

## Articles

*Research article*

УДК: 340

DOI:10.17323/2713-2749.2023.3.4.22

# 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”.



## Keywords

artificial intelligence; object-based legal language; legal regulation of AI behavior; algorithm; driverless vehicle; AI theory.

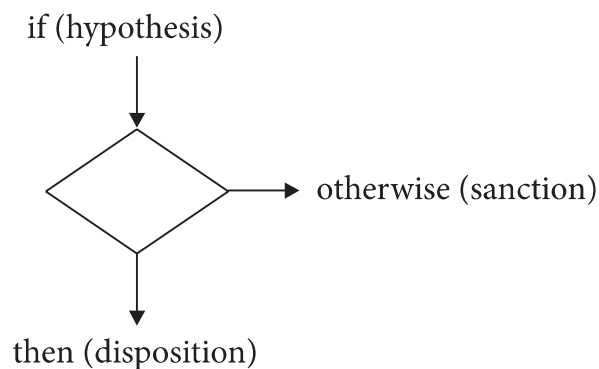
---

**For citation:** Baturin Yu.M. (2023) Specialized Legal Language–Guided AI. *Legal Issues in the Digital Age*, vol. 4, no. 3, pp. 4–22. DOI: 10.17323/2713-2749.2023.3.4.22

---

## 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.



**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<sup>1</sup> 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.

---

<sup>1</sup> 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<sup>2</sup> 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

---

<sup>2</sup> 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.<sup>3</sup> 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

---

<sup>3</sup> 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 (“1<sup>st</sup> 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 (“2<sup>nd</sup> 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),<sup>4</sup> but also for the requirements to road obstacle detection algorithms (GOST R 70251-2022),<sup>5</sup> testing requirements to road sign identification (recognition) algorithms (GOST R 70255-2022),<sup>6</sup> crossroad structure detection and reconstruction algorithms (GOST R 70253-2022),<sup>7</sup> roadside and traffic lane control algorithms (GOST R 70256-2022),<sup>8</sup> road user behavior prediction algorithms (GOST R 70254-2022)<sup>9</sup> and low-level data merge algorithms

---

<sup>4</sup> GOST R 70249-2022 AI-enabled road transport systems. Intelligent vehicles. Terms and definitions. Moscow, 2022.

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

<sup>6</sup> GOST R 70255-2022 AI-enabled road transport systems. Vehicle driving systems. Test requirements to road sign detection and recognition algorithms.

<sup>7</sup> 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.

<sup>8</sup> GOST R 70256-2022 AI-enabled road transport systems. Vehicle driving systems. Test requirements to roadside and traffic lane control algorithms.

<sup>9</sup> GOST R 70254-2022 AI-enabled road transport systems. Vehicle driving systems. Test requirements to road user behavior prediction algorithms.

(GOST R 70252-2022).<sup>10</sup> Finally, GOST R 70250-2022<sup>11</sup> 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.<sup>12</sup>

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].

---

<sup>10</sup> GOST R 70252-2022 AI-enabled road transport systems. Vehicle driving systems. Test requirements to low-level data merge algorithms.

<sup>11</sup> GOST R 70250-2022. AI-enabled road transport systems. Application options and composition of functional AI sub-systems.

<sup>12</sup> 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, still 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<sup>13</sup>

---

<sup>13</sup> 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 κοινή 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 (in 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 obrazovanii*=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.

The article was submitted to editorial office 01.10.2023; approved after reviewing 05.10.2023; accepted for publication 05.10.2023.