



Research article

DOI: <https://doi.org/10.21202/jdtl.2023.24>

The Possibility and Necessity of the Human-Centered AI in Legal Theory and Practice

Andrey V. Rezaev ✉

Saint Petersburg State University
Saint Petersburg, Russian Federation

Natalia D. Tregubova

Saint Petersburg State University
Saint Petersburg, Russian Federation

Keywords

Algorithm,
artificial intelligence,
artificial sociality,
digital economy,
digital technologies,
human,
human-centered artificial
intelligence,
law,
regulation,
sociology

Abstract

Objective: the paper aims to define the problems juridical theory and practice face with the progress of AI technologies in everyday life and correlate these problems with the human-centered approach to exploring artificial intelligence (Human-Centered AI).

Methods: the research critically analyzes the relevant literature from various disciplines: jurisprudence, sociology, philosophy, and computer sciences.

Results: the article articulates the prospects and problems the legal system confronts with the advancement of digital technologies in general and the tools of AI specifically. The identified problems are correlated with the provisions of the human-centered approach to AI. The authors acknowledge the necessity for AI inventors, as well as the owners of companies participating in the race to develop artificial intelligence technologies, to place humans, not machines, into the focus of attention as a primary value. In particular, special effort should be directed towards collecting and analyzing high-quality data for the organization of artificial intelligence tools development, taking into account that nowadays, the tools of AI are as practical as the data on which they are trained are effective.

✉ Corresponding author

© Rezaev A. V., Tregubova N. D., 2023

The English translation of the original text has been provided by the Editorial Office of the Journal of Digital Technologies and Law.

This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (CC BY 4.0) (<https://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

The authors formulate three principles of human-centered AI for the legal sphere: 1) a human as a necessary link in the chain of making and executing legal decisions; 2) the need to regulate artificial intelligence at the international law level; 3) formulating “a taboo” for introducing the artificial intelligence technologies.

Scientific novelty: the article manifests one of the first attempts in the Russian-language scientific literature to outline the prospects of developing human-centered AI methodology in jurisprudence. Based on an analysis of special literature, the authors formulate three principles of including artificial intelligence into juridical theory and practice according to the assumptions of a human-centered approach to AI.

Practical significance: the principles and arguments the article advances can be helpful in the legal regulation of artificial intelligence technologies and their harmonious inclusion into legal practices.

For citation

Rezaev, A. V., Tregubova, N. D. (2023). The Possibility and Necessity of the Human-Centered AI in Legal Theory and Practice. *Journal of Digital Technologies and Law*, 1(2), 564–580. <https://doi.org/10.21202/jdtl.2023.24>

Contents

Introduction

1. Digital technologies and law

2. Artificial intelligence in legal practice and theory: *pro et contra*

3. Human-centered artificial intelligence

Conclusion

References

Introduction

In 1948, Norbert Wiener, a founder of cybernetics, wrote: “we are already in a position to construct artificial machines of almost any degree of elaborateness of performance. Long before Nagasaki and the public awareness of the atomic bomb, it had occurred to me that we were here in the presence of another social potentiality of unheard-of importance for good and evil” (Wiener, 1983).

Today, “artificial machines” are already solving (or will be able to solve soon) multiple problems humanity faces. However, these machines undoubtedly, created new problems,

too¹. What Wiener called “artificial machines” is now, in this or that form, a part of the life of society, and we can hardly imagine our life without artificial intelligence technologies. Therefore, it is not a surprise that in recent years there has been a lot of information «noise» around artificial intelligence and its potential to radically change the world we live and work in.

The objective of our considerations here is to show that as artificial intelligence technologies are being developed and introduced into our daily life, the necessity is proportionally increasing for the software developers, designers, and owners of the companies participating in the race to introduce new AI tools, to have humans and their needs but not machines and their efficiency, as the primary value and goal of advancement. It is not the goals of one person or company, not the technologies or machines, but a human and the human attitude that serves as the measure of morals and humanness. Realizing that good-hearted calls for humanness may sound abstract in the logic of technologically oriented development, we would like to discuss more specifically how the need to work with artificial intelligence within the approach called Human-Centered AI (HCAI).

The problem of the artificial intelligence orientation towards the good of humans is acute in all spheres of life but especially sensitive in some of them. These include education, medicine, and jurisprudence, where the price of a mistake – of a human or an algorithm – is the highest. Juridical solutions regulate human life and their relations with others and sometimes refer to existential issues – life, death, and justice.

In this article, we consider the problems the juridical theory and practice face with the advancement of AI in everyday life and how these problems correlate with the human-centered approach to AI. We define artificial intelligence as “an ensemble of rational, logical, and formalized instrumental rules developed and coded by human beings that organize the processes and activities to emulate rational/intellectual structures and fabricate and reproduce goal-oriented practices as well as the mechanisms for constructing further coding and decision making.” (Rezaev & Tregubova, 2019).

Today, one of the factors determining the development of artificial intelligence technologies is online culture – “an ensemble of communication networks, devices, algorithms, formal and informal rules of interaction, patterns of behavior, cultural symbols, which allow and structure people’s activity in the internet and similar networks, providing remote access to creating, exchanging and obtaining information” (Rezaev & Tregubova, 2019).

¹ To confirm this, one may cite a recent statement by Sam Altman, a founder of OpenAI company which developed a famous ChatGPT chatbot: “I think where we are right now is not where we want to be. The way this should work is that there are extremely wide bounds of what these systems can do, that are decided by, not Microsoft or OpenAI, but society, governments, something like that, some version of that, people, direct democracy. <...> It’s very new technology. We don’t know how to handle it.” *Bing’s Revenge and Google’s AI FacePlant*. <https://www.nytimes.com/2023/02/10/podcasts/bings-revenge-and-googles-ai-face-plant.html>

The Internet provides vast data on which artificial intelligence algorithms are trained and the “platform” for these algorithms to act.

As a result of the simultaneous development of computation capacities of artificial intelligence online culture, artificial intelligence is increasingly involved in everyday life and human relationships. “Artificial sociality” appears: artificial intelligence becomes an active mediator and participant in social interactions (Rezaev & Tregubova, 2019).

From its inception, the AI project had an a-disciplinary character. The artificial intelligence developers strived to reproduce human intelligence, hence, boldly borrowed the necessary provisions from mathematics, psychology, cybernetics, etc. (Russell & Norvig, 2007). However, while developing the machines reproducing the functioning of the human mind required turning to the achievements from various fields of knowledge, this is all the more true to understand how these machines enter the everyday life of society and are built into social relations. In other words, researching the problems of artificial sociality has an interdisciplinary and potentially – “a-disciplinary” character. That is why, in this article, we rely on both the philosophical and sociological analysis of the problems of AI, and the results of research in jurisprudence and law.

Further reflections are organized as follows. First, we will pay attention to several vital aspects associated with the introduction of digital technologies into legal practices. Then we will consider the problems and prospects of the rapid penetration of AI into the everyday life of society, which changes the characteristics of juridical work and the structure of legal systems. Finally, we will turn to the human-centered approach to AI and its consequences for the legal sphere.

1. Digital technologies and law

Summarizing the influence of digital technologies on the legal system, one should emphasize the following.

First, digital technologies simplified access to legal information via online databases, legal search systems, and other online resources. This fact, accordingly, created opportunities for nearly every Internet user (that is, almost 90% of the Russian population)² to perform online research to obtain legal information. The Internet revolutionized search in all spheres of human life (Utekhin, 2019) and legal information is no exception.

Second, digital technologies improved communication between lawyers, clients, and other actors in the legal system. For example, videoconferencing allows lawyers to communicate with their clients distantly. Thus, access to legal services for residents of remote districts and regions is improved.

² Dmitriy Chernyshenko: Russia has about 130 million Internet users today – almost 90% of the population. <http://government.ru/news/46639/>

Third, digital technologies made it possible to submit and store legal documents electronically. In other words, organizing juridical practices with digital technologies significantly decreases the need for paper documents and simplifies information search and exchange ([Rusakova, 2020](#); [Stepanov et al., 2021](#)).

Fourth, digital technologies have led to automating many legal processes, such as routine checks of documents and compiling contracts (including the so-called smart contracts ([Efimova et al., 2020](#))). Accordingly, the demand for routine manual labor of lawyers and their assistants has significantly decreased.

Fifth, using digital technologies in legal practices gave rise to new branches of law, such as cyberlaw/law in cyberspace ([Mazhorina, 2020](#)), intellectual property law, and data protection law ([Voinikanis, 2020](#)).

Thus, digital technologies have already significantly influenced the development of law, making legal services and practices more accessible, efficient, and effective. At the same time, practicing lawyers, special literature, mass media, and everyday legal service consumers almost unanimously emphasize that digital technologies generate new problems for the legal system development. These are, first of all, the issues of confidentiality ([Talapina, 2022](#)) and accessibility ([Panchenko, 2012](#)) of legal databases and the problem of critical assessment of the information obtained from the Internet ([Greger, 2017](#)).

The current stage of digital technologies development in online culture suggests paying attention to how artificial intelligence transforms and shapes further development of legal practices. What are the advantages and disadvantages of using AI technologies in routine legal practice?

2. Artificial intelligence in legal practice and theory: *pro et contra*

Using artificial intelligence technologies in routine legal practice has both advantages and disadvantages. The main benefits are the following:

- Effective organization of the lawyers' practices. The AI instruments automate and accelerate the performance of such tasks as document review, preliminary juridical examination of literature sources, and analysis of contracts ([Talapina, 2021](#)).
- Artificial intelligence may perform specific tasks more accurately than people, for example, find regularities in data or check documents for factual mistakes, and grammatical or stylistic inconsistencies ([Andreev et al., 2020](#)).
- The use of AI technologies, reducing the need for manual work, saves the funds of juridical companies and their clients.

- Artificial intelligence technologies provide lawyers with more complete, comprehensive, and detailed information allowing them to make more grounded decisions.

The main disadvantages of using artificial intelligence are the following:

- The common disadvantage of using artificial intelligence for all professions is that some professions disappear while others will appear and come to the fore (Lee, 2019). A broad use of AI technologies in legal practice is still a potential, but very soon, it will inevitably lead to a review of jobs nomenclature within the juridical system structure; this will especially touch upon paralegals and other auxiliary staff (Lessig, 2019).

- artificial intelligence systems are, to a certain extent, translators of bias and prejudices characteristic of their creators (Gorokhova, 2021). The artificial intelligence algorithms may be biased or erroneous due to, at least, two circumstances: a) if they are based on and were developed with biased or erroneous data arrays; b) if they are misused. Hence, the introduction of AI technologies implies searching for ways to ensure just and bias-free artificial intelligence systems.

- artificial intelligence technologies, like any other technologies, bear safety risks. AI technologies cannot guarantee complete cybersecurity (O'Neil, 2018). Artificial intelligence may minimize but not eliminate data leakage or hacking. Accordingly, confidentiality – the cornerstone of legal practices – is threatened when using artificial intelligence technologies.

Thus, using AI technologies in everyday legal practices provides multiple advantages, but they should be weighed with potential risks and drawbacks. Lawyers must not only realize the capabilities of AI, but also thoroughly review their use and see their limitations and potential risks.

Besides the problems with using the algorithms and machines which are already manifested in everyday life, one should also keep in mind the actual problems generated by the ubiquitous penetration of artificial intelligence into legal practices:

- confidentiality problems. The effective performance of artificial intelligence systems often requires access to large amounts of personal data, which causes concerns about confidentiality and data protection. Regulators and legislators must find a balance between privacy protection and promoting innovations in the sphere of artificial intelligence (Gorokhova, 2021).

- legal liability for the actions performed by artificial intelligence. As artificial intelligence systems become more autonomous and make decisions without human interference, questions arise about who is responsible for their actions (Vavilin, 2021; Baturin & Polubinskaya, 2022). For example, if an AI driverless car caused an accident, should

a developer, a user, or the artificial intelligence system per se be liable (Rudenko, 2020)? Who will bear responsibility if something goes wrong when AI instruments are used? Who will be responsible for accidents or mistakes caused by an artificial intelligence system: a programmer, an owner (of what?) or the AI designer? These already are the juridical questions of today.

- the issues related to intellectual property rights to the products created by artificial intelligence technologies (Lee et al., 2021). For example, who will be deemed an inventor or an artist if an artificial intelligence system creates a work of art or invents a new technology?

- a critical element is using artificial intelligence tools (for example, ChatGPT) for juridical interpretation of documents and applying legal norms, especially regarding the complex and nuanced character of juridical substantiation of a certain decision. There are grounds to fear that artificial intelligence will not be able to comprehensively grasp human considerations and judgments necessary for effective juridical decision-making (Tsvetkov, 2021).

- lack of communication and real-life human contacts. This is a significant disadvantage for legal practices, which may, by default, touch upon existential matters of life and death, restriction of freedom. Judges note: justice is impossible without a holistic view of the situation, including moral and emotional aspects, which is inaccessible for AI (Bykov & Narskaya, 2022).

Noteworthy, the very question of whether AI technologies should be regulated and how is the object of discussion (Etzioni & Etzioni, 2017). The frameworks are just starting to be elaborated in this sphere, the “pioneers” often being the legislators of the European Union (Hickman & Petrin, 2021; Fink & Finck, 2022; Ulnicane, 2022). In Russia, legal regulation of artificial intelligence technologies is also being developed. In 2019, the National strategy of artificial intelligence development up to 2030³ was adopted, specifying the basic definitions and general principles of using AI technologies.

Thus, the development of artificial sociality poses both a practical and conceptual problem for jurisprudence. The practical/functional problem is how artificial intelligence technologies will change legal practices, while the conceptual problem refers to their legal regulation.

We believe that a set of problems which have already emerged and which are bound to emerge in the future in legal practice can be solved more effectively with the approach called a human-centered AI.

³ On the development of artificial intelligence in the Russian Federation: Executive Order of the President of the Russian Federation. <http://static.kremlin.ru/media/events/files/ru/AH4x6HgKWANwVtMOfPDhcbRpvd1HCCsv.pdf>

3. Human-centered artificial intelligence

The approach called a Human-Centered AI in the scientific literature (Ford et al., 2015; Shneiderman, 2021)⁴ implies, first of all, understanding the straightforward fact that people and machines are not the same⁵. There is no need to aim at making an artificial intelligence tool similar to a human being. On the contrary, success will probably be achieved in the opposite direction when a human stays a human with their intellect, consciousness, subconsciousness, and emotional and spiritual world. At the same time, machines and algorithms will be developed by a human and, during “self-training,” follow their own logic of development, different from that of a human.

Unfortunately, this circumstance is being neglected, just like the human-centered approach to AI in general. Most technological leaders in the USA and other countries continue spending a lot on developing software that can do just what people can do. Developers very well realize that they can earn easy money by selling their products to corporations having no other orientations in their development except those set by the logic of the market and profit (Zuboff, 2022). Everyone is focused on using artificial intelligence to reduce costs for the working force while caring little about the essence of the social progress and development of a moral human being and a just society.

Human-centered AI requires immediate attention to collecting and analyzing high-quality data to organize the development of artificial intelligence tools. The artificial intelligence algorithms are as effective as the data on which they are trained are effective. In contrast, partial or incomplete data may lead to not only unjust/false results but to ones opposite to the initial goal. Collecting information for self-learning models must be diverse and representative; the data must reflect the real world we live in and the people we work with, regardless of their social and class differences.

Elsewhere, we have already emphasized the fact that, under the current stage of capitalism development, under the extreme orientation towards financial indicators, profit, and functional efficiency, it is practically impossible to solve these problems (Rezaev, 2021)⁶. Nevertheless, it would be wrong for the tactics of social sciences development not to consider them at all and not to attempt to propose variants of their solution.

⁴ Notably, in 2019 a Human-Centered AI Institute was established at Stanford University (USA) – the largest research center in this area.

⁵ This statement has been repeatedly made in philosophy and social sciences. See (Dreyfus, 1978; Wolfe, 1993; Esposito, 2017).

⁶ For example, Elon Musk (who sponsored OpenAI company which developed ChatGPT) said with obvious regret that he donated money (US\$ 1 billion) to create an open platform aimed at free open access, while now ChatGPT is an opposite model – close and fully aimed at profits. However, Elon Musk executes no control over OpenAI or ChatGPT at the moment. *Elon Musk at the 2023 World Government Summit in Dubai*. https://www.youtube.com/watch?v=jmNrlNgXx_U&ab_channel=ElonAlerts

The market has never been and cannot be (even under artificial sociality) the touchstone of beauty, goodness, and truth. Strategically, social knowledge substantiated the impossibility of a harmonious, moral, and just world without exploitation of human by human and without social and cultural inequality within the framework of a capitalist economy⁷. However, what the problems for society are and what the variants of the social development trajectories are under the still uncontrolled spreading of AI tools – these topics are just starting to be considered belatedly.

Characterizing the features of artificial intelligence development, one should remember that AI technologies are not neutral. Humans create them, and algorithms reproduce their creators' values, biases, and prejudices. Thus, AI designers and producers must adhere to ethical and human-oriented approaches. It means, among other things, accounting for various viewpoints and opinions in the design process, providing transparency and accountability, and paying priority attention to the human personality and wellbeing of society in general, not that of individual subjects or technological systems.

The key point is the understanding that AI instruments are already a powerful means of solving some of the most burning problems facing society, but they are not a panacea. While defining and formulating the directions of social development, one should not rely exclusively on artificial intelligence when solving social, economic, cultural, and political problems. Even under "artificial sociality," people must remain within the reality of human experience and admit that social progress requires more than just technological solutions.

⁷ An example is "The Wealth of Nations" by Adam Smith. Although Smith is often called the first theoretician of political economy and an advocate of capitalism, his works are critical regarding many aspects of a capitalist economy. For example, he postulates that a rush for profits may lead to a lack of concern about the well-being of workers and society as a whole and that to provide a just and equal society a certain form of state intervention is necessary. In his work "The Great Transformation" Karl Polanyi states that capitalism is a historically recent phenomenon that generated absolutely negative consequences for social development, including turning labor into a fictitious commodity, destroying a traditional way of life, and rising nationalistic and fascist movements. Thorstein Veblen in his book "The Theory of the Leisure Class" showed that capitalism creates a conspicuous consumption and wastefulness culture when people are praised for their ability to consume and demonstrate their wealth, not for their contribution to society. A contemporary Canadian researcher Naomi Klein asserts that capitalism is often imposed on society by violence and coercion and is often used by the wealthy elite to maintain their political and economic power (Klein, 2007). See for more details (Harvey, 2014).

Conclusion

This essay began with a citation from Norbert Wiener, the founder of cybernetics. We want to conclude it with the judgments presented by one of the founders of artificial intelligence research – Joseph Weizenbaum. Weizenbaum argued that the use of computers should be banned or at least restricted in two cases (Weizenbaum, 1982). The first case is connected with attempts to replace a human with a machine in the areas related to interpersonal relationships, love, and understanding. The second case is using computers in a situation where it can lead to irreversible consequences. In our opinion, Weizenbaum correctly formulated the basic principles of human-centered artificial intelligence, which relate to the general spread of AI technologies and their use in the theory and practice of jurisprudence, in particular.

In conclusion, we will formulate three principles of including AI into the legal theory and practice according to the methodological principles of human-centered approach (HCAI).

First. A human being must always remain within the chain of making/executing legal decisions. Legal scientists have persistently formulated this thesis. Artificial intelligence technologies may take on many tasks in legal practice, but it is a human being that must control, check, conceptualize, and weigh the actions and decisions of the artificial intelligence.

Second. Today, we need to elaborate the laws determining a rational and understandable *modus vivendi* for the activity of artificial intelligence in social systems aimed at a human being, not at profit and market. This is almost impossible within one state, especially a capitalist one. That is why the world faces the need to create international law for AI involvement in society. Like any rule, the law may be violated – by mistake or out of malice. But violation of the law does not repeal the law itself; it just reveals the malicious persons who distorted the law.

Third. The progress of AI in everyday life of people poses the need for prohibitions, including juridical ones, a taboo for using artificial intelligence in certain spheres of human life (Rezaev, 2021). These are, first of all, spheres associated with existential issues. For example, an important issue is whether one should use artificial intelligence to determine if a person is lying (Oravec, 2022), or whether artificial intelligence may serve as an autonomous weapon (International Committee, 2020). Defining such spheres at international, national, and local levels, formulating legal prohibitions, and creating law-enforcement mechanisms is one of the priority tasks for Human-Centered AI.

References

- Andreev, V. K., Laptev, V. A., & Chucha S. Yu. (2020). Artificial intelligence in the system of electronic justice by consideration of corporate disputes. *Vestnik of Saint Petersburg University. Law*, 11(1), 19–34. (In Russ.). <https://doi.org/10.21638/spbu14.2020.102>

- Baturin, Yu. M., & Polubinskaya, S. V. (2022). Artificial intelligence: legal status or legal regime? *Gosudarstvoiparvo*, 10, 141. (In Russ.). <https://doi.org/10.31857/s102694520022606-7>
- Bykov, A. V., & Narskaya, A. I. (2022). Law, Morality, and Machine Learning: Judges' Perspective on the Essence of Justice and the Prospects of Its Robotization. *Monitoring of Public Opinion: Economic and Social Changes Journal (Public Opinion Monitoring)*, 5, 278–298. (In Russ.). <https://doi.org/10.14515/monitoring.2022.5.2137>
- Dreyfus, H. (1978). *What computers can't do: A critique of artificial reason*. Moscow: Progress. (In Russ.).
- Efimova, L., Mikheeva, I., & Chub, D. (2020). Comparative Analysis of Doctrinal Concepts of Legal Regulating Smart Contracts in Russia and Foreign States. *Journal of the Higher School of Economics*, 4, 78–105. (In Russ.).
- Esposito, E. (2017). Artificial Communication? The Production of Contingency by Algorithms. *Zeitschrift für Soziologie*, 46(4), 249–265. <https://doi.org/10.1515/zfsoz-2017-1014>
- Etzioni, A., & Etzioni, O. (2017). Should Artificial Intelligence Be Regulated? *Issues in Science and Technology*, 33(4), 32–36.
- Fink, M., & Finck, M. (2022). Reasoned A(I) administration: explanation requirements in EU law and the automation of public administration. *European Law Review*, 47(3), 376–392.
- Ford, K. M., Hayes, P. J., Glymour, C., & Allen, J. (2015). Cognitive Orthoses: Toward Human-Centered AI. *AI Magazine*, 36(4), 5–8. <https://doi.org/10.1609/aimag.v36i4.2629>
- Gorokhova, S. S. (2021). Artificial intelligence: an instrument ensuring cybersecurity of the financial sphere or a cyber threat to banks? *Banking Law*, 1, 35–46. (In Russ.). <https://doi.org/10.18572/1812-3945-2021-1-35-46>
- Greger, R. (2017). Judge as an Internet Surfer. Identification of the Circumstances of the Case on the Internet. *Herald of Civil Procedure*, 7(4), 161–173. (In Russ.). <https://doi.org/10.24031/2226-0781-2017-7-4-161-173>
- Harvey, D. (2014). *Seventeen Contradictions and the End of Capitalism*. Oxford: Oxford University Press.
- Hickman, E., & Petrin, M. (2021). Trustworthy AI and Corporate Governance: The EU's Ethics Guidelines for Trustworthy Artificial Intelligence from a Company Law Perspective. *European Business Organization Law Review*, 22, 593–625. <https://doi.org/10.1007/s40804-021-00224-0>
- International Committee of the Red Cross (2020). Artificial intelligence and machine learning in armed conflict: A human-centred approach. *International Review of the Red Cross*, 102(913), 463–479. <https://doi.org/10.1007/s40804-021-00224-0>
- Klein, N. (2007). *The Shock Doctrine: The Rise of Disaster Capitalism*. New York: Henry Holt.
- Lee, J.-A., Hilty, R. M., & Liu, K.-C. (Eds.). (2021). *Artificial Intelligence and Intellectual Property*. Oxford: Oxford University Press.
- Lee, K.-F. (2019). *AI Superpowers: China, Silicon Valley and the new world order*. Moscow: Mann, Ivanov i Ferber. (In Russ.).
- Lessig, L. (2019). Artificial intelligence is going to oust a wide circle of lawyers. *Zakon*, 5, 8–30. (In Russ.).
- Mazhorina, M. (2020). Cyberplace and Methodology of International Private Law. *Journal of the Higher School of Economics*, 2, 230–253. (In Russ.).
- O'Neil, C. (2018). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Moscow: AST. (In Russ.).
- Oravec, J. A. (2022). The emergence of “truth machines”? Artificial intelligence approaches to lie detection. *Ethics and Information Technology*, 24, 6. <https://doi.org/10.1007/s10676-022-09621-6>
- Panchenko, V. Yu. (2012). Information availability of legal assistance: ideal and real state. *Agrarnoye i zemelnoye parvo*, 11(95), 95–102. (In Russ.).
- Rezaev, A. V. (2021). Twelve Theses on Artificial Intelligence and Artificial Sociality. *Monitoring of Public Opinion: Economic and Social Changes*, 1, 20–30. <https://doi.org/10.14515/monitoring.2021.1.1894>
- Rezaev, A. V., Tregubova, N. D. (2019). Artificial intelligence, On-line Culture, Artificial Sociality: Definition of the Terms. *Monitoring of Public Opinion: Economic and Social Changes*, 6, 35–47. <https://doi.org/10.14515/monitoring.2019.6.03>
- Rudenko, N. (2020). Sociotechnical barriers to developing autonomous vehicles in Russia. In L. Zemnukhova, K. Glazkov, O. Logunova, A. Maksimova, D. Sivkov, & N. Rudenko, *The Adventures of Technologies: Digitalization Barriers in Russia* (17–70). Moscow – Saint Petersburg: FNISTS RAN. (In Russ.). <https://doi.org/10.31119/978-5-89697-339-3>
- Rusakova, E. (2020). The integration of modern digital technologies to the legal proceedings of People's Republic of China and Singapore, *Gosudarstvo i parvo*, 9, 102. (In Russ.). <https://doi.org/10.31857/s102694520011323-6>
- Russell, S., & Norvig, P. (2007). *Artificial Intelligence: A Modern Approach* (2d ed.). Moscow: Vilyams. (In Russ.).
- Shneiderman, B. (2021). Human-centered AI. *Issues in Science and Technology*, 37(2), 56–61.

- Stepanov, O., Pechegin, D., & Diakonova, M. (2021). Towards the Issue of Digitalization of Judicial Activities. *Journal of the Higher School of Economics*, 5, 4–23. (In Russ.). <https://doi.org/10.17323/2072-8166.2021.5.4.23>
- Talapina, E. V. (2021). Artificial intelligence and legal expertise in public administration. *Vestnik of Saint Petersburg University. Law*, 12(4), 865–881. (In Russ.). <https://doi.org/10.21638/spbu14.2021.404>
- Talapina, E. V. (2022). The right to informational self-determination: on the edge of public and private. *Law. Journal of the Higher School of Economics*, 15(5), 24–43. (In Russ.).
- Tsvetkov, Yu. A. (2021). Artificial Intelligence in Justice. *Zakon*, 4, 91–107. (In Russ.).
- Ulnicane, I. (2022). Artificial Intelligence in the European Union: policy, ethics and regulation. In T. Hoerber, I. Cabras, G. Weber (Eds.). *Routledge Handbook of European Integrations* (pp. 254–269). London: Routledge. <https://doi.org/10.4324/9780429262081-19>
- Utekhin, I. (2019). Search and Interfaces for Search. *Laboratorium: Russian Review of Social Research*, 11(1), 152–165. (In Russ.). <https://doi.org/10.25285/2078-1938-2019-11-1-152-165>
- Vavilin, E. V. (2021). Artificial intelligence as a participant in civil relations: the transformation of law. *Vestnik Tomskogo gosudarstvennogo universiteta. Pravo*, 42, 135–146. (In Russ.). <https://doi.org/10.17223/22253513/42/11>
- Voinikanis, E. A. (2020). Regulation of big data and intellectual property right: common approaches, problems and prospects of development. *Zakon*, 7, 135–156. (In Russ.).
- Weizenbaum, J. (1982). *Computer Power and Human Reason: From Judgment to Calculation*. Moscow: Radio i svyaz. (In Russ.).
- Wiener, N. (1983). *Cybernetics: Or Control and Communication in the Animal and the Machine* (2d ed.). Moscow: Nauka; Glavnaya redaktsiya izdaniy dlya zarubezhnykh stran. (In Russ.).
- Wolfe, A. (1993). *The Human Difference: Animals, Computers, and the Necessity of Social Science*. Berkeley: University of California Press.
- Zuboff, Sh. (2022). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Moscow: Izd-vo Instituta Gaidara. (In Russ.).

Authors information



Andrey V. Rezaev – Doctor of Philosophical Sciences, Professor, Head of the International Research Laboratory TANDEM at the Faculty of Sociology, St Petersburg State University

Address: 1/3 Smolnogo Str., 191124 Saint Petersburg, Russia

E-mail: rezaev@hotmail.com

ORCID ID: <https://orcid.org/0000-0002-3918-835X>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=13004674100>

Web of Science Researcher ID:

<https://www.webofscience.com/wos/author/record/K-3472-2013>

Google Scholar ID: <https://scholar.google.ru/citations?user=Uzv39ccAAAAJ>

РИНЦ Author ID: https://elibrary.ru/author_items.asp?authorid=648768



Natalia D. Tregubova – PhD (Sociology), Associate Professor of the Department of Comparative Sociology, Saint Petersburg State University

Address: 1/3 Smolnogo Str., 191124 Saint Petersburg, Russia

E-mail: n.tregubova@spbu.ru

ORCID ID: <https://orcid.org/0000-0003-3259-5566>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=56645016900>

Web of Science Researcher ID:

<https://www.webofscience.com/wos/author/record/K-3487-2013>

Google Scholar ID: <https://scholar.google.com/citations?user=8dhGr3gAAAAJ&hl>

РИНЦ Author ID: https://elibrary.ru/author_items.asp?authorid=832705

Authors' contributions

A. V. Rezaev and N. D. Tregubova contributed equally to the formulation of the article's key provisions and preparation of the manuscript for publication.

Conflict of interests

The authors declare no conflict of interests.

Financial disclosure

The study was supported by RFBR and MOST, the research project No. 21-511-52002.

Thematic rubrics:

OECD: 5.05 / Law

PASJC: 3308 / Law

WoS: OM / Law

Article history

Date of receipt – March 3, 2023

Date of approval – May 4, 2023

Date of acceptance – June 16, 2023

Date of online placement – June 20, 2023



Научная статья

УДК 340.143:004.8

EDN: <https://elibrary.ru/sadrzw>

DOI: <https://doi.org/10.21202/jdtl.2023.24>

Возможность и необходимость человеко-ориентированного искусственного интеллекта в юридической теории и практике

Андрей Владимирович Резаев ✉

Санкт-Петербургский государственный университет
г. Санкт-Петербург, Российская Федерация

Наталья Дамировна Трегубова

Санкт-Петербургский государственный университет
г. Санкт-Петербург, Российская Федерация

Ключевые слова

Алгоритм,
искусственная
социальность,
искусственный интеллект,
право,
регулирование,
социология,
цифровая экономика,
цифровые технологии,
человек,
человеко-ориентированный
искусственный интеллект

Аннотация

Цель: определение проблем, которые ставит распространение технологий искусственного интеллекта перед юридической теорией и практикой, и соотнесение этих проблем с человеко-ориентированным подходом к искусственному интеллекту (Human-Centered AI).

Методы: исследование основано на критическом анализе релевантной литературы из разных дисциплинарных областей: юриспруденции, социологии, философии, компьютерных наук.

Результаты: в статье сформулированы основные перспективы и проблемы, которые ставит перед правовой системой развитие цифровых технологий в целом и алгоритмов искусственного интеллекта в частности. Выделенные проблемы соотнесены с положениями человеко-ориентированного подхода к искусственному интеллекту. Авторы утверждают необходимость того, чтобы разработчики программ искусственного интеллекта, владельцы компаний, участвующих в гонке по внедрению технологий искусственного интеллекта, сосредоточивались на том, чтобы люди и человек как базовая ценность общества находились в центре внимания. В частности, специальные усилия следует направить на сбор и анализ высококачественных данных для организации разработок инструментов искусственного интеллекта, поскольку сегодня алгоритмы искусственного интеллекта эффективны

✉ Контактное лицо

© Резаев А. В., Трегубова Н. Д., 2023

Статья находится в открытом доступе и распространяется в соответствии с лицензией Creative Commons «Attribution» («Атрибуция») 4.0 Всемирная (CC BY 4.0) (<https://creativecommons.org/licenses/by/4.0/deed.ru>), позволяющей неограниченно использовать, распространять и воспроизводить материал при условии, что оригинальная работа упомянута с соблюдением правил цитирования.

настолько, насколько эффективны данные, на которых они обучаются. Авторы формулируют три принципа человеко-ориентированного искусственного интеллекта для правовой сферы: 1) человек как необходимое звено в цепочке принятия и исполнения правовых решений; 2) необходимость регулирования искусственного интеллекта на уровне международного права; 3) формулировка «табу» для внедрения технологий искусственного интеллекта.

Научная новизна: статья представляет собой первую в русскоязычной научной литературе попытку обозначить перспективы развития области человеко-ориентированного искусственного интеллекта в юриспруденции. На основании анализа специальной литературы авторы формулируют три принципа включения искусственного интеллекта в юридическую теорию и практику с точки зрения человеко-ориентированного подхода к искусственному интеллекту.

Практическая значимость: принципы, сформулированные в статье, будут полезны как для правового регулирования технологий искусственного интеллекта, так и для гармоничного их включения в юридические практики.

Для цитирования

Резаев, А. В., Трегубова, Н. Д. (2023). Возможность и необходимость человеко-ориентированного искусственного интеллекта в юридической теории и практике. *Journal of Digital Technologies and Law*, 1(2), 564–580. <https://doi.org/10.21202/jdtl.2023.24>

Список литературы

- Андреев, В. К., Лаптев, В. А., Чуча, С. Ю. (2020). Искусственный интеллект в системе электронного правосудия при рассмотрении корпоративных споров. *Вестник Санкт-Петербургского университета. Право*, 1, 19–34. <https://doi.org/10.21638/spbu14.2020.102>
- Батурин, Ю. М., Полубинская, С. В. (2022). Искусственный интеллект: правовой статус или правовой режим? *Государство и право*, 10, 141–154. <https://doi.org/10.31857/s102694520022606-7>
- Быков, А. В., Нарская, А. И. (2022). Закон, мораль и машинное обучение: взгляд судей на сущность и перспективы роботизации правосудия. *Мониторинг общественного мнения: экономические и социальные перемены*, 5, 278–298. <https://doi.org/10.14515/monitoring.2022.5.2137>
- Вавилин, Е. В. (2021). Искусственный интеллект как участник гражданских отношений: трансформация права. *Вестник Томского государственного университета. Право*, 42, 135–146. <https://doi.org/10.17223/22253513/42/11>
- Вейценбаум, Дж. (1982). *Возможности вычислительных машин и человеческий разум. От суждений к вычислениям*. Москва: Радио и связь.
- Винер, Н. (1983). *Кибернетика, или управление и связь в животном и машине* (2-е изд.). Москва: Наука; Главная редакция изданий для зарубежных стран.
- Войниканис, Е. А. (2020). Регулирование больших данных и право интеллектуальной собственности: общие подходы, проблемы и перспективы развития. *Закон*, 7, 135–156.
- Горохова, С. С. (2021). Искусственный интеллект: инструмент обеспечения кибербезопасности финансовой сферы или киберугроза для банков. *Банковское право*, 1, 35–46. <https://doi.org/10.18572/1812-3945-2021-1-35-46>
- Грегер, Р. (2017). Судья как интернет-серфер. Выяснение обстоятельств дела в Интернете. *Вестник гражданского процесса*, 7(4), 161–173. <https://doi.org/10.24031/2226-0781-2017-7-4-161-173>
- Дрейфус, Х. (1978). *Чего не могут вычислительные машины: Критика искусственного разума*. Москва: Прогресс.
- Ефимова, Л. Г., Михеева, И. В., Чуб, Д. В. (2020). Сравнительный анализ доктринальных концепций правового регулирования смарт-контрактов в России и в зарубежных странах. *Право. Журнал Высшей школы экономики*, 4, 78–105. <https://doi.org/10.17323/2072-8166.2020.4.78.105>

- Зубофф, Ш. (2022). *Эпоха надзорного капитализма. Битва за человеческое будущее на новых рубежах власти*. Москва: Изд-во Института Гайдара.
- Лессиг, Л. (2019). Искусственный интеллект вытеснит широкий пласт юристов. *Закон*, 5, 8–30.
- Ли, К.-Ф. (2019). *Сверхдержавы искусственного интеллекта. Китай, Кремниевая долина и новый мировой порядок*. Москва: Манн, Иванов и Фербер.
- Мажорина, М. В. (2020). Киберпространство и методология международного частного права. *Право. Журнал Высшей школы экономики*, 2, 230–253. <https://doi.org/10.17323/2072-8166.2020.2.230.253>
- О'Нил, К. (2018). *Убийственные большие данные. Как математика превратилась в оружие массового поражения*. Москва: АСТ.
- Панченко, В. Ю. (2012). Информационная доступность юридической помощи: идеальная модель и реальное состояние. *Аграрное и земельное право*, 11(95), 95–102.
- Рассел, С., Норвиг, П. (2007). *Искусственный интеллект: современный подход* (2-е изд.). Москва: Вильямс.
- Резаев, А. В., Трегубова, Н. Д. (2019). «Искусственный интеллект», «онлайн-культура», «искусственная социальность»: определение понятий. *Мониторинг общественного мнения: Экономические и социальные перемены*, 6, 35–47. <https://doi.org/10.14515/monitoring.2019.6.03>
- Руденко, Н. И. (2020). Социотехнические барьеры разработки беспилотных автомобилей в России. В кн. Л. В. Земнухова и др. *Приключения технологий: барьеры цифровизации в России* (с. 17–70). Москва – Санкт-Петербург: ФНИСЦ РАН. <https://doi.org/10.31119/978-5-89697-339-3>
- Русакова, Е. П. (2020). Интегрирование современных цифровых технологий в судопроизводство Китайской Народной Республики и Сингапура. *Государство и право*, 9, 102–109. <https://doi.org/10.31857/s102694520011323-6>
- Степанов, О. А., Печегин, Д. А., Дьяконова, М. О. (2021). К вопросу о цифровизации судебной деятельности. *Право. Журнал Высшей школы экономики*, 5, 4–23. <https://doi.org/10.17323/2072-8166.2021.5.4.23>
- Талапина, Э. В. (2021). Искусственный интеллект и правовые экспертизы в государственном управлении. *Вестник Санкт-Петербургского университета. Право*, 4, 865–881. <https://doi.org/10.21638/spbu14.2021.404>
- Талапина, Э. В. (2022). Право на информационное самоопределение: на грани публичного и частного. *Право. Журнал Высшей школы экономики*, 15(5), 24–43.
- Утехин, И. (2019). Поиск и его интерфейсы. *Laboratorium: журнал социальных исследований*, 11(1), 152–165. <https://doi.org/10.25285/2078-1938-2019-11-1-152-165>
- Цветков, Ю. А. (2021). Искусственный интеллект в правосудии. *Закон*, 4, 91–107.
- Esposito, E. (2017). Artificial Communication? The Production of Contingency by Algorithms. *Zeitschrift für Soziologie*, 46(4), 249–265. <https://doi.org/10.1515/zfsoz-2017-1014>
- Etzioni, A., & Etzioni, O. (2017). Should Artificial Intelligence Be Regulated? *Issues in Science and Technology*, 33(4), 32–36.
- Fink, M., & Finck, M. (2022). Reasoned A(l)administration: explanation requirements in EU law and the automation of public administration. *European Law Review*, 47(3), 376–392.
- Ford, K. M., Hayes, P. J., Glymour, C., & Allen, J. (2015). Cognitive Orthoses: Toward Human-Centered AI. *AI Magazine*, 36(4), 5–8. <https://doi.org/10.1609/aimag.v36i4.2629>
- Harvey, D. (2014). *Seventeen Contradictions and the End of Capitalism*. Oxford: Oxford University Press.
- Hickman, E., & Petrin, M. (2021). Trustworthy AI and Corporate Governance: The EU's Ethics Guidelines for Trustworthy Artificial Intelligence from a Company Law Perspective. *European Business Organization Law Review*, 22, 593–625. <https://doi.org/10.1007/s40804-021-00224-0>
- International Committee of the Red Cross (2020). Artificial intelligence and machine learning in armed conflict: A human-centred approach. *International Review of the Red Cross*, 102(913), 463–479. <https://doi.org/10.1007/s40804-021-00224-0>
- Klein, N. (2007). *The Shock Doctrine: The Rise of Disaster Capitalism*. New York: Henry Holt.
- Lee, J.-A., Hilty, R. M., & Liu, K.-C. (Eds.). (2021). *Artificial Intelligence and Intellectual Property*. Oxford: Oxford University Press.
- Oravec, J. A. (2022). Oravec, J. A. (2022). The emergence of “truth machines”? Artificial intelligence approaches to lie detection. *Ethics and Information Technology*, 24, 6. <https://doi.org/10.1007/s10676-022-09621-6>
- Rezaev, A. V. (2021). Twelve Theses on Artificial Intelligence and Artificial Sociality. *Monitoring of Public Opinion: Economic and Social Changes*, 1, 20–30. <https://doi.org/10.14515/monitoring.2021.1.1894>
- Shneiderman, B. (2021). Human-centered AI. *Issues in Science and Technology*, 37(2), 56–61.
- Ulnicane, I. (2022). Artificial Intelligence in the European Union: policy, ethics and regulation. In T. Hoerber, I. Cabras, G. Weber (Eds.). *Routledge Handbook of European Integrations* (pp. 254–269). London: Routledge. <https://doi.org/10.4324/9780429262081-19>
- Wolfe, A. (1993). *The Human Difference: Animals, Computers, and the Necessity of Social Science*. Berkeley: University of California Press.

Сведения об авторах



Резаев Андрей Владимирович – доктор философских наук, профессор, руководитель Международной исследовательской лаборатории ТАНДЕМ, Санкт-Петербургский государственный университет

Адрес: 191124, Российская Федерация, г. Санкт-Петербург, ул. Смольного, 1/3

E-mail: rezaev@hotmail.com

ORCID ID: <https://orcid.org/0000-0002-3918-835X>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=13004674100>

Web of Science Researcher ID:

<https://www.webofscience.com/wos/author/record/K-3472-2013>

Google Scholar ID: <https://scholar.google.ru/citations?user=Uzv39ccAAAAJ>

РИНЦ Author ID: https://elibrary.ru/author_items.asp?authorid=648768



Трегубова Наталья Дамировна – кандидат социологических наук, доцент кафедры сравнительной социологии, Санкт-Петербургский государственный университет

Адрес: 191124, Российская Федерация, г. Санкт-Петербург, ул. Смольного, 1/3

E-mail: n.tregubova@spbu.ru

ORCID ID: <https://orcid.org/0000-0003-3259-5566>

Scopus Author ID: <https://www.scopus.com/authid/detail.uri?authorId=56645016900>

Web of Science Researcher ID:

<https://www.webofscience.com/wos/author/record/K-3487-2013>

Google Scholar ID: <https://scholar.google.com/citations?user=8dhGr3gAAAAJ&hl>

РИНЦ Author ID: https://elibrary.ru/author_items.asp?authorid=832705

Вклад авторов

А. В. Резаев и Н. Д. Трегубова внесли равный вклад как в формулировку ключевых положений статьи, так и в подготовку рукописи к публикации.

Конфликт интересов

Авторы заявляют об отсутствии конфликта интересов.

Финансирование

Исследование выполнено при финансовой поддержке РФФИ и Министерства по науке и технологиям Тайваня в рамках научного проекта № 21-511-52002.

Тематические рубрики

Рубрика OECD: 5.05 / Law

Рубрика ASJC: 3308 / Law

Рубрика WoS: OM / Law

Рубрика ГРНТИ: 10.07.49 / Планирование и прогнозирование в праве

Специальность ВАК: 5.1.1 / Теоретико-исторические правовые науки

История статьи

Дата поступления – 3 марта 2023 г.

Дата одобрения после рецензирования – 4 мая 2023 г.

Дата принятия к опубликованию – 16 июня 2023 г.

Дата онлайн-размещения – 20 июня 2023 г.