being recognised, or that image, video or audio content is being generated. This will increase public confidence in AI systems [Chamberlain J., 2023: 5].

Title IX of the Draft stipulates that developers of such AI systems are encouraged to elaborate voluntary Codes of Conduct that reflect how the principles envisaged for all the AI systems discussed earlier are to be implemented. Then it will be clear to users how to operate the system correctly and what measures the developers have taken to make the products safe.

At the same time, researches show that the perception of AI systems and the effect of their application depends very much on what information the user has about them [Pataranutaporn P. et al., 2023: 3]. It is quite easy to mislead people and lower their alertness through proper advertising and overly positive product descriptions. Thus, there is a need to demand that companies develop adequate and understandable rules for the use of the AI system that contain notifications of possible negative consequences. The same should apply to the interfaces that users interact with.

3. The Risk-Based Approach in the Context of Foundation Models and General Purpose AI Systems

The key issue in the finalisation of the Draft is the choice of regulatory approach to the development and application of foundation models and general purpose AI systems. The three versions of its text contain different provisions: only general requirements that apply to all AI systems by risk level and no additional requirements (European Commission text); additional requirements are established for general purpose AI systems (European Council text); individual requirements are established for foundation models, while general purpose AI systems are subject to general requirements on risk levels (European Parliament text). All of the approaches have a number of debatable and ambiguous controversial provisions. Considering the high relevance of the content of the Draft for political forces, business, and the public, it is still difficult to predict unequivocally whether any of the approaches considered will be chosen as the main one, or whether the final text will to some extent combine all of them. Moreover, in some cases, finding the most balanced solution is complicated by the lobbying of large technology companies[19] that have the power to influence the process of drafting and discussing regulations.

Scholars have also taken other positions on the place of general purpose AI systems and foundation models in a risk-based approach. For instance, researchers at The Future Society[20] suggest that all general purpose AI systems should be categorised into three broad groups based on the levels of risk they pose to human rights: Generative AI systems (400+ providers); Group 1 general purpose AI systems (foundation models) (~14 providers); Group 2 general purpose AI systems (frontier foundation models) (~10 providers). Each group will have a different set of legal requirements. Group 3 will be characterised by the most extensive regulatory requirements, which include, in addition to the requirements for all other groups, requirements such as: internal and external independent audits, regular interaction with the AI Office, full transparency, etc. At the same time, this approach is clearly weak as it offers a division into too few groups and is too reliant on current technological realities.

Thus, we believe it is necessary to divide the requirements for foundation models and general purpose AI systems into different groups. Specific requirements should be applied to foundation models, taking into account

---

[19] See: The lobbying ghost in the machine. Big Tech's covert defanging of Europe's AI Act. Available at: URL: https://corporateeurope.org/sites/default/files/2023-03/The%20 Lobbying%20Ghost%20in%20the%20Machine.pdf (accessed: 08.10.2023)

[20] Heavy is the Head that Wears the Crown. Available at: URL: https://thefuturesociety. org/heavy-is-the-head-that-wears-the-crown/ (accessed: 08.10.2023)

that different actors will distribute them at all stages of their life cycle. Requirements for general purpose AI systems should vary based on a risk-based approach. Placing the frontier general purpose AI systems on the market must be based on more extensive regulatory requirements.

The latest version of the Draft retains the term "general purpose AI systems", but extends the requirements to the development and application of the foundation models and the entire AI value chain. The new Article 28b established a number of requirements for AI systems that they must meet before they can enter the market. These include: take measures to mitigate possible negative consequences from their application, use pre-trained and validated data sets, develop only models that can be safe, transparent and predictable throughout their lifecycle, keep relevant technical documents about the model for at least 10 years from the date of its release to the market, etc. Generative AI systems must meet additional requirements: comply with transparency requirements, build and train models in such a way that they cannot potentially be used for infringing purposes, and disclose details of the use of copyrighted material in datasets. All these measures are designed to place additional obligations on the developers of AI systems and thereby offset the shortcomings of the risk-based approach that involves only setting requirements for entities bringing AI systems to market.

The requirement to disclose datasets causes the greatest controversies. This issue is extremely painful because its regulation requires a balance between support for content creators and technology development [Hacker P., 2021: 259]. At the same time, the Draft stipulated long lead times for the preparation of datasets by technology companies when these create new products. Some companies already voluntarily use only legally clean data to create their products nowadays[21].

Another important measure that is widely discussed in academia and society is the right of an individual to prohibit the use of their data or their property to train AI systems. Requirements in this regard have not yet been reflected in the Draft, but some people in the business community have expressed their willingness to offer such waivers[22].

---

[21] Adobe's Firefly has been fully trained on legally clean data (on its Adobe Stock dataset and on open licence works and public domain content whose copyrights have expired). Also, the company has a whole team of moderators who check new data for copyright infringement risks before adding it to datasets.

[22] For example, StabilityAI voluntarily accepts applications from authors demanding that their content be removed from datasets. OpenAI has announced that it will not collect

Another pressing issue is access to AI systems and foundation models. The current text only allows to test them in regulatory sandboxes. Meanwhile, legislating an obligation to leave open access to AI systems and foundation models for scholars and researchers would be a rational measure. This would ensure the necessary level of transparency in the functioning of such systems because independent experts could monitor the quality of AI systems and identify potential threats in a timely manner.

A number of issues regarding the distribution of foundation models and AI systems under open licences also remain unspecified. In particular, a group of companies that distribute advisory software have suggested that lawmakers should provide a clear definition of AI components. The latest version of the text of the Draft (European Parliament version) contains such a term regarding open-source (Articles 5e и 12a-c), but does not give it an exhaustive definition.[23] Another rational solution in helping small businesses may be to differentiate requirements for foundation models suppliers depending on their use cases, development methods, and market position. Scholars suggest using, e.g., a staggered system for bringing foundation models to market. It implies that hazard levels of the system should be defined to grant access to the system under open licences [Solaiman I., 2023: 119]. This means that, e.g., foundation models with market-leading features will be prohibited for distribution via open-source due to high risks of leakage and misuse.[24]

Conclusion

Although the Draft has been actively developed and discussed for several years, there are still a number of issues that have not been clearly resolved. Moreover, the constant changes in technology create new problems

---

data labelled "Do Not Train". A whole range of US companies that are part of the Content Authenticity Initiative have developed and are implementing Content Credentials. The technology allows for the addition of a "Do Not Train" tag to metadata, which should allow the data not to be included in future datasets, digitally tag the data for authorship, and separate generated content from copyrighted content (in order to protect human-created elements with copyright).

[23] Supporting Open Source and Open Science in the EU AI Act. Available at: URL: https://huggingface.co/blog/assets/eu_ai_act_oss/supporting_OS_in_the_AIAct.pdf (accessed: 08.10.2023)

[24] Open-Sourcing Highly Capable Foundation Models. Available at: URL: https://www.governance.ai/research-paper/open-sourcing-highly-capable-foundation-models (accessed: 08.10.2023)

and challenges for lawmakers. In addition, the extensive adoption of foundation models over the past year requires refinement of the original concept of the risk-based approach.

In order to build a flexible regulatory approach, requirements need to be differentiated both by risk levels and across all stages of the life cycle of AI systems and their degree of autonomy. This will allow a wider range of legislative requirements to apply to different groups of systems. This approach also makes it possible to take into account the distribution of these systems and foundation models by different actors and to properly regulate all stages of their life cycle.

The provisions related to the classification of such systems by risk levels need to be refined. First, the range of prohibited systems should be clearly defined on the basis of clear criteria. It will help developers to better understand the regulatory requirements, and to eliminate arbitrary practices in the decisions taken by law enforcement agencies. Second, classifying a small group of systems as high-risk systems may have a positive impact on innovation and technology development. However, some of the most capable systems and foundation models may eventually need more stringent requirements, such as external independent auditing and licensing, to be placed on the market. Third, legal requirements are needed to develop adequate and understandable rules for the use of systems and their interfaces, which should notify the user of possible negative consequences.

An analysis of the requirements for placing foundation models on the market has shown that the existing approach can be improved by implementing a number of additional regulatory requirements. First, regulatory requirements for foundation models should take into account their distribution by different actors at all stages of their life cycle, and requirements for general purpose AI systems should take into account their risk level. Second, users should be able to unilaterally opt out of having their data used to train these systems. Third, researchers should be given access to the systems and foundation models to ensure their security. Fourth, additional requirements for placing AI systems on the market under open licences should be provided.

References

1. Bradford A. (2012) The Brussels Effect. *Northwestern University Law Review,* vol. 107, no. 1, pp. 1–64.

2. Chamberlain J. (2023) The Risk-Based Approach of the European Union's Proposed Artificial Intelligence Regulation: Some Comments from a Tort Law Perspective. *European Journal of Risk Regulation,* vol. 14, no. 1, pp. 1–13.

3. Gstrein O. (2022) European AI Regulation: Brussels Effect versus Human Dignity? *Zeitschrift für Europarechtliche Studien,* vol. 4, pp. 755–772.

4. Greenleaf G. (2021) The "Brussels Effect" of the EU's "AI Act" on Data Privacy Outside Europe. *Privacy Laws & Business International Report,* issue 171, pp. 3–7.

5. Hacker P. (2021) A legal framework for AI training data—from first principles to the Artificial Intelligence Act. *Law, Innovation and Technology,* vol. 13, no. 2, pp. 257–301.

6. Mahler T. (2021) Between risk management and proportionality: The risk-based approach in the EU's Artificial Intelligence Act Proposal. In: Publicerad i Nordic Yearbook of Law and Informatics 2020–2021: Law in the Era of Artificial Intelligence, Mars, pp. 247–270.

7. Mökander J. et al. (2023) Operationalising AI governance through ethics-based auditing: an industry case study. *AI and Ethics,* vol. 3, issue 2, pp. 451–468.

8. Mökander J. et al. (2022) Conformity Assessments and Post-market Monitoring: A Guide to the Role of Auditing in the Proposed European AI Regulation. *Minds & Machines,* vol. 32, issue 2, pp. 241–268.

9. Mökander J. et al. (2021) Ethics-Based Auditing of Automated Decision-Making Systems: Nature, Scope, and Limitations. *Science and Engineering Ethics,* vol. 27, issue 4, pp. 1–30.

10. Mökander J. et al. (2023) Auditing large language models: a three-layered approach. Available at: https://doi.org/10.1007/s43681-023-00289-2

11. Neuwirth R. (2023) The EU Artificial Intelligence Act: Regulating Subliminal AI Systems. L.: Routledge, 144 p.

12. Neuwirth R. (2023) Prohibited artificial intelligence practices in the proposed EU Artificial Intelligence Act (AIA). *Computer Law & Security Review,* vol. 48, pp. 1–41.

13. Novelli C. et. al. (2023) Taking AI risks seriously: a new assessment model for the AI Act. *AI & Society*, vol. 38, no. 3, pp. 1–5.

14. Pataranutaporn P. et. al. (2023) Influencing human–AI interaction by priming beliefs about AI can increase perceived trustworthiness, empathy and effectiveness. Nat Mach Intell. Available at: https://doi.org/10.1038/s42256-023-00720-7.

15. Schuett J. (2023) Risk Management in the Artificial Intelligence Act. *European Journal of Risk Regulation*, February*, pp. 1–19.

16. Solaiman I. (2023) The Gradient of Generative AI Release: Methods and Considerations. In: Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency. N.Y.: Association for Computing Machinery, p. 111–122.

17. Veale M. et. al. (2021) Demystifying the Draft EU Artificial Intelligence Act. *Computer Law Review International*, vol. 22, issue 4, pp. 97–112.

18. Vetter D. et. al. (2023) Lessons Learned from Assessing Trustworthy AI in Practice. *Digital Society*, vol. 2, issue 3, pp. 1–25.

Information about the authors:

D.L.Kuteinikov — Candidate of Sciences (Law).
O.A.Izhaev — Candidate of Sciences (Law), Senior Researcher.

Legal Issues in the DIGITAL AGE

The submitted articles should be original, not published before in other printed editions. The articles should be topical, contain novelty, have conclusions on research and follow the guidelines given below. If an article has an inappropriate layout, it is returned to the article for fine-tuning. Articles are submitted Word-processed to the address: lawjournal@hse.ru

Article Length

Articles should be between 60,000 and 80,000 characters. The size of reviews and the reviews of foreign legislation should not exceed 20,000 characters.

The text should be in Times New Roman 14 pt, 11 pt for footnotes, 1.5 spaced; numbering of footnotes is consecutive.

Article Title

The title should be concise and informative.

Author Details

The details about the authors include:

· Full name of each author

· Complete name of the organization — affiliation of each author and the complete postal address

· Position, rank, academic degree of each author

· E-mail address of each author

Abstract

The abstract of the size from 150 to 200 words is to be consistent (follow the logic to describe the results of the research), reflect the key features of the article (subject matter, aim, methods and conclusions).

The information contained in the title should not be duplicated in the abstract. Historical references unless they represent the body of the paper as well as the description of the works published before and the facts of common knowledge are not included into the abstract.

Keywords

Please provide keywords from 6 to 10 units. The keywords or phrases are separated with semicolons.

References

The references are arranged as follows: [Smith J., 2015: 65]. See for details http://law-journal.hse.ru.

A reference list should be attached to the article.

Footnotes

The footnotes include legal and jurisprudencial acts and are to be given paginaly.

The articles are peer-reviewed. The authors may study the content of the reviews. If the review is negative, the author is provided with a motivated rejection.

Вопросы права **В ЦИФРОВУЮ ЭПОХУ**

«Вопросы права в цифровую эпоху» — научный ежеквартальный электронный журнал, направленный на всесторонний анализ права в цифровую эпоху. Его главная цель заключается в рассмотрении вопросов, связанных с правовыми последствиями постоянно меняющихся информационных технологий.

Цифровая эпоха — это эпоха информационных и коммуникационных технологий, обусловливающих дальнейшее общественное развитие, в том числе с использованием цифровых данных. Но вместе с тем цифровое развитие выявляет пробелы в праве и потребность в новых правовых решениях.

"Вопросы права в цифровую эпоху" — журнал, который предоставляет возможность юристам — ученым и практикам — обмениваться мнениями. В том числе журнал поощряет междисциплинарные дискуссии по темам, находящимся на стыке права, технологий, экономики и политики в современном мире.

"Вопросы права в цифровую эпоху" — рецензируемый журнал. В нем применяется двойное "слепое" рецензирование присылаемых материалов.

Журнал приглашает авторов присылать статьи, отражающие результаты научных исследований регулирования цифровой среды. Редакция приветствует теоретические и компаративистские подходы, исследование перспектив правового развития в различных странах.

Legal Issues in the **DIGITAL AGE**

Вопросы права
в цифровую эпоху

3/2023

Статьи

СТАТЬИ

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ, ОРИЕНТИРОВАННЫЙ СПЕЦИАЛИЗИРОВАННЫМ ЮРИДИЧЕСКИМ ЯЗЫКОМ

Юрий Михайлович Батурин
Высшая школа государственного аудита (факультет), Московский государственный университет им. М.В. Ломоносова,
Российская Федерация, 119992, Москва, Ленинские горы,
д. 1, стр. 13, 4-й учебный корпус гуманитарных факультетов
baturin@ihst.ru, SPIN-код: 9179-9362, ORCID: 0000-0003-1481-5369

Аннотация

Исходя из структурного подобия законодательства и компьютерных программ, предлагается для нормативного регулирования поведения искусственного интеллекта сделать юриста обязательным участником проектирования и разработки систем искусственного интеллекта. Выдвигается идея разработки с этой целью объектного специализированного юридического языка. Обсуждаются базовые элементы такого языка. Подчеркивается, что искусственный интеллект должен уметь самостоятельно формулировать и описывать свои цели на том же объектном языке, что необходимо для обратной связи с его создателями и пользователями. На примере принятых в Российской Федерации нормативных документов и ГОСТов о беспилотных транспортных средствах демонстрируется сложность задачи создания специализированного юридического языка искусственного интеллекта, в том числе потому что для формализации юридических суждений требуется их контекстная градация. Прогнозируется появление семейства объектных юридических языков разработчиков искусственного интеллекта. Ставится вопрос о создании теории искусственного интеллекта, которая должна объяснить данные и факты, которыми будет оперировать сильный искусственный интеллект. Рекомендуется скорректировать определение искусственного интеллекта, уже воспринятое утвержденными концепциями и стратегиями, как системы, осуществляющей поиск решений без заранее заданного алгоритма, так чтобы оно не исключало использования алгоритмов. Показывается важная роль алгоритмов для искусственного интеллекта. Сильный искусственный интеллект интерпретируется как ориентируемый объектным языком. Анализируются различия между сильным искусственным интеллектом и человеком. Обсуждается внутреннее представление о мире и о себе способного к ответственному поведению искусственного интеллекта, в терминах которого выражалась бы входная для него информация. Делается вывод, что концептуальные, лингвистические

и практические трудности, с которыми столкнутся юристы-разработчики сильного искусственного интеллекта, не должны останавливать «юридизацию проектирования» искусственного интеллекта.

Информация об авторе:

Ю.М. Батурин — доктор юридических наук, профессор, член-корреспондент РАН, зав. кафедрой компьютерного права и информационной безопасности

Научная статья

«ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ» КАК ОБЪЕКТ ИМУЩЕСТВЕННОГО ОБОРОТА: ОСОБЕННОСТИ ГРАЖДАНСКО-ПРАВОВОГО РЕЖИМА

Людмила Юрьевна Василевская
Университета имени О.Е. Кутафина (МГЮА), Россия 125993, Москва, ул. Садовая-Кудринская, д. 9, ljvasilevskaja@msal.ru

Аннотация

С опорой на методологию цивилистического исследования анализируются понятие «искусственного интеллекта», его правовая природа и особенности включенности в имущественный оборот. В качестве предмета исследования выступают доктрина по рассматриваемой проблематике, нормы законодательства об «искусственном интеллекте» и практика их применения. Цель исследования состоит в теоретическом обосновании научного представления об «искусственном интеллекте» как сложного объекта гражданских прав, в структуре которого могут быть представлены различные результаты интеллектуальной деятельности. Правовая природа «искусственного интеллекта» как идеального явления предопределяет особенности его гражданско-правового режима — специфику правового оформления отношений с этим объектом. В статье использованы системно-структурный и сравнительно-правовой методы познания. В частности, использование системно-структурного метода позволило рассмотреть «искусственный интеллект» не только как сложно-структурное образование, но и выделить в

нем основной структурообразующий элемент — компьютерные программы. Использование сравнительно-правового метода дало возможность выработать представление о механизме правового регулирования отношений с «искусственным интеллектом», показать особенности толкования правовых норм об интеллектуальной собственности применительно к рассматриваемому объекту имущественного оборота. В статье применяются и такие специальные методы познания, как логический и формально-юридический. Применение формально-юридического метода позволило сформулировать понятие «искусственного интеллекта» и рассмотреть его основные правовые характеристики. Методологической особенностью исследования является сочетание теоретического и эмпирического уровней познания. Использование указанных методов позволило проанализировать правовые вопросы об «искусственном интеллекте» во взаимосвязи с основными положениями догматики науки гражданского права. Сделан вывод, что основными и часто применяемыми на практике договорными конструкциями по распоряжению исключительным правом на «искусственный интеллект» выступают такие договоры, как договор об отчуждении исключительного права на «искусственный интеллект» и лицензионный договор. Анализируется суть законодательных изъятий из общих правил о договорах об отчуждении исключительного права и о лицензионных договорах применительно к «искусственному интеллекту».

Информация об авторе:

Л.Ю. Василевская — доктор юридических наук, профессор.

Научная статья

ТЕХНОЛОГИИ «ИСКУССТВЕННОГО ИНТЕЛЛЕКТА»: ПРОБЛЕМЫ КВАЛИФИКАЦИИ И ПРАВОВОГО РЕЖИМА

Екатерина Борисовна Подузова
Московский государственный юридический университет имени О.Е. Кутафина (МГЮА), Россия, 125993, Москва, ул. Садовая-Кудринская, 9, ekaterinak7785@yandex.ru, https: // orcid.org/ 0000-0002-1945-4761

Аннотация

На основе цивилистической методологии проведен анализ понятия, правовой природы, а также правового режима технологии «искусственного интеллекта». Предметом исследования выступают: цивилистическая доктрина; нормы российского и зарубежного законодательства; правоприменительная практика. Цель исследования состоит в теоретическом обосновании правовой природы и правового режима технологии «искусственного интеллекта». В статье применялись сравнительно-правовой, системно-структурный методы познания, а также метод моделирования. В частности, использование сравнительно-правового метода позволило рассмотреть определенные элементы технологии «искусственного интеллекта» в контексте внутригосударственного и международного регулирования отношений. Применение данного метода также дало возможность обосновать правовой режим технологии «искусственного интеллекта» как результата интеллектуальной деятельности. Использование системно-структурного метода позволило аргументировать правовую природу технологии «искусственного интеллекта», а также выявить ее структурные элементы. Применение данного метода дало возможность установить сферу применения технологий «искусственного интеллекта» (оказание информационных и медицинских услуг; разработка и использование робототехнических устройств, применяемых в нефтегазовой промышленности; проектирование многоквартирных домов и др.). Использование метода моделирования позволило раскрыть соотношение понятий «технология «искусственного интеллекта» и «форма ее закрепления», а также предложить решение вопроса о правомерности раздела исключительного права на результат интеллектуальной деятельности на доли. Методологической особенностью настоящего исследования является сочетание теоретического и эмпирического уровней познания. Использование указанного комплекса методов позволило рассмотреть вопросы правовой квалификации и правового режима технологии «искусственного интеллекта» в контексте неоднозначных доктринальных и практических подходов к их решению. В результате доказано, что технология «искусственного интеллекта», будучи сложным техническим объектом, является по своей правовой природе результатом интеллектуальной деятельности (интеллектуальной собственностью). Установлено, что доктринальные подходы к квалификации технологии «искусственного интеллекта» в качестве субъекта гражданского права («цифрового» субъекта) или вещи невозможно признать обоснованными. Аргументируется вывод, что следует признавать исключительное право на технологию «искусственного интеллекта» как на результат интеллектуальной деятельности.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ VS СУДЕЙСКОЕ УСМОТРЕНИЕ: ПЕРСПЕКТИВЫ И РИСКИ АВТОМАТИЗАЦИИ СУДЕБНОЙ ПРАКТИКИ

В.А. Родикова
Институт управления РАНХиГС, Россия 620144, Екатеринбург, ул. 8 Марта, 66, rodikovajus@gmail.com, SCIENCE INDEX: 4890-5030

Аннотация

Статья является продолжением широкой общественной дискуссии по вопросу внедрения интерфейса искусственного интеллекта в систему российского правосудия, рискам и правовым уязвимостям данного процесса на конкретных примерах правоприменительной практики. Анализируется действующее отечественное законодательство по вопросам использования искусственного интеллекта, а также последние зарубежные новеллы по данному вопросу, включая принятый Европейским советом компромиссный текст AI Act. В статье рассматриваются риск-факторы, реализуемые вследствие наличия судейского усмотрения и использования Judicial AI как в сравнении, так и самостоятельно. Изучаются контрадикторные точки зрения правоприменителей, зарубежных и российских исследователей, разработчиков систем искусственного интеллекта на перспективы использования Judicial AI в судебной системе. Обосновываются выводы об отсутствии кратко- и среднесрочных перспектив внедрения искусственного интеллекта с учетом рисков и геополитической обстановки, а также состояния законодательной базы и принципов работы системы правосудия Российской Федерации.

Ключевые слова

искусственный интеллект; нейросеть; судейское усмотрение; сильный и слабый искусственный интеллект; риск-ориентированный подход; система правосудия; прогнозирование; биометрическая идентификация.

Информация об авторе:
В.А. Родикова — аспирантка.

УПРАВЛЕНИЕ ИСКУССТВЕННЫМ ИНТЕЛЛЕКТОМ И ОПЫТ КИТАЯ В РАМКАХ КОНЦЕПЦИИ «СООБЩЕСТВА ЕДИНОЙ СУДЬБЫ ЧЕЛОВЕЧЕСТВА»

Цзя Шаосюе

Центр Международной юридической подготовки и сотрудничества для ШОС, Шанхайский политико-юридический университет, Китайская Народная Республика 201701, Шанхай, район Ципу, проспект Вэйцисун, 7989, ys02302041@126.com

Аннотация

В последние годы искусственный интеллект, опираясь на большие данные и Интернет, стремительно развивается и определяет будущее направление мирового развития науки и техники. Несмотря на то что искусственный интеллект благоприятствует научно-технической революции и индустриальной модернизации человечества, он также привел к появлению новых рисков. Приходится обращать все больше внимания на потенциальные риски искусственного интеллекта, которые следует регулировать. Риски искусственного интеллекта характеризуются разнообразием технологических угроз, сходством рисков искусственного интеллекта, с которыми сталкиваются разные страны и высокой сложностью управления, и требуют согласованных усилий всех стран. Необходимо осуществлять развитие искусственного интеллекта в стране с точки зрения общих интересов человечества, обеспечивать безопасность и управляемость искусственного интеллекта, укреплять международное сотрудничество. В настоящее время западные страны отстаивают концепцию технологической гегемонии и технологической монополии, а развивающиеся страны имеют мало возможностей выражения мнения в управлении искусственным интеллектом, и китайская концепция «Сообщества единой судьбы человечества» необходима для управления искусственным интеллектом. Начиная с данной концепции, в статье отмечаются новые китайские опыты и предложения в области внутреннего и международного управления искусственным интеллектом. В ответ на проблему чрезмерного использования и злоупотребления новыми технологиями Китай предлагает создать систему управления искусственным интеллектом, включающую совместное управление со стороны различных субъектов, открытое и прозрачное регулирование, всесторонние консультации, разработку эффективных законов, чтобы способствовать благотворному развитию искусственного интеллекта в будущем и вносить вклад в углубление управления искусственным интеллектом с помощью китайского предложения.

Ключевые слова

глобальное управление; искусственный интеллект; риски и угрозы; сообщество единой судьбы человечества; китайский опыт.

127

Информация об авторе:

Цзя Шаосюе — доктор юридических наук, доцент.

Научная статья

АНАЛИЗ РИСК-ОРИЕНТИРОВАННОГО ПОДХОДА В ПРОЕКТЕ РЕГЛАМЕНТА ЕС ОБ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ

Дмитрий Леонидович Кутейников
Осман Аликович Ижаев

[1,2] Институт государства и права Тюменского государственного университета, Россия, Тюмень, 625003, ул. Володарского, 6
[1] kuteynikov@me.com, https://orcid.org/0000-0003-1448-3085, https://elibrary.ru/author_profile.asp?id=776358
[2] izhaev.osman@gmail.com, https://orcid.org/0000-0003-3777-8927, https://elibrary.ru/author_profile.asp?id=827391

Аннотация

В статье анализируется риск-ориентированный подход, который лежит в основе проекта Регламента ЕС об ИИ. Предполагается, что данный акт будет принят до конца 2023 года и станет центральным звеном в системе правого регулирования отношений, связанных разработкой и применением систем искусственного интеллекта в Европейском союзе. Вместе с тем постоянное изменение технологических реалий создает новые проблемы и вызовы для законодателей, сбалансированное решение которых займет еще годы. Также широкое распространение за последний год базовых моделей и систем искусственного интеллекта общего назначения потребует доработки первоначальной концепции риск-ориентированного подхода.

В настоящем исследовании проанализированы общие проблемы, присущие риск-ориентированному подходу, такие как определение круга систем искусственного интеллекта, их классификация по уровням риска для прав человека, а также распространение оптимального набора юридических требований для каждой группы таких систем. Делается вывод, что для более гибкого подхода к нормативному правовому регулированию необходимо дифференцировать требования как по уровням риска, так и по всем этапам их жизненного цикла и уровням автономности. Отдельно анализируются проблемы распространения риск-ориентированного подхода на базовые модели и системы искусственного интеллекта общего назначения.

Информация об авторах:

Д.Л. Кутейников — кандидат юридических наук руководитель лаборатории «Социально-правовые подходы использования искусственного интеллекта и робототехники»
N-код: 1877-8584, AuthorID: 776358
О.А. Ижаев — кандидат юридических наук, старший научный сотрудник лаборатории «Социально-правовые подходы использования искусственного интеллекта и робототехники»

АВТОРАМ

Представленные статьи должны быть оригинальными, не опубликованными ранее в других печатных изданиях. Статьи должны быть актуальными, обладать новизной, содержать выводы исследования, а также соответствовать указанным ниже правилам оформления. В случае ненадлежащего оформления статьи она направляется автору на доработку.

Статья представляется в электронном виде в формате Microsoft Word по адресу: lawjournal@hse.ru
Адрес редакции: 109028, Москва, Б. Трехсвятительский пер, 3, оф. 113
Рукописи не возвращаются.

Объем статьи
Объем статей до 1,5 усл. п.л., рецензий — до 0,5 усл. п.л.

При наборе текста необходимо использовать шрифт «Times New Roman». Размер шрифта для основного текста статей — 14, сносок — 11; нумерация сносок сплошная, постраничная. Текст печатается через 1,5 интервала.

Название статьи
Название статьи приводится на русском и английском языке. Заглавие должно быть кратким и информативным.

Сведения об авторах
Сведения об авторах приводятся на русском и английском языках:

- фамилия, имя, отчество всех авторов полностью
- полное название организации — места работы каждого автора в именительном падеже, ее полный почтовый адрес.
- должность, звание, ученая степень каждого автора
- адрес электронной почты для каждого автора

Аннотация
Аннотация предоставляется на русском и английском языках объемом 250–300 слов.
Аннотация к статье должна быть логичной (следовать логике описания результатов в статье), отражать основное содержание (предмет, цель, методологию, выводы исследования).

Сведения, содержащиеся в заглавии статьи, не должны повторяться в тексте аннотации. Следует избегать лишних вводных фраз (например, «автор статьи рассматривает…»).

Исторические справки, если они не составляют основное содержание документа, описание ранее опубликованных работ и общеизвестные положения, в аннотации не приводятся.

Ключевые слова
Ключевые слова приводятся на русском и английском языках. Необходимое количество ключевых слов (словосочетаний) — 6–10. Ключевые слова или словосочетания отделяются друг от друга точкой с запятой.

Сноски
Сноски постраничные.
Сноски оформляются согласно ГОСТ Р 7.0.5-2008 «Система стандартов по информации, библиотечному и издательскому делу. Библиографическая ссылка. Общие требования и правила составления», утвержденному Федеральным агентством по техническому регулированию и метрологии. Подробная информация на сайте http://law-journal.hse.ru.

Тематическая рубрика
Обязательно — код международной классификации УДК.

Список литературы
В конце статьи приводится список литературы. Список следует оформлять по ГОСТ 7.0.5-2008.

Статьи рецензируются. Авторам предоставляется возможность ознакомиться с содержанием рецензий. При отрицательном отзыве рецензента автору предоставляется мотивированный отказ в опубликовании материала.

Плата с аспирантов за публикацию рукописей не взимается.