Research article
DOI: https://doi.org/10.21202/jdtl.2023.42

Achieving Algorithmic Transparency and Managing Risks of Data Security when Making Decisions without Human Interference: Legal Approaches

Anna K. Zharova

Institute of State and Law of the Russian Academy of Sciences
Moscow, Russian Federation

Keywords

algorithmic transparency,
artificial intelligence,
confidentiality,
data protection,
data security,
digital technologies,
GDPR,
information technologies,
law,
personal data

Abstract

Objective: to compare modern approaches in law to the use of program codes and algorithms in decision-making that meet the principles of transparency and openness, as well as the increasingly stringent requirements for ensuring the security of personal and other big data obtained and processed algorithmically.

Methods: the main methods for researching the principle of transparency in algorithmic decision-making were formal-legal and comparative analysis of legal acts and international standards of information security, as well as the principles and legal constructions contained in them.

Results: it was determined that the development of information security standardization, inclusion in legal acts of requirements for the development of information technologies that comply with the principles of transparency and openness of applied algorithms will minimize the risks associated with the unlawful processing of users' big data and obtaining information about their privacy. Proposals were identified, related to the implementation of algorithmic transparency in the field of data processing legal regulation. Recommendations were formulated, based on which the legislator can solve the problem of ensuring the openness of the logic of information technology algorithms with regard to modern standards of information security.

Scientific novelty: it consists in the substantiation of new trends and relevant legal approaches, which allow revealing the logic of data processing by digital and information technologies, based on the characterization of European standards of the "privacy by design" concept in new digital and information technologies of decision-making and data protection, as well as on the new legal requirements for artificial intelligence systems, including the requirement to ensure algorithmic transparency, and criteria for personal data and users' big data processing. This said, data protection is understood as a system of legal, technical and organizational principles aimed at ensuring personal data confidentiality.

Practical significance: it is due to the need to study the best Russian and international practices in protecting the privacy of users of digital and information technologies, as well as the need for legislative provision of requirements for the use of algorithms that meet the principles of transparency and openness of personal data processing, taking into account the need to ensure confidentiality at all stages of the life cycle of their processing, which will ensure the continuity of security management.

For citation

Zharova, A. K. (2023). Achieving Algorithmic Transparency and Managing Risks of Data Security when Making Decisions without Human Interference: Legal Approaches. *Journal of Digital Technologies and Law, 1*(4), 973–993. https://doi.org/10.21202/jdtl.2023.42

Contents

Introduction

Information technology models are becoming increasingly complex, and more and more human-related data is being processed by them. For example, the technologies of the Internet of Things collect a large amount of data, which may contain various types, including users' big data. A user of information technologies (further – IT) is concerned about the impossibility to control the actions performed by information technologies, as well as the inability to understand the logic of the algorithms processing their data, what data are processed and what the final result of data analysis is. The desire to understand the criteria for analyzing processed data, as well as the need to ensure control over the list of such data, have led the legislator to the idea of including in legal acts the requirement to use algorithms that meet the principles of transparency and openness. In other words, it is necessary to reveal the logic of data processing by information technologies.

In computer science, the term "algorithmic transparency" is used to describe the transparency of processes occurring during the information technologies functioning. Due to the need to ensure the protection of user data, this term was borrowed by legal science.

In the Russian legislation, the term "algorithmic transparency" is used to describe the regulation of relations occurring during the application of artificial intelligence systems (further – AI)[1]. However, researchers believe that the term "algorithmic transparency" can be used to describe a wider range of relations that go beyond AI functioning (Kuteinikov et al., 2020; Gulemin, 2022).

Due to such polysemy, it is first of all necessary to study the "algorithmic transparency" concept, and then, based on the obtained results, to research the proposals related to the algorithmic transparency implementation in the field of legal regulation of data processing.

1. Notion of algorithmic transparency

The Concept of developing the regulation of relations in the sphere of artificial intelligence and robotics up to 2024[2] refers the problem of algorithmic transparency of artificial intelligence systems to the conceptual problem areas of regulation of relations in the said sphere[3].

"Algorithmic transparency" of the information model allows understanding the logic of the information model functioning implemented by AI with the given input data. However, it is essential that the AI algorithms complexity does not allow any AI algorithm to be described in such a way that its logic becomes understandable to an average person. The algorithmic

1 Decree of the President of the Russian Federation No. 490 of 10.11.2019 (with the National Strategy for the Development of Artificial Intelligence up to 2030). (2019). Collection of legislation of the Russian Federation, 41, Art. 5700.

2 Order of the Government of the Russian Federation No. 2129-r of 19.08.2020. (2020). Collection of legislation of the Russian Federation, 35, Art. 5593.

3 Ibid.

stages are the most easily perceived in a linear AI model and its mathematical interpretation. The most difficult for understanding is the logic of AI model functioning in a deep AI architecture[4].

In other words, algorithmic transparency is the explainability of the work of "AI and the process of it achieving results, non-discriminatory access of the users of products, created using artificial intelligence technologies, to information about the work of artificial intelligence algorithms applied in these products"[5]. Some researchers believe that technological developments in AI and algorithms have become an integral part of public administration (Feijóo et al., 2020; Carlsson & Rönnblom, 2022; Balasubramaniam et al., 2023; Green, 2022).

2. Can user data be protected through understanding the logic of data processing algorithm?

In 1999, L. Lessig was one of the first authors who, paying tribute to legal and social norms in providing legal regulation of relations arising in the ICT sphere, recognized the software code as an equal component in regulating information relations. A program code defines the ICT space architecture and allows achieving the best result in regulating relations arising in the information sphere (Lessig, 1999).

A program code formalizes the logic of algorithm operation. Proposals to provide algorithmic transparency in software are increasingly common. "Transparency of algorithms becomes a type of control, and transparency of algorithmic decision-making serves to ensure that unfair discriminations can be detected and challenged" (Talapina, 2020). However, there is an opposing point of view. For example, some scholars believe that the requirement of algorithmic transparency is aimed at IT developers' interaction with the authorities in order to normalize citizens' behavior (Wang, 2022).

In our opinion, the algorithmic transparency requirement will not bring the desired result, since it is not always possible even for a specialist to comprehend the logic of an AI algorithm. In this regard, unfair discriminations embedded in AI algorithms cannot always be detected and challenged.

Analyzing the tendency of algorithmic transparency implementation, we will attempt to present positions for and against the algorithm logic disclosure.

---

4 Koreshkova, T. (2020, December 29). Explainable artificial intelligence. GRFC Scientific-technical center. https://clck.ru/36h6w6

5 Decree of the President of the Russian Federation No. 490 of 10.11.2019 (with the National Strategy for the Development of Artificial Intelligence up to 2030). (2019). Collection of legislation of the Russian Federation, 41, Art. 5700.

An average person, hoping for the algorithm logic disclosure, believes it will allow them to understand that logic. However, such proposals, although not groundless, have their disadvantages. Firstly, for the majority of people incompetent in the field of programming and information technology development, the algorithm logic disclosure will not provide any information. Secondly, if the algorithm (for example, an AI) user understands its logic, they will be unable to change the algorithm because this will require revision of the entire mathematical toolkit embedded therein. Thirdly, the algorithm logic disclosure must not contradict the intellectual property law, since the intellectual rights to algorithms belong to its developers. Accordingly, the algorithm logic disclosure may occur in strictly limited cases only.

Other researchers suggest replacing the algorithm logic disclosure with insurance of risks associated with information security. For example, a user boarding an airplane and entrusting their life to a carrier does not study the logic of the airplane software beforehand. All risks are assumed by the carrier, an insurer and other persons responsible for passenger transportation (Ostroumov, 2015). Cannot we use the same legal framework of regulating relations in the case of data processing?

"Carriers" of user data are various information intermediaries like providers and operators of personal data processing. Taking into account the high probability that algorithm logic disclosure will actually give nothing to the user, would not it be more effective to apply an insurance system to relations in the ICT sphere, as in the case of air transportation? In this case, the risks of "loss" or unauthorized access to user data, as well as the liability of information intermediaries or personal data operators would be insured.

However, there are pitfalls in this case as well. For example, in the case of air transportation, all stages from the creation of an airplane to its flight are strictly regulated by legal and technical norms. This is not so with the creation of algorithms, information models and their use. Standardization of information technologies to ensure information security is voluntary. Only such information systems as critical information infrastructure of the Russian Federation and systems processing personal data are subject to mandatory standardization.

In this connection, drawing an analogy between ensuring IT user security through algorithmic transparency and air transportation security is only possible in case of using legal and technical norms fro strict regulation of IT creation and its use.

Trust in the field of information security occupies the minds of scholars from different countries (Bujold et al. 2022; Cui et al., 2022; Zhu et al., 2023). Trust is defined as a cultural value that may sometimes conflict with national AI policies (Li, 2022; Robinson, 2020; Xu et al., 2022).

3. Comparative analysis of the principles of algorithmic transparency

Experts from the Community for Fairness, Accountability, and Transparency in Machine Learning (FAT) define five principles of algorithmic transparency – Fairness, Auditability, Explainability, Responsibility and Accuracy[6]. As one may see, the explainability of AI logic is only ranked third in these proposals. This is probably due to the fact that algorithmic openness may not solve the issues of IT user security in all cases.

Researchers propose to supplement these five principles with the principle of making changes to the AI algorithm's operating logic in case of disagreement with its functioning (Malyshkin, 2019; Gordon et al., 2022). However, such proposals raise concerns, as in this case it is possible to violate intellectual property laws. Most companies are reluctant to disclose their algorithms and make them transparent, "citing potential gaming by users that may negatively affect the algorithm's predictive power" (Qiaochu et al., 2020; Stahl et al., 2022).

For our part, we would like to emphasize that the absence in the formulated five principles of "the possibility of changing the AI logic" is understandable. Such algorithms are developed by a team of programmers; changing the logic of operation of one part of an AI algorithm will make the mathematical model of the whole algorithm inoperable (Varsha, 2023; Lang & Shan, 2000; Akter et al., 2022). If the human brain could solve the problem of processing large, unstructured data, AI algorithms would be useless.

In the Russian Federation, in accordance with the National Strategy for the AI development up to 2030, the mandatory principles of AI development and use are: protection of human rights and freedoms, technological sovereignty, integrity of the innovation cycle, reasonable frugality, support for competition, security, and transparency[7].

Security is understood as "inadmissibility of the use of artificial intelligence with the objective of intentionally causing harm to citizens and legal entities, as well as prevention and minimization of risks of negative consequences of the use of artificial intelligence technologies"[8]. Transparency is defined as "explainability of the work of artificial intelligence and the process of achieving its results, non-discriminatory access of the users of products created with the use of artificial intelligence technologies to information about the algorithms of artificial intelligence used in these products" (p. 19)[9].

It should be emphasized that while the Community for Fairness, Accountability and Transparency in Machine Learning defines algorithmic transparency through five

6 Principles for Accountable Algorithms and a Social Impact Statement for Algorithms. https://clck.ru/36h7GL

7 Decree of the President of the Russian Federation No. 490 of 10.11.2019 (with the National Strategy for the Development of Artificial Intelligence up to 2030). (2019). Collection of legislation of the Russian Federation, 41, Art. 5700.

8 Ibid.

9 Ibid.

principles, the Russian legislation has included algorithmic transparency in the principles of AI development and use. The principle of non-discriminatory access of AI users to information about the AI algorithms applied overlaps with the principles formulated by the FAT Community.

The Russian legislation also does not contain the requirement to publish rules defining the basic algorithmic processing of user data, unlike the legislation of France. In accordance with the French Law "On the digital republic" of October 7, 2016, such rules must be published on a public authority website (Talapina, 2020).

In the European Union, the security of personal data of EU residents (Su et al., 2023), as well as the transparency of their processing by algorithms (Matheus et al., 2021; Kempeneer, 2021), including AI (Kempeneer et al., 2023; de Bruijn et al., 2022), is regulated by the General Data Protection Regulation (further – GDPR) (Stöger et al., 2021), which entered into force in 2018 (Mourby et al., 2021).

4. Elaboration and adoption of the General Data Protection Regulation

Although privacy protection is not directly related to personal data protection, the use of personal data allows for the identification of an individual and, consequently, the acquisition of information about his or her private life. Thus, well-protected personal data reduces the risks of obtaining information about a person's private life (Bolton et al., 2021; Leerssen, 2023). The need to combat various breaches of personal data and individual's privacy legislation and to minimize the consequences of such breaches led to the GDPR development and adoption (Willems et al., 2022; Custers & Heijne, 2022).

One of the first cases involving an unlawful acquisition of information about an individual's privacy through illegal access to a personal data base was the Uber case. The Uber's database was hacked in 2014 and 2016, allowing attackers to track the real-time location of every Uber user. In 2017, the Federal Trade Commission (FTC) accused Uber of failing to properly control employees' access to Uber user and driver databases, as well as breaching its information security system. Uber and the FTC subsequently signed an agreement, according to which Uber committed to conducting third-party audits for twenty years and implementing a privacy protection program[10].

In 2017, the FTC included additional provisions in the agreement obliging Uber to audit their system and submit reports to the FTC, as well as to disclose the fees and terms of agreements between Uber and the third parties that monitor vulnerabilities in Uber's software. Under the latest version of the agreement, Uber:

[10] Uber criminal complaint raises the stakes for breach response. https://clck.ru/37AXba

– may be subject to civil penalties if it fails to notify the FTC of incidents involving unauthorized access to Uber user and driver information;

– is prohibited from misrepresenting the system's level of information protection, the privacy, security, and integrity of personal information, and how it controls internal access to consumers' personal information;

– must implement a comprehensive privacy program and receive biennial, independent third-party assessments of the security of its information system for 20 years. Uber must submit these assessments to the FTC and confirm compliance with the adopted privacy program, while the latter must contain the security terms stipulated in its agreement with the FTC;

– must store user location information on the system, protected by a password and encryption;

– must provide annual training to employees responsible for handling personal information on its data protection and security practices and apply the latest security control techniques;

– must use the best data protection practices to protect drivers' personal information;

– must designate one or more employees to coordinate and oversee the security and privacy program, and conduct regular evaluations of the effectiveness of its internal controls and procedures related to the protection of personal and geographic location information of its employees and customers;

– is obliged to use multi-factor authentication before any employee can access sensitive customer personal information, as well as to use other strong data security practices[11].

In connection with the breaches that occurred in 2014 and 2016, Uber paid a $148 million fine[12]. On August 20, 2020, a criminal case was filed against its former chief security officer, charging him with obstruction of justice and allegedly attempting to cover up a data breach that occurred in 2016.

5. Implementation of the General Data Protection Regulation

GDPR defines the every person's right to protection of their personal data in accordance with part 1 of Article 16 of the Treaty on the Functioning of the European Union (TFEU)[13] and

11 Ibid.

12 Uber to Pay $148 Million Fine for Massive Data Breach That Exposed 57 Million Users' Personal Info. https://clck.ru/36h7RG

13 Treaty on the Functioning of the European Union [Rus., Eng.] (with the «List stipulated by Art. 38...", "Overseas countries and territories to which the provisions of Part Four of the Treaty apply...") (signed in Rome on 25.03.1957) (amended and restated as of 13.12.2007). SPS KonsultantPlyus.

part 1 of Article 8 of the Charter of Fundamental Rights of the European Union[14] (further – the Charter) [15], as well as the "right to privacy" (Article 7 of the Charter) [16].

GDPR requires from the companies processing personal data of EU residents or conducting their activities in the territory of EU states to comply not only with legal requirements, but also with organizational and technical requirements. This must be taken into account by developers at the stage of designing information technologies, and is called "privacy in design" (Article 3 of GDPR). The requirement to ensure privacy in the digital world through "privacy by design and by default" is approved by the European Data Protection Board (EDPB) in Guidelines 4/2019 on Article 25 Data Protection by Design and by Default[17].

The implementation of these requirements has raised many questions among companies about their implementation procedures. Therefore, the European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS) have provided clarifications. EDPB Statement 03/2021 "Privacy Regulation", adopted by the Privacy and Confidentiality Board for Electronic Communications Services, states that the proposed Regulation should not under any circumstances reduce the level of protection defined in the current Directive 2002/58/EC[18]. The Regulation should complement GDPR by providing additional strong privacy safeguards and protection for all types of electronic communications[19]. Directive 2002/58/EC covers the processing of personal data and privacy protection, including requirements to ensure the security of networks and services; confidentiality of communications; access to stored data; processing of traffic and location data; identification; publicly available subscriber directories, and prohibition of commercial communications (spam)[20]. EDPB pays special attention to the security of personal data processed by employers. It clearly defines the instances and conditions under which employers may access employees' personal data, as well as liability for excessive data collection through data analysis and processing technologies. It includes, for example, an employer using geolocation systems and technologies to continuously monitor an employee's movements and behavior.

---

14 Charter of Fundamental Rights of the European Union (2007/C 303/01) [Rus., Eng.] (with "Explanations..." (2007/C 303/02)) (adopted in Strasbourg on 12.12.2007). SPS KonsultantPlyus.

15 Ibid.

16 Charter of Fundamental Rights of the European Union. https://clck.ru/36h7Tn

17 Guidelines 4/2019 on Article 25 Data Protection by Design and by Default Version 2.0 Adopted on 20 October 2020. https://clck.ru/36h7Ug

18 E-privacy Directive 2009/136/EC. https://clck.ru/36h7VG

19 Statement 03/2021 on the ePrivacy Regulation Adopted on 9 March 2021. https://clck.ru/36h7WF

20 E-privacy Directive 2009/136/EC. https://clck.ru/36h7VG

Since EDPB and EDPS functions overlap, a Memorandum of Understanding[21] between EDPB and EDPS was adopted to differentiate their activities, according to which EDPB ensures the integrity of GDPR law enforcement practice, and EDPS ensures the common approaches of national supervisory authorities. At the same time, EDPB and EDPS may issue joint documents on personal data protection issues.

6. "Privacy by design" concept of information technologies

The "privacy by design" concept was developed long before the GDPR adoption. In 1995, the Data Protection Directive 95/46 / EC[22] included a provision that "to protect data security, technical and organizational measures shall be defined and adopted at the stage of planning the data processing system" (Article 46 of Directive 95/46 / EC).

On June 22, 2011, EDPS put forward the concept of changing the approach to the regulation of personal data protection and privacy[23] as a public opinion of this organization. Given the necessity and appropriateness of taking into account the requirements of personal data protection through privacy by design, the scholars proposed a change in the concept of personal data protection, outlining it in the following seven principles[24]:

1. Privacy by design measures should be preventive and take into account possible risks and threats, rather than being a reactive response to privacy breaches.

2. Privacy solutions for information systems should be implemented in the system at the design level, rather than being an option for the user.

3. Possible risks and threats should be considered at the technology design stage, as well as be stipulated in information security standards and take into account the data context. Personal data security methods should be continuously updated.

4. Confidentiality should be implemented at all stages of the personal data processing life cycle, as it will ensure the continuity of security management. The applied security standards should guarantee the confidentiality, integrity and availability of personal data throughout their life cycle, as well as the implementation of secure data destruction, encryption, access control and logging.

5. Privacy policies and procedures shall be monitored, evaluated and enforced; openness and transparency shall be maintained, in order to meet the principle of accountability and enable the trust of personal data subjects and counterparties,

21 Memorandum of Understanding. https://clck.ru/36h7ic

22 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. https://clck.ru/36h7jM

23 The History of the General Data Protection Regulation. https://clck.ru/36h7jw

24 The Seven Principles. https://goo.su/mn8ob7

as well as to harmonize business practices. Information on personal information management policies and practices, compliance and grievance mechanisms should be available to individuals.

6. There should be no compromising between security and functionality.

7. The rights and interests of personal data subjects should be the basis for privacy design.

These principles were among the first to take into account virtually all possible risks of a personal data processing breach. However, other concepts have also been proposed[25].

7. Notions of "personal data" and "privacy" in compliance with the EU legislation

In accordance with GDPR, personal data means any information relating to an identifiable natural person. In accordance with Article 4 of the GDPR, such data may include, for example, a reference to an identifier such as name, identification number, location data, online identifier, any factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

The GDPR text does not define the concept of "privacy", but refers to Directive 2002/58/EC of the European Parliament and of the Council of July 12, 2002, concerning personal data processing and privacy protection in the electronic communications sector (further – Directive 2002/58/EC)[26]. GDPR uses the concept of "personal data". According to EDPB[27], EDPS[28], and ENISA[29], in the context of projected privacy, the concepts of "personal data" and "privacy" should be considered as synonyms. In addition, EDPB, EDPS, and ENISA guidelines stipulate that for situations of low importance, no distinction should be made between personal data protection and privacy by design and by default.

---

25 Langheinrich, M. Privacy by Design – Principles of Privacy-Aware Ubiquitous Systems: Distributed Systems Group Institute of Information Systems, IFW Swiss Federal Institute of Technology, ETH Zurich 8092 Zurich, Switzerland. https://clck.ru/36h7qq

26 Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) (OJ L 201, 31.7.2002, p. 37).

27 EDPB is an independent body of the European Union, established and functioning on the basis of the GDPR. EDPB helps to ensure the harmonized application of the GDPR, for which purpose it has a number of powers stipulated by Art. 70 of GDPR. In particular, this body is authorized to issue guidelines, recommendations and best practices for the GDPR application.

28 EDPS is an independent EU body controlling the activities of the national supervisory authorities established in compliance with VGDPR section.

29 The European Union Agency for Cybersecurity (ENISA). https://clck.ru/N598K

However, this common understanding of personal data protection and privacy by design and by default is not accepted in all cases. The published EDPS opinion on Privacy by Design and Default[30] distinguishes between two concepts – privacy by design and data protection by design. The notion of privacy by design is used to refer to a system of technological measures aimed at ensuring privacy, developed in the course of international debates over the last few decades. This notion defines the legal regime of information, which consists in restricting access to it, and means "data protection by technological design" [31] (Zharova, 2020).

"Data protection by design" refers to a preliminary solution on data protection and privacy at the stage of technology design for all user actions[32] (Zharova, 2019).

Discussions over the extent to which these terms differ continue to this day. For example, the developers of explanations on the application of GDPR[33] write that there is still uncertainty about what privacy by design means and how it can be implemented. This problem arises due to the fact that, on the one hand, Directive 95/46/EC[34] is not fully implemented in some member states. On the other hand, according to the privacy by design principle contained in GDPR, data security guidelines require that organizational and technical measures should be adopted as early as at the stage of planning the information system. For example, the GDPR principle of integrity and confidentiality determines the need to protect data against unauthorized access or unlawful processing, as well as against accidental loss, destruction or damage[35]. However, the EU legislation leaves completely open the question of the protective measures taken by the parties responsible. For example, is anonymization of a person's name sufficient to fulfill the legislation requirements?

GDPR proposes using data encryption or anonymization as a possible privacy by design measure. However, this suggestion does not make it clear how this measure would further align with the GDPR's user authentication requirement and the technical implementation of the right to object.

---

30 Regulation (EC) No 45/2001 of the European Parliament and of the Council of 18 December 2000 on the protection of individuals with regard to the processing of personal data by the Community institutions and bodies and on the free movement of such data. https://clck.ru/36h7v3

31 GDPR Privacy by Design. https://clck.ru/36h7vZ

32 Data protection by design and default. https://goo.su/Hxoh2d

33 GDPR Privacy by Design. https://clck.ru/36h7vZ

34 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. https://clck.ru/36h7jM

35 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance) of 27 April 2016. https://clck.ru/34U2FN

As a result, the authors of the clarifications on GDPR application defined the term "privacy by design" as "data protection through technological design". They believe that "in data processing procedures, data protection is best adhered to when it is already integrated into the technology during its design phase".

Conclusion

The problems of personal data protection, control of users' big data processing principles, and protection of individual's privacy are only getting more acute every year. The need to ensure personal data protection and privacy of an individual as an IT user poses a challenge to the legislator to ensure the openness of the information technology algorithms' logic. To achieve this objective, GDPR stipulates the requirement to implement privacy by design in IT development, which was proposed back in 1995. The requirement to implement the principle of algorithmic transparency in AI systems was proposed much later – in 2019 by Russian lawmakers and in 2018 by foreign lawmakers.

Algorithms of data processing are becoming more and more complex. Hence, legislative proposals to reveal the logic of their functioning, for example, in AI systems, are made more and more often. However, one should understand that such proposals cannot be implemented for all algorithms. It is hardly possible to explain complex mathematical tools in simple words that will be understandable to every common person.

However, this does not mean that there is no solution to this complex technical and legal problem. We believe that the development of information security standards and the inclusion of requirements in legal acts on the IT development in compliance with standardization requirements will minimize the risks associated with the unlawful processing of users' big data and obtaining privacy information.

References

Akter, Sh., Dwivedi, Y. K., Sajib, Sh., Biswas, K., Bandara, R. J., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. *Journal of Business Research*, *144*, 201–216. https://doi.org/10.1016/j.jbusres.2022.01.083

Balasubramaniam, N., Kauppinen, M., Rannisto, A., Hiekkanen, K., & Kujala, S. (2023). Transparency and explainability of AI systems: From ethical guidelines to requirements. *Information and Software Technology*, *159*, 107197. https://doi.org/10.1016/j.infsof.2023.107197

Bolton, M., Raven, R., & Mintrom, M. (2021). Can AI transform public decision-making for sustainable development? An exploration of critical earth system governance questions. *Earth System Governance*, *9*, 100116. https://doi.org/10.1016/j.esg.2021.100116

Bujold, A., Parent-Rocheleau, X., & Gaudet, M.-C. (2022). Opacity behind the wheel: The relationship between transparency of algorithmic management, justice perception, and intention to quit among truck drivers. *Computers in Human Behavior Reports*, *8*, 100245. https://doi.org/10.1016/j.chbr.2022.100245

Carlsson, V., & Rönnblom, M. (2022). From politics to ethics: Transformations in EU policies on digital technology. *Technology in Society*, *71*, 102145. https://doi.org/10.1016/j.techsoc.2022.102145

Cui, M., Mariani, M. S., & Medo, M. (2022). Algorithmic bias amplification via temporal effects: The case of PageRank in evolving networks. *Communications in Nonlinear Science and Numerical Simulation*, *104*, 106029. https://doi.org/10.1016/j.cnsns.2021.106029

Custers, B., & Heijne, A.-S. (2022). The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice. *Computer Law & Security Review*, *46*, 105727. https://doi.org/10.1016/j.clsr.2022.105727

de Bruijn, H., Warnier, M., & Janssen, M. (2022). The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making. *Government Information Quarterly*, *39*(2), 101666. https://doi.org/10.1016/j.giq.2021.101666

Feijóo, C., Kwon, Y., Bauer, J. M., Bohlin, E., Howell, B., Jain, R., Potgieter, P., Vu, K., Whalley, J., & Xia, J. (2020). Harnessing artificial intelligence (AI) to increase wellbeing for all: The case for a new technology diplomacy. *Telecommunications Policy*, *44*(6), 101988. https://doi.org/10.1016/j.telpol.2020.101988

Gordon, G., Rieder, B., & Sileno, G. (2022). On mapping values in AI Governance. *Computer Law & Security Review*, *46*, 105712. https://doi.org/10.1016/j.clsr.2022.105712

Green, B. (2022). The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review*, *45*, 105681. https://doi.org/10.1016/j.clsr.2022.105681

Gulemin, A. (2022). Limits of big data processing for the purposes of obtaining information about a person: a legal aspect. In Elektronnoe prilozhenie k "Rossiiskomu yuridicheskomu zhurnalu", 6, 52–57. (In Russ.). http://doi.org/10.34076/22196838_2022_6_52

Kempeneer, Sh. (2021). A big data state of mind: Epistemological challenges to accountability and transparency in data-driven regulation. *Government Information Quarterly*, *38*(3), 101578. https://doi.org/10.1016/j.giq.2021.101578

Kempeneer, Sh., Pirannejad, A., & Wolswinkel, J. (2023). Open government data from a legal perspective: An AI-driven systematic literature review. *Government Information Quarterly*, 101823. https://doi.org/10.1016/j.giq.2023.101823

Kuteynikov, D. L., Izhaev, O. A., Zenin, S. S., & Lebedev, V. A. (2020). Algorithmic transparency and accountability: legal approaches to solving the "black box" problem. *Lex russica*, *73*(6), 139–148. (In Russ.). https://doi.org/10.17803/1729-5920.2020.163.6.139-148

Lang, H., & Shan, C. (2000). Bias phenomenon and compensation in multiple target tracking algorithms. *Mathematical and Computer Modelling*, *31*(8–9), 147–165. https://doi.org/10.1016/S0895-7177(00)00063-7

Leerssen, P. (2023). An end to shadow banning? Transparency rights in the Digital Services Act between content moderation and curation. *Computer Law & Security Review*, *48*, 105790. https://doi.org/10.1016/j.clsr.2023.105790

Lessig, L. (1999). *Code and other laws of cyberspace*. New York: Basic Books.

Li, Z. (2022). Affinity-based algorithmic pricing: A dilemma for EU data protection law. *Computer Law & Security Review*, *46*, 105705. https://doi.org/10.1016/j.clsr.2022.105705

Malyshkin, A. V. (2019). Integration of artificial intelligence into public life: some ethical and legal problems. *Vestnik of Saint Petersburg University. Law*, *10*(3), 444–460. (In Russ.). https://doi.org/10.21638/spbu14.2019.303

Matheus, R., Janssen, M., & Janowski, T. (2021). Design principles for creating digital transparency in government. *Government Information Quarterly*, *38*(1), 101550. https://doi.org/10.1016/j.giq.2020.101550

Mourby, M., Ó Cathaoir, K., & Bjerre Collin, C. (2021). Transparency of machine-learning in healthcare: The GDPR & European health law. *Computer Law & Security Review*, *43*, 105611. https://doi.org/10.1016/j.clsr.2021.105611

Ostroumov, N. N. (2015). Legal regime of international air transportation. Moscow: Statut. (In Russ.).

Qiaochu, W., Yan, H., Stefanus, J., & Param Vir, S. (2020, July 15). *Algorithmic Transparency with Strategic Users*. http://dx.doi.org/10.2139/ssrn.3652656

Robinson, S. C. (2020). Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI). *Technology in Society*, *63*, 101421. https://doi.org/10.1016/j.techsoc.2020.101421

Stahl, B. C., Rodrigues, R., Santiago, N., & Macnish, K. (2022). A European Agency for Artificial Intelligence: Protecting fundamental rights and ethical values. *Computer Law & Security Review*, *45*, 105661. https://doi.org/10.1016/j.clsr.2022.105661

Stefik, M. (1996). Letting loose the light: Igniting commerce in electronic publication. In M. Stefik (Ed.), *Internet dreams: Archetypes, myths, and metaphors* (pp. 219–253). Cambridge, MA: MIT Press.

Stöger, K., Schneeberger, D., Kieseberg, P., & Holzinger, A. (2021). Legal aspects of data cleansing in medical AI. *Computer Law & Security Review*, *42*, 105587. https://doi.org/10.1016/j.clsr.2021.105587

Su, Zh., Bentley, B. L., McDonnell, D., Cheshmehzangi, A., Ahmad, J., Šegalo, S., Pereira da Veiga, C., & Xiang, Yu-Tao. (2023). China's algorithmic regulations: Public-facing communication is needed. *Health Policy and Technology*, *12*(1), 100719. https://doi.org/10.1016/j.hlpt.2022.100719

Talapina, E. V. (2020). Algorithms and artificial intelligence in the human rights context. *Journal of Russian Law*, *10*, 25–39. (In Russ.). https://doi.org/10.12737/jrl.2020.118.

Varsha, P. S. (2023). How can we manage biases in artificial intelligence systems – A systematic literature review. *International Journal of Information Management Data Insights*, *3*(1), 100165. https://doi.org/10.1016/j.jjimei.2023.100165

Wang, H. (2022). Transparency as Manipulation? Uncovering the Disciplinary Power of Algorithmic Transparency. *Philosophy & Technology*, *35*, 69. https://doi.org/10.1007/s13347-022-00564-w

Willems, J., Schmidthuber, L., Vogel, D., Ebinger, F., & Vanderelst, D. (2022). Ethics of robotized public services: The role of robot design and its actions. *Government Information Quarterly*, *39*(2), 101683. https://doi.org/10.1016/j.giq.2022.101683

Xu, J., Xiao, Yu., Wang, W. Hu., Ning, Yu., Shenkman, E. A., Bian, J., & Wang, F. (2022). Algorithmic fairness in computational medicine. *eBioMedicine*, *84*, 104250. https://doi.org/10.1016/j.ebiom.2022.104250

Zharova, A. (2019). Ensuring the information security of information communication technology users in Russia. *International Journal of Cyber Criminology*, *13*(2), 255–269. DOI: https://doi.org/10.5281/zenodo.3698141

Zharova, A. (2020). The protect mobile user data in Russia. *International Journal of Electrical and Computer Engineering*, *10*(3), 3184–3192. https://doi.org/10.11591/ijece.v10i3.pp3184-3192

Zhu, H., Sallnäs Pysander, E.-L., & Söderberg, I.-L. (2023). Not transparent and incomprehensible: A qualitative user study of an AI-empowered financial advisory system. *Data and Information Management*, 100041. https://doi.org/10.1016/j.dim.2023.100041

Author information

Anna K. Zharova – Dr. Sci. (Law), Associate Professor, Senior Researcher, Institute of State and Law of the Russian Academy of Sciences
Address: 10 Znamenka Str., 420100 Moscow, Russian Federation
E-mail: anna_jarova@mail.ru
ORCID ID: https://orcid.org/0000-0002-2981-3369
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56964137900
WoS Researcher ID: https://www.webofscience.com/wos/author/record/H-4012-2015
Google Scholar ID: https://scholar.google.com/citations?user=g8ij3BsAAAAJ
RSCI Author ID: https://elibrary.ru/author_items.asp?authorid=151076

Conflict of interest

The author is an Editor-in-Chief of the Journal; the article has been reviewed on general terms.

Financial disclosure

The research was not sponsored.

Thematic rubrics

OECD: 5.05 / Law
PASJC: 3308 / Law
WoS: OM / Law

Article history

Date of receipt – May 22, 2023
Date of approval – August 21, 2023
Date of acceptance – November 30, 2023
Date of online placement – December 15, 2023

Научная статья
УДК 34:004:346.1:006.44:004.8
EDN: https://elibrary.ru/oppobg
DOI: https://doi.org/10.21202/jdtl.2023.42

Достижение алгоритмической прозрачности и управление рисками информационной безопасности при принятии решений без вмешательства человека: правовые подходы

Анна Константиновна Жарова

Институт государства и права Российской академии наук
г. Москва, Российская Федерация

Ключевые слова

GDPR,
алгоритмическая прозрачность,
защита данных,
информационная безопасность,
информационные технологии,
искусственный интеллект,
конфиденциальность,
персональные данные,
право,
цифровые технологии

Аннотация

Цель: сравнение современных подходов в праве к использованию в процессе принятия решений программных кодов и алгоритмов, отвечающих принципам прозрачности и открытости, а также возрастающим требованиям к обеспечению безопасности персональных и иных больших данных, полученных и обработанных алгоритмическим путем.

Методы: основными методами исследования принципа прозрачности алгоритмизированного принятия решений являлись формально-юридический и сравнительный анализ правовых актов и международных стандартов информационной безопасности, содержащихся в них принципов и правовых конструкций.

Результаты: определено, что развитие области стандартизации информационной безопасности, включение в правовые акты требований о разработке информационных технологий, соответствующих принципам прозрачности и открытости применяемых алгоритмов, позволит минимизировать риски, связанные с неправомерными обработкой больших пользовательских данных и получением информации об их частной жизни; выявлены связанные с реализацией алгоритмической прозрачности предложения в области правового регулирования обработки данных; сформулированы рекомендации, с опорой на которые законодатель может решать задачу обеспечения открытости логики работы алгоритмов информационных технологий с учетом современных стандартов информационной безопасности.

Научная новизна: состоит в обосновании новых тенденций и формируемых в соответствии с ними правовых подходов, позволяющих раскрыть логику обработки данных цифровыми и информационными технологиями, на основе характеристики общеевропейских стандартов концепции конфиденциальности при проектировании новых цифровых и информационных технологий принятия решений и защиты данных, новых правовых требований, предъявляемых к системам искусственного интеллекта, включая требование об обеспечении алгоритмической прозрачности, критериев обработки персональных данных, а также больших пользовательских данных. При этом защита данных рассматривается как система правовых, технических и организационных принципов, направленная на обеспечение конфиденциальности персональных данных.

Практическая значимость: обусловлена необходимостью изучения передового отечественного и международного опыта защиты частной жизни пользователей цифровых и информационных технологий, а также законодательного обеспечения требований об использовании алгоритмов, отвечающих принципам прозрачности и открытости обработки персональных данных с учетом необходимости обеспечения конфиденциальности на всех этапах жизненного цикла их обработки, что позволит обеспечить непрерывность управления безопасностью.

Для цитирования

Жарова, А. К. (2023). Достижение алгоритмической прозрачности и управление рисками информационной безопасности при принятии решений без вмешательства человека: правовые подходы. *Journal of Digital Technologies and Law*, *1*(4), 973–993. https://doi.org/10.21202/jdtl.2023.42

Список литературы

Гулемин, А. Н. (2022). Пределы обработки больших объемов данных для целей получения информации о человеке: правовой аспект. *Электронное приложение к Российскому юридическому журналу*, *6*, 52–57. http://doi.org/10.34076/22196838_2022_6_52

Кутейников, Д. Л., Ижаев, О. А., Зенин, С. С., Лебедев, В. А. (2020). Алгоритмическая прозрачность и подотчетность: правовые подходы к разрешению проблемы «черного ящика». *Lex russica (Русский закон)*, *73*(6), 146. https://doi.org/10.17803/1729-5920.2020.163.6.139-148

Малышкин, А. В. (2019). Интегрирование искусственного интеллекта в общественную жизнь: некоторые этические и правовые проблемы. Вестник Санкт-Петербургского университета. *Право*, *10*(3), 444–460. https://doi.org/10.21638/spbu14.2019.303

Остроумов, Н. Н. (2015). *Правовой режим международных воздушных перевозок*. Москва: Статут. https://elibrary.ru/ulcfpl

Талапина, Э. В. (2020). Алгоритмы и искусственный интеллект сквозь призму прав человека. *Журнал российского права*, *10*, 25–39. https://doi.org/10.12737/jrl.2020.118

Akter, Sh., Dwivedi, Y. K., Sajib, Sh., Biswas, K., Bandara, R. J., & Michael, K. (2022). Algorithmic bias in machine learning-based marketing models. *Journal of Business Research*, *144*, 201–216. https://doi.org/10.1016/j.jbusres.2022.01.083

Balasubramaniam, N., Kauppinen, M., Rannisto, A., Hiekkanen, K., & Kujala, S. (2023). Transparency and explainability of AI systems: From ethical guidelines to requirements. *Information and Software Technology*, *159*, 107197. https://doi.org/10.1016/j.infsof.2023.107197

Bolton, M., Raven, R., & Mintrom, M. (2021). Can AI transform public decision-making for sustainable development?

An exploration of critical earth system governance questions. *Earth System Governance*, *9*, 100116. https://doi.org/10.1016/j.esg.2021.100116

Bujold, A., Parent-Rocheleau, X., & Gaudet, M.-C. (2022). Opacity behind the wheel: The relationship between transparency of algorithmic management, justice perception, and intention to quit among truck drivers. *Computers in Human Behavior Reports*, *8*, 100245. https://doi.org/10.1016/j.chbr.2022.100245

Carlsson, V., & Rönnblom, M. (2022). From politics to ethics: Transformations in EU policies on digital technology. *Technology in Society*, *71*, 102145. https://doi.org/10.1016/j.techsoc.2022.102145

Cui, M., Mariani, M. S., & Medo, M. (2022). Algorithmic bias amplification via temporal effects: The case of PageRank in evolving networks. *Communications in Nonlinear Science and Numerical Simulation*, *104*, 106029. https://doi.org/10.1016/j.cnsns.2021.106029

Custers, B., & Heijne, A.-S. (2022). The right of access in automated decision-making: The scope of article 15(1)(h) GDPR in theory and practice. *Computer Law & Security Review*, *46*, 105727. https://doi.org/10.1016/j.clsr.2022.105727

de Bruijn, H., Warnier, M., & Janssen, M. (2022). The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making. *Government Information Quarterly*, *39*(2), 101666. https://doi.org/10.1016/j.giq.2021.101666

Feijóo, C., Kwon, Y., Bauer, J. M., Bohlin, E., Howell, B., Jain, R., Potgieter, P., Vu, K., Whalley, J., & Xia, J. (2020). Harnessing artificial intelligence (AI) to increase wellbeing for all: The case for a new technology diplomacy. *Telecommunications Policy*, *44*(6), 101988. https://doi.org/10.1016/j.telpol.2020.101988

Gordon, G., Rieder, B., & Sileno, G. (2022). On mapping values in AI Governance. *Computer Law & Security Review*, *46*, 105712. https://doi.org/10.1016/j.clsr.2022.105712

Green, B. (2022). The flaws of policies requiring human oversight of government algorithms. *Computer Law & Security Review*, *45*, 105681. https://doi.org/10.1016/j.clsr.2022.105681

Kempeneer, Sh. (2021). A big data state of mind: Epistemological challenges to accountability and transparency in data-driven regulation. *Government Information Quarterly*, *38*(3), 101578. https://doi.org/10.1016/j.giq.2021.101578

Kempeneer, Sh., Pirannejad, A., & Wolswinkel, J. (2023). Open government data from a legal perspective: An AI-driven systematic literature review. *Government Information Quarterly*, 101823. https://doi.org/10.1016/j.giq.2023.101823

Lang, H., & Shan, C. (2000). Bias phenomenon and compensation in multiple target tracking algorithms. *Mathematical and Computer Modelling*, *31*(8–9), 147–165. https://doi.org/10.1016/S0895-7177(00)00063-7

Leerssen, P. (2023). An end to shadow banning? Transparency rights in the Digital Services Act between content moderation and curation. *Computer Law & Security Review*, *48*, 105790. https://doi.org/10.1016/j.clsr.2023.105790

Lessig, L. (1999). *Code and other laws of cyberspace*. New York: Basic Books.

Li, Z. (2022). Affinity-based algorithmic pricing: A dilemma for EU data protection law. *Computer Law & Security Review*, *46*, 105705. https://doi.org/10.1016/j.clsr.2022.105705

Matheus, R., Janssen, M., & Janowski, T. (2021). Design principles for creating digital transparency in government. *Government Information Quarterly*, *38*(1), 101550. https://doi.org/10.1016/j.giq.2020.101550

Mourby, M., Ó Cathaoir, K., & Bjerre Collin, C. (2021). Transparency of machine-learning in healthcare: The GDPR & European health law. *Computer Law & Security Review*, *43*, 105611. https://doi.org/10.1016/j.clsr.2021.105611

Qiaochu, W., Yan, H., Stefanus, J., & Param Vir, S. (2020, July 15). *Algorithmic Transparency with Strategic Users*. http://dx.doi.org/10.2139/ssrn.3652656

Robinson, S. C. (2020). Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI). *Technology in Society*, *63*, 101421. https://doi.org/10.1016/j.techsoc.2020.101421

Stahl, B. C., Rodrigues, R., Santiago, N., & Macnish, K. (2022). A European Agency for Artificial Intelligence: Protecting fundamental rights and ethical values. *Computer Law & Security Review*, *45*, 105661. https://doi.org/10.1016/j.clsr.2022.105661

Stefik, M. (1996). Letting loose the light: Igniting commerce in electronic publication. In M. Stefik (Ed.), *Internet dreams: Archetypes, myths, and metaphors* (pp. 219–253). Cambridge, MA: MIT Press.

Stöger, K., Schneeberger, D., Kieseberg, P., & Holzinger, A. (2021). Legal aspects of data cleansing in medical AI. *Computer Law & Security Review*, *42*, 105587. https://doi.org/10.1016/j.clsr.2021.105587

Su, Zh., Bentley, B. L., McDonnell, D., Cheshmehzangi, A., Ahmad, J., Šegalo, S., Pereira da Veiga, C., & Xiang, Yu-Tao. (2023). China's algorithmic regulations: Public-facing communication is needed. *Health Policy and*

Technology, *12*(1), 100719. https://doi.org/10.1016/j.hlpt.2022.100719

Varsha, P. S. (2023). How can we manage biases in artificial intelligence systems – A systematic literature review. *International Journal of Information Management Data Insights*, *3*(1), 100165. https://doi.org/10.1016/j.jjimei.2023.100165

Wang, H. (2022). Transparency as Manipulation? Uncovering the Disciplinary Power of Algorithmic Transparency. *Philosophy & Technology*, *35*, 69. https://doi.org/10.1007/s13347-022-00564-w

Willems, J., Schmidthuber, L., Vogel, D., Ebinger, F., & Vanderelst, D. (2022). Ethics of robotized public services: The role of robot design and its actions. *Government Information Quarterly*, *39*(2), 101683. https://doi.org/10.1016/j.giq.2022.101683

Xu, J., Xiao, Yu., Wang, W. Hu., Ning, Yu., Shenkman, E. A., Bian, J., & Wang, F. (2022). Algorithmic fairness in computational medicine. *eBioMedicine*, *84*, 104250. https://doi.org/10.1016/j.ebiom.2022.104250

Zharova, A. (2019). Ensuring the information security of information communication technology users in Russia. *International Journal of Cyber Criminology*, *13*(2), 255–269. EDN: https://elibrary.ru/ltmesv. DOI: https://doi.org/10.5281/zenodo.3698141

Zharova, A. (2020). The protect mobile user data in Russia. *International Journal of Electrical and Computer Engineering*, *10*(3), 3184–3192. EDN: https://www.elibrary.ru/juzboh. DOI: https://doi.org/10.11591/ijece.v10i3.pp3184-3192

Zhu, H., Sallnäs Pysander, E.-L., & Söderberg, I.-L. (2023). Not transparent and incomprehensible: A qualitative user study of an AI-empowered financial advisory system. *Data and Information Management*, 100041. https://doi.org/10.1016/j.dim.2023.100041

Сведения об авторе

Жарова Анна Константиновна – доктор юридических наук, доцент, старший научный сотрудник, Институт государства и права Российской академии наук
Адрес: 420100, Российская Федерация, г. Москва, ул. Знаменка, 10
E-mail: anna_jarova@mail.ru
ORCID ID: https://orcid.org/0000-0002-2981-3369
Scopus Author ID: https://www.scopus.com/authid/detail.uri?authorId=56964137900
WoS Researcher ID: https://www.webofscience.com/wos/author/record/H-4012-2015
Google Scholar ID: https://scholar.google.com/citations?user=g8ij3BsAAAAJ
РИНЦ Author ID: https://elibrary.ru/author_items.asp?authorid=151076

Конфликт интересов

Автор является главным редактором журнала, статья прошла рецензирование на общих основаниях.

Финансирование

Исследование не имело спонсорской поддержки.

Тематические рубрики

Рубрика OECD: 5.05 / Law
Рубрика ASJC: 3308 / Law
Рубрика WoS: OM / Law
Рубрика ГРНТИ: 10.19.61 / Правовое регулирование информационной безопасности
Специальность ВАК: 5.1.2 / Публично-правовые (государственно-правовые) науки

История статьи

Дата поступления – 22 мая 2023 г.
Дата одобрения после рецензирования – 21 августа 2023 г.
Дата принятия к опубликованию – 30 ноября 2023 г.
Дата онлайн-размещения – 15 декабря 2023 г.