Legal Issues in the **DIGITAL AGE**

2/2025

ISSUED QUARTERLY VOLUME 6

Publisher

National Research University Higher School of Economics

ISSN 2713-2749

The journal is registered in the Federal Service of Supervision of Communications, Information Technology and Mass Media. Certification of registration of mass media серия Эл № ФС77-83367

Address:
3 Bolshoy Triohsviatitelsky Per., Moscow 109028, Russia
Tel.:
+7 (495) 220-99-87
e-mail lida@hse.ru

Designer
Andrei Pavlov
Pre-press
Natalya Puzanova

Legal Issues in the **DIGITAL AGE**

Legal Issues in the **DIGITAL AGE**

ISSUED QUARTERLY

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

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

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

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

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

Publication in the journal is free of charge.

All materials are available for free download.

Legal Issues in the Digital Age. 2025. Vol. 6, no. 2.
Вопросы права в цифровую эпоху. 2025. Том 6. № 2.

Digital Platforms and Law

Research article
JEL: K43
UDK: 349
DOI:10.17323/2713-2749.2025.2.4.49

In Search of the Regulatory Optimum for Digital Platforms: A Comparative Analysis

Alexey S. Koshel[1], Yaroslav I. Kuzminov[2], Ekaterina V. Kruchinskaia[3], Bogdan V. Lesiv[4]

[1, 2, 3, 4] Higher School of Economics (HSE University), 20 Myasnitskaya Str., Moscow 101000, Russia,

[1] koshel@hse.ru, ORCID: 0000-0002-4517-8326

[2] kouzminov@hse.ru, ORCID:0000-0003-4598-0631

[3] ekruchinskaya@hse.ru, ORCID:0000-0003-4778-3287

[4] blesiv@hse.ru, ORCID: 0000-0003-3085-3983

Abstract

The rapid growth of digital platforms and ecosystems has become a significant economic phenomenon on a global scale. This growth is due to the ability of these platforms to provide additional and flexible opportunities that are mutually beneficial for sellers, buyers, and platform workers. Because of it the activities of digital platforms have a positive impact on the overall gross domestic product of countries worldwide. The focus of the study is made on the regulatory frameworks for digital platforms both in Russia and around the world, including the rights and obligations of owners, operators, and users resulting from their participation in market transactions. The study does not include digital platforms used in the public sector or social media and messaging services. Scholar methods: comparative legal, formal logic, formal doctrinal, historical legal, as well as analytical, synthetic, and hermeneutical methods are systematically and integrally applied in the research. Based on the sources material, a hypothesis has been proposed regarding three stages of platform regulation growth globally and in Russia. Upon the results of an analysis of the three-stage evolutionary process of legal regulation for e-commerce, it has been

found that there is commonly inconsistent impact of various branches of law on the different areas of social relations or different types of platforms. Among this inconsistency are legal gaps and conflicts of legal rules, which make benefits for stakeholders spontaneous rather than the result of systematic interaction within the regulatory framework. Authors of the article identify a major source of legal uncertainty: the absence of standardized terms and harmonized regulatory principles that account for the unique nature of cross-industry digital economy. Lessons from global jurisdictions and three stages of e-commerce regulation reveal that, in its latest phase, the platform economy necessitates system of tailored legal definitions to manage its multifaceted activities. The survey proposes such conceptual structures that may be employed in Russian legal system. They reflect the multidimensional nuances of civil, tax, competition, information, and administrative laws. Additionally, a balanced scheme of general principles has been developed that would ensure the transparent interaction of digital platforms with society, the state, and economic entities.

Keywords

e-commerce; digital platforms; platform economy; Big Tech; platforms' intermediary role; legal glossary; *primum non nocere.*

For citation: Koshel A.S., Kuzminov Ya.I., Kruchinskaya E.V., Lesiv B.V. (2025) In Search of the Regulatory Optimum for Digital Platforms: A Comparative Analysis. *Legal Issues in the Digital Age*, vol. 6, no. 2, pp. 4–49. DOI:10.17323/ 2713-2749.2025.2.4.49

Introduction

History offers numerous examples that support the thesis that the development of socio-economic formations often outpaces the development of the legal institutions that regulate them. For instance, during the period of active capitalist development in Europe, there was a discrepancy between the needs of the burgeoning market economy and the archaic feudal law that governed property and trade relations. Currently, digital platforms represent one of the most striking examples of this kind on a similar scale, as they have already had a significant impact on the structure and principles of trade, introducing the transnational principle into the ways where goods and services are acquired.

Like many other economic and technological innovations before them, digital platforms have been emerging within legal regimes that were developed earlier without considering their specifics. Therefore, the functioning of such economically significant institutions outside a properly adjusted legal framework inevitably leads to conflicts.

For the optimal development of an institution that has a direct and substantial impact on everyday market relations, it is necessary to maintain a sort of rational alignment between the objectives of legal regulation and the goals pursued by the corresponding innovative economic institutions. Such alignment will optimize the impact of legal norms on economic processes, entrenching predictability and stability, while maintaining the potential for innovative development.

The lack of a congruent economic and legal model to regulate digital platforms, taking into account their structural and organizational features, may increase the risks of loss in terms of both stability and progress. This principle aligns with the tenets of rational choice theory and new institutionalism [North D.C., 1990]; [Haggard S., Tiede L., 2011).

Incongruent economic and legal models in the case of a digital platform may lead to market failures due to futile regulatory measures. Excessive regulatory stringency, disproportionate to the potential benefit, may hinder the utilization of useful properties of a product or service, at the same time the objectives of law (stability, security, etc.) may be achieved through less invasive means. Over regulation may lead to an artificial increase in prices or a decrease in the accessibility of goods.

Conversely, the absence of proper regulation may produce harm to consumers as actors, whole market and, finally, society. *Non liquet*[1] situations create conditions for abuse or opportunism on the part of digital platform owners and operators, whose large-scale actions may threaten national security. Thus, what matters is not just the presence of regulation *per se*, but its adequacy and relevance to the innovative aspect of the platform economy.

In view of what has been said, we are able define the current state of the legal regulation of digital platforms. Firstly, for objective reasons, the activity of digital platforms, one of the most prominent phenomena of the recent technological advancement, did not initially have comprehensive legal regulation, although such activity certainly requires regulation due to the risks of opportunism, monopolization, and market control by major business players. Secondly, such regulation must be congruent with the goals of the corresponding economic sector so as not to become an artificial inhibiting factor in its development. Thirdly, the legal regulation should be in terms of *lex specialis,* clear and specific, since setting an ex-

[1] Literally — it is not clear (Latin). A legal lacuna or absence of clear legal regulation.

cessively broad "normative framework" will not contribute to achieving the stated goals, at the same time leading to either a broad interpretation of the norms or giving rise to circumventive schemes.

As a methodological guide to work on such a complex interdisciplinary problem the principle of *primum non nocere* is widely and justifiably applied, which means that the optimum regulatory framework begins where the effective development of the economic institution continues [Lofstedt R.E., 2003: 36–43]; [Rylova M.A., 2014: 30–42), the effectiveness meaning the interests of both the institution and the customers are taken into account. The main task of the work is answering the question of how to design harmonious and balanced law-making initiatives for the economic phenomenon under consideration. The answer requires application of inductive method and critical analysis of the legal experience in the field in foreign jurisdictions.

In contrast to traditional views that primarily consider digital platforms as tools (whether seen as complex software systems, innovative business models, or technological infrastructure), the article proposes a fundamentally different conceptualization. Digital platforms are viewed as an innovative form of market organization and economic activity with an intermediary function.

This approach implies that platforms do not simply facilitate economic processes, but help to form new market structures, redefine relationships between participants while creating entirely new types of value. For example, marketplaces are not average online stores; they represent complex digital products — market mechanisms in which millions of sellers and buyers interact with each other, forming a global market accessible to everyone. At the same time, the term "marketplace" has not yet been legally enshrined in any country in the world.

Thus, the main research problem is the absence of established *lex specialis* legal norms regarding the platform economy in Russia, and the study is focused on the legal regulation of the platform economy; the authors consider a number of specific types (kinds) of digital platforms operating in the field of electronic commerce (e-commerce).

The relevance of regulatory issues in the field of the digital platform and ecosystem market is explained by several facts. In many countries the platform economy serves as one of the leading drivers of economic development and growth, creating new trade flows, accelerating the inflow of resources, and stimulating entrepreneurial activity [Paun C., Ivascu C. et al., 2024].

On the one hand, platforms lower entry barriers for new market participants. On the other hand, they unite various economic entities, from small businesses to large enterprises, unifying the rules of competition [Hossain M.B. et al., 2022: 162–178]. The spread of platform economic forms occurs more discretely than evolutionarily, which raises questions that require clarification within the existing legal system.

The very activity of digital platforms does not contradict legal norms and develops as a legitimate way of conducting trade within the current framework of civil law. But, as they grow, platforms develop their own complex economic structures functioning according to their own specific internal principles and producing noticeable effect on the market. In this regard, *lex specialis* regulation becomes a necessary step to protect a balance of interests and to prevent abuses caused by the dominant position of the platform.

Though an innovative market form with the function of intermediary, the phenomenon of digital platforms mirrors the success of electronic commerce in the 2010s in its form of dissemination of goods. Even then it was emphasized that online commerce (Kenney M. et al., 2016: 61–69) is a significant driver of not only competition but also innovation: as companies competed for consumer attention, breakthrough technologies for recommendation systems and user-friendly interfaces were developed (Deldjoo Y. et al., 2024: 69–108). Online commerce experienced significant growth during the COVID-19 pandemic. According to Statista data[2], the dynamics of revenue from retail trade in e-commerce show a pronounced peak in growth during lockdowns, followed by a slowdown as restrictive measures were lifted.

In the markets of the People's Republic of China (hereinafter — PRC), the European Union (hereinafter — EU), the Russian Federation (hereinafter — Russia, RF), and the Republic of Korea (hereinafter — RK), the ratio of online and offline sales is approaching equilibrium from 2026 onwards. In contrast, the United States of America (hereinafter–USA) market demonstrates sustained growth in e-commerce and the platform economy.

Overall, there is a global trend towards an increase in the share of online commerce, albeit with varying intensity in different regions. For example, in China the share of online commerce will grow from 12.3% in

[2] Available at: URL: https://www.statista.com/markets/413/e-commerce/ (accessed: 10.04.2025)

2017 to almost 40% by 2029, and in the United States — from 17.7% to more than 43% over the same period. In Russia and the EU countries, this growth is more moderate, but even there the online segment shows a steady rise (from 4.1% to 8.8% in Russia and from 8.4% to 17.2% in the EU). These differences may be due to several factors, including the level of infrastructure development, consumer behavior patterns, and cultural traditions.

Given the multifaceted nature of the issue, further research requires addressing a key question: how is platform economy regulation understood and what are the boundaries of the applicability of various measures aimed at achieving common economic well-being? The relevance of the question is due to the lack of clear definitions of digital platforms and their characteristics in contemporary scholarly publications [Heimburg V., Wiesche M., 2023: 72–85].

The lack of systematization of regulatory approaches and the absence of uniform criteria for assessing their effectiveness in the USA, PRC, EU, and RK, caused by legal and technical difficulties in distinguishing participants in market processes, hinders the development of regulatory acts and limits the use of foreign legal experience, impeding the formation of a consistent and predictable legal environment. In this regard, it seems that the principle of *primum non nocere* should underlie the regulation of platforms, minimizing unforeseen negative consequences, and ensuring economic growth, and the development of a high-quality legal glossary is a necessary condition for balanced regulation of the platform economy.

To maintain the correctness and validity of legal terminology, it is necessary to study the experience of jurisdictions where the platform economy has become widespread and regulated. The absence of a specifically adjusted regulatory framework (comprising both reliable legal definitions and principles providing a solid normative ground for subsequent legal rules) may lead to abuse of a dominant position, concentration of market power, and unfair competition on behalf of Big-Tech. Alternately, excessively strict regulation can reduce market share and the quality of platform functioning. The lacuna that this research aims to fill is the deficiency of a systematic analysis of both foreign and Russian experience that may be valuable in elaboration of adequate definitions, and subsequently, approaches to the regulation of the platform economy.

1. Stages of Development of E-Commerce Regulation in Foreign Jurisdictions

In developed countries the transformation of legal regulation of the platform economy represents an evolutionary process of adapting legislation to the dynamic development of e-commerce. In this case, e-commerce is understood as a phenomenon preceding the emergence of the platform economy. At the same time, the development of the economic institution of platforms itself in many jurisdictions outpaces the formation of unified approaches to legal regulation [Shelepov A.V., Kolmar O.I., 2024: 110–126]. This mismatch is expressed in the heterogeneity of regulatory strategies, due to differences in the pace of digital service development and in the priorities of national policy. A comprehensive study of this phenomenon is a separate analytical task, requiring the identification of an optimal balance between stimulating competition, protecting consumer rights, promoting innovation, and ensuring national economic security through the development of management models adapted to local economic conditions [Lafuente E. et al., 2024: 36–43].

Despite differences in national approaches, three stages in the development of e-commerce regulation can be identified; they reflect the general patterns of increasing complexity of legal constructions requiring an appropriate response from the legislator. A detailed analysis of these stages is necessary to predict further evolution of legal norms, to prepare in time the legislative framework for new challenges of the platform economy.

The first stage has started in the early 21st century with the formation of a primary legal base for typical trading operations on the Internet. The key task was to protect the consumer as the most vulnerable party in retail trade relations, considering the specifics of online transactions. During this period, norms were developed to ensure the transparency of transaction terms, protection against unscrupulous sellers, and dispute resolution mechanisms. Legislative bodies sought to protect consumer rights and promote the development of new digital forms of economic activity.

An example is the EU E-Commerce Directive (Directive 2000/31/EC), aimed at preserving legal certainty, protecting consumer rights, and creating a legal framework for the free movement of goods and information. Later, the EU Consumer Rights Directive (Directive 2011/83/EU)

was passed, representing a comprehensive regulatory act aimed at protecting consumer rights, including in the digital environment.

Similar trends were observed in the RK, where the Act on Promotion of Information and Communication Network Utilization and Information Protection and the Act on Consumer Protection in Electronic Commerce were approved. The purpose of the latter was to strengthen the protection of the rights and interests of consumers by establishing fair trade rules and promoting sustainable development of the national economy. Similar norms ensuring consumer protection were adopted in the USA (Restore Online Shoppers' Confidence Act (ROSCA; 2010) and the PRC (E-Commerce Law of the People's Republic of China of 2018).

It is important the initial measures taken by the Korean and Chinese authorities, despite their apparent «belatedness,» were a response to the same challenges that the mentioned above acts of other states dealt with and therefore these measures in South Korea and China are consistent with the concept of the first stage of e-commerce regulation.

At the first stage the regulatory approach reflected the recognition of the relationship between the development of digital platforms and e-commerce and ensuring consumer confidence, as well as providing them with guarantees of protection against potential risks associated with remote transactions. Legislators sought a balance between innovation and security, introducing norms governing issues of transparency, consumer rights to return goods and services, as well as providing dispute resolution mechanisms.

The legal frameworks developed for early e-commerce proved insufficient for regulating evolving digital platforms. The initial emphasis on regulating e-commerce, specifically focused on «seller-buyer» type of transactions, proved insufficient to cover the entire spectrum of interactions and risks arising within the platform economy. The need to adapt legal regulation was due to the increasing complexity of the structure of new digital forms of economic activity, including the emergence of platform ecosystems, the use of artificial intelligence and algorithmic trading, the application of technologies for tracking personal preferences, and the integration of social networks into sales channels.

The transition to the next stage of development of legal regulation was due to the realization of the inadequacy of existing norms to ensure comprehensive protection of consumer rights; there arose the need to expand the scope of regulation to cover new types of activities and

their emerging risks, particularly those related to the quality control of products distributed through digital platforms. The measures taken at subsequent stages only partially filled the existing gaps, as will be shown below.

The second stage in 2010s is characterized by the expansion of the scope of legal regulation to cover issues of consumer personal data protection, as a response to the growth of online commerce and the increase in the volume of user data. The goal was to create legal mechanisms that guarantee the confidentiality and protection of users personal information.

In the EU, the General Data Protection Regulation 2016/679 (GDPR) was adopted in 2016, establishing uniform standards for the processing of personal data. In RK, the Personal Information Protection Act was passed in 2011, and in PRC, the Personal Information Protection Law came into force in 2021.

In the USA, the regulation of the protection of personal data of digital platform users is mainly carried out at the state level. For example, in California, where this issue is regulated in terms of protecting the constitutional right to privacy (California Online Privacy Protection Act of 2003), supplemented in 2013, California Consumer Privacy Act (2018), and California Privacy Rights Act (2020), approved by local referendum, that strengthens the regulation of the previous act. These laws establish increased guarantees and stricter requirements for the processing of personal data, akin to the approaches laid down in the GDPR (EU).

Thus, the second stage is characterized by the recognition of the importance of personal data protection as the influence of digital platforms grew and by the adoption of relevant legislation. However, the approach to regulation based on understanding a digital platform as a tool for transactions, rather than as an independent intermediary, creates risks of incorrect law enforcement. Insufficient attention to the role of the platform as an intermediary operating with personal data, as well as the lack of continuity with the previous stage of regulation and the absence of a comprehensive approach potentially reduces the force of personal data protection and creates transaction costs for consumers and entrepreneurs.

The third, current stage of legal regulation development of the platform economy that started in the early 2020s represents a response to the increasing complexity of the platform economy structure and strengthening the dominant positions of large digital platforms. This stage logi-

cally continues answering the questions that arose at the previous stage. The increase in computing power and the volume of processed data and the spread of intelligent algorithms have allowed platforms to accumulate not only users personal data, but also to control significant arrays of information, whose leakage may be a threat to national security.

Platforms have gained the ability to use the accumulated data and to apply artificial intelligence for commercial and other purposes. This advancement has created the preconditions for obtaining a monopoly, or dominant position in the markets, and displacing competitors who is still using traditional forms of trade. This situation contradicts the principles of achieving general economic equilibrium and efficient allocation of resources, which, for instance, the Cournot model describes as the advantages of perfect competition in comparison to an oligopolistic market.

A visible trend of monopolization is currently traced in large digital economies, which confirms the expediency of the third stage of regulation. Due to the «scale effect» and «network effects,» one dominant company owning a popular digital platform may capture a significant market share, from 20% to 45%. This can be seen in China where T-mall has captured 45%, in the USA where 30% of the market belongs to Amazon (30%), in the EU countries Amazon's share is 20% (see Figure 1). In contrast, in the smaller RK market, competition remains more balanced, and there are several large players present.

An unprecedented increase in the concentration of market power and global inequality observed in the PRC, USA, and EU countries, may lead to long-term instability. Thus, antitrust regulation, which has already played an important role in regulating traditional market relations between buyer and seller, must be adapted and strengthened to ensure balance and sustainable development in the digital environment. It is within the framework of solving antitrust problems, responding to the challenges posed by the dominant positions of large companies that the jurisdictions under consideration are working on the legal issues unresolved at the second regulation stage, insufficient attention to the specific intermediary role of platforms in the market being among them.

Countries with the highest market monopolization, PRC and EU, have adopted laws that correlate with the third stage of regulation. As part of the third stage, the EU has introduced a regulatory framework for the digital sector through the Digital Markets Act (2022) and the Digital Services Act (2022; hereinafter — DSA). The Digital Markets

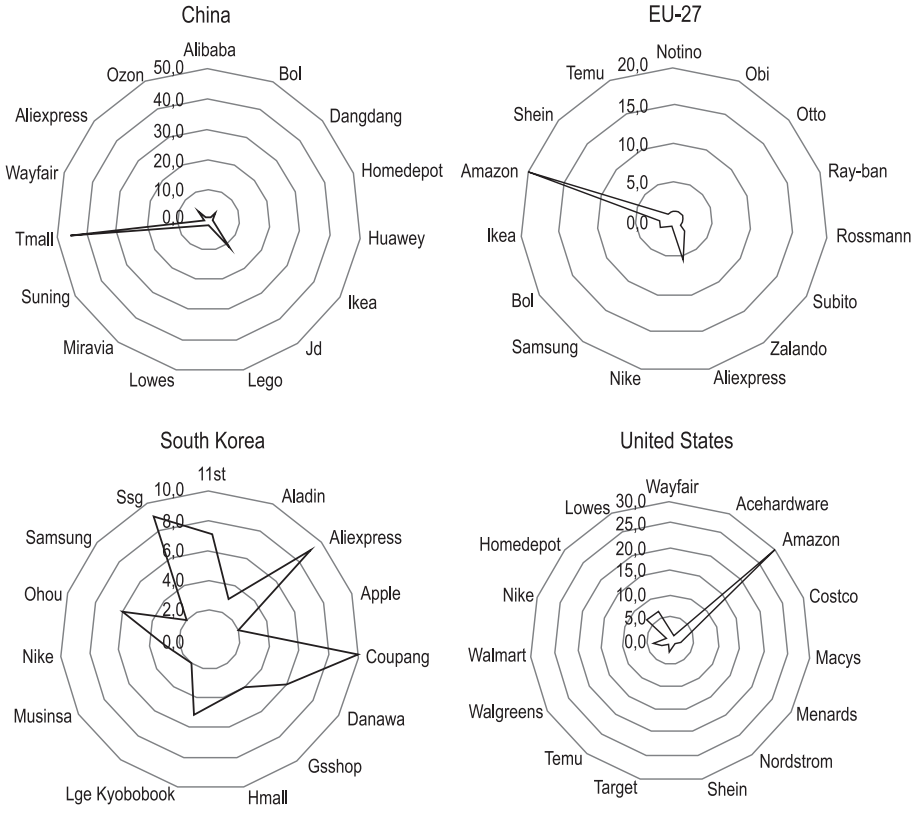

Figure 1. Distribution of major players in the jurisdictions under consideration

Source: Statista.

Act (hereinafter–DMA) gives the European Commission the power to supervise large digital platforms (hereinafter — LDPs), defined as «gatekeepers.» The DMA aims to create fair and competitive conditions for business and end users. «Gatekeepers» are defined as entities providing «core platform services» (hereinafter — CPS), listed in Art. 2, they are to meet the quantitative thresholds for revenue and active user reach, which are specified in Art. 3, with the aim of presuming the materiality of their impact on the market. In addition, to confirm this official status a decision by the European Commission is required, the status is received by the platforms after notifying the commission of reaching the specified thresholds.

Both acts mentioned above manifest the EU's comprehensive approach, based on assessing the scale of digital platform activities and

imposing additional obligations on LDPs due to their significant market power and potential risks to the stability of civil commerce. The DMA and DSA pay particular attention to the operational activities, duties and responsibilities of large, dominant digital platforms (hereinafter — DDPs), with targeted legal approaches based on the specifics of the regulated legal relations, which clearly demonstrates the regulator's special attention to creating a fair competitive environment.

In China, the regulation of digital platforms relies on the Anti-Monopoly Law of the People's Republic of China (反垄断法), passed in 2008. Key amendments affecting the activities of digital platforms came into force on August 1, 2022. The Anti-monopoly Law implies the possibility of adopting subordinate acts and interpretations. In 2021 the State Council of China has passed the Anti-Monopoly Guidelines (guidance) for the platform economy.

Currently China is working on detailing special requirements for digital platform operators through several subordinate acts that are in the public consultation stage. In October 2021, a draft of the Guiding Principles was published, proposing a classification of digital platforms; the criteria to differentiate platforms are such as the main scope of activity, the number of active users, and market capitalization. In accordance with the Principles for Classification and Categorization of Internet Platforms, issued by the State Administration for Market Regulation, six types of digital platforms are distinguished: online sales intermediary platforms, life services platforms, social entertainment platforms, information platforms, financial services platforms, and computing services platforms.

Along with this, the Chinese authorities have recently undertaken several comprehensive changes to subordinate antitrust regulation to take into consideration the specifics of relations developing in the digital economy. All changes were the subject of open discussion with the participation of authorities, experts, and representatives of the real sector. Thus, in 2023, the State Administration for Market Regulation of the PRC has issued the Provisions on Prohibiting Abuse of Dominant Market Positions (禁止滥用市场支配地位行为规定), which has prohibited dominant platforms from using the data they obtained or their algorithms, technologies, or rules to take actions aimed at abusing their dominant position in the market.

According to this document, a dominant market position is recognized if the operator can control prices, volumes, or other transactional

conditions, or could prevent or influence the entry of other operators into the market. In addition, regulatory changes were also expressed in the Provisions on the Prohibition of Monopoly Agreements (禁止垄断协议规定) prohibiting digital platforms from using the data obtained, existing algorithms, technologies, and platform rules to enter into horizontal and vertical monopoly agreements through the communication of their intentions in any form, the exchange of confidential information, or the establishment of coordinated actions.

For example, collusion with the aim of applying the same algorithms and platform rules to calculate the prices of different sellers, profiting from maintaining fixed prices, as well as the coordinated distribution of sales or procurement markets are not allowed. In recent years, the PRC has also paid attention to one of the identified problematic aspects concerning data protection. On September 30, 2024, the State Council of the PRC has issued the Regulations on Network Data Security Management, it came into force on January 1, 2025.

The document focuses on issues of cyber security, data confidentiality, and introduces rules prohibiting operators from using data to discriminate against users or suppliers. The rules have extraterritorial effect and apply to platforms that carry out network data processing both within the territory of the PRC and abroad, if this activity may harm national security, public interest, or the legitimate interests of citizens of the PRC. The rules cover the protection of not only personal information, but also any other information processed and generated through Internet networks, depending on three data categories — general data, important data, and core data.

General standards for information protection in China involve cumulative measures taken by data operators depending on the three indicated data categories. Special attention is paid to the protection of important data; that includes information that, in the event of forgery, destruction, leakage, illegal acquisition, or illegal use, could threaten national security, economic activity, social stability, healthcare, and the safety of the population of the PRC. Owners of important data are subject to increased information security requirements: they are to conduct a full-scale risk assessment annually and before each transfer of such data to a third party, as well as to report on the technologies used to protect such data.

The Regulations introduce special rules emphasizing the need for special attention to dominant platforms:

platforms processing data of more than 10 million people are required to comply with the standards applicable to owners of «important data»;

«big (large) platforms» are required to refrain from using data, algorithms, or user conditions to carry out unfair and misleading actions, such as forced data processing or discrimination against users; to oversee personal data protection they are required to establish an independent supervisory body consisting of both employees and external experts; they are to publish an annual report on social responsibility in terms of personal data protection.

The regulation of digital platforms in the PRC is characterized by a comprehensive and strict approach aimed at ensuring competition and data protection. Its features include the recognition of significance of platforms as an independent type of activity, an emphasis on antitrust regulation, strict requirements for data protection, extraterritorial effect of legislation, and increased control over large platforms. This reflects a desire for comprehensive control and the specifics of the country's political and economic system.

The Republic of Korea, in turn, demonstrates a softer approach to regulating digital platforms. In 2021 amendments were imposed into the Telecommunications Business Act, prohibiting app store operators from unfair practices against app developers. Currently, the Korea Fair Trade Commission (hereinafter — KFTC) and the government are discussing additional regulatory measures, including the requirement for the operators of foreign platforms to appoint a representative in Korea and the encouragement of self-regulation.

On January 12, 2023 the KFTC has published the «Guidelines for Reviewing Abuses of Dominant Market Positions by Online Platform Operators» (Online Platform Monopoly Guidelines), clarifying antitrust legislation. Moreover, on September 11, 2024, the KFTC has announced a new regulatory roadmap, proposing the following amendments to the recently approved Monopoly Regulation and Fair Trade Act (hereinafter — MRFTA). At the same time, the Chairman of the Commission has noted that, despite the initially different concept, a decision was made to take an alternative path of developing precise threshold values for establishing a presumption of market dominance, after which a strict list of prohibited actions is applied (as opposed to the EU approach, where platform regulation operates on the principle of «ex-ante,» i.e., preventive intervention to prevent abuses). The development of regulation will be conducted for six types of platforms: transaction in-

termediary platforms, search engines, video platforms, social networks, operating systems and advertising platforms.

A distinctive feature of regulation in the Republic of Korea is their active promotion of self-regulation of the platform economy industry. In July 2022, the Platform Regulatory Council was established, and in August of the same year, the Non-Governmental Self-Regulatory Organization for Digital Platforms (Platform SRO), operating in four divisions: platform and business user relations, consumers, data and artificial intelligence, and innovation and management. Despite the advantages of self-regulation, the KFTC insists on the need to introduce additional rules for large digital platforms. The delay in legislative decisions may be due to the dominance of USA platforms, which introduces a cross-border element and a clash of economic and political interests.

In the USA, the third stage of regulation development also fell on the period from the early 2020s and it is characterized by a reaction to market abuses by digital giants and by conflicting interactions between stakeholders, as well as obstacles to legislative decisions due to opposition from dominant corporations. In 2020, the Judiciary Committee of the US House of Representatives has published a report presenting evidence of anti-competitive behavior by large technology companies. The US Congress today has two bipartisan bills approved by specialized committees: the American Innovation and Choice Online Act (AICOA) and the Open App Markets Act (OAMA).

AICOA affects resources with more than 50 million active users per month in the USA (or at least 100,000 active business users per month) and a market capitalization of at least $600 billion, related to: the creation and exchange of content, search, or the sale, delivery, or advertising of goods or services. In fact, the bill is to affect Google, Apple, Amazon, Microsoft, etc. OAMA concerns app stores and related operating systems, such as iOS, Apple App Store, Google Play and Android, Microsoft Store on Windows, and so on.

In the USA, two definitive types of digital platforms are designed — in relation to app stores and in relation to other digital platforms, including digital giants. OAMA is focused on suppressing self-preferencing practices of app stores, and AICOA contains a number of more detailed and highly specialized requirements aimed at a wider segment of the platform economy. In addition to prohibiting self-preference, AICOA contains prescriptions that include prohibitions on discrimination against users, restricting access without using other platform products,

using non-public data of business users to compete with their products, hindering access to user-generated data, and hindering the removal of pre-installed software or changing default settings.

Although AICOA and OAMA have not yet come into force, the first decentralized steps have already been taken in the USA to curb the anti-competitive activities of platforms. In 2020 the Federal District Court for the Northern District of California considered two disputes in a lawsuit filed by Epic Games against Google and Apple, the decisions on which became precedents. In the case against Google, the jury has found Google's actions to be anti-competitive practices with the properties of a monopoly. The court has issued a permanent injunction prohibiting the obstruction of the installation of alternative app stores for Android, as well as the payment of incentives or the provision of discounts to developers who release applications exclusively through the Play Store. In the case against Apple, the court has found the company's practice of preventing users from switching to other sites and app stores (anti-steering policies) to be a violation of antitrust law.

A comparison of the digital platform regulation in the RK, USA, PRC and EU demonstrates a variety of approaches: from an emphasis on self-regulation in Korea to attempts to adopt large-scale legislative acts in the USA and strict state control and the extraterritorial effect of legislation in the PRC. The most comprehensive, but also the most controversial, is the approach to regulating and defining large platforms in the EU (DMA and DSA).

The experience reviewed shows that all jurisdictions strive to ensure competition, data protection, and consumer rights. Thus, legal systems are adapting to the challenges of the platform economy — albeit with varying degrees of stringency. The above mentioned court decisions demonstrate the relevance of the problem: the absence of proper legislative regulation leads to fragmented and complicated resolution of fair trade issues involving digital giants, which does not provide proportionate restraint of dominant position abuses. Court decisions confirm the illegitimacy of burdensome conditions imposed by digital giants on counterparts, which ultimately affects consumers negatively.

The third stage is regulatory consistent: the rules of antitrust and civil law regulation, as well as the protection of user data, form independent characteristics of digital platform services. As the quality of these services affects the state of competition, reasonably high special requirements, including strict prohibitions, must be imposed on their provision, which corresponds to the regulatory policy initiated in the 21st century.

Generally, the third stage reveals the relationship between regulatory and terminological problems: the lack of clear-cut definitions of platform economy concepts makes it difficult to distinguish the interests of participants and to regulate platform economy with purpose. The need for high-quality legal definitions is due to the need to distinguish platform activities from other activity types, to identify the features of the independence of platform services, to take into account their impact on competition, and to determine the criteria for significant types of digital platforms. The introduction of a criteria-based distinction between a digital platform, its types, operators, business users, and consumers is an urgent task, since the absence of such a distinction is a source of collisions.

The analysis of foreign experience mentioned demonstrates the evolution of digital platform regulation, where each stage consistently solved arising problems, but the dynamics of the platform economy led to the emergence of new challenges. Despite the progress in antitrust regulation, data protection, and terminological certainty, a comprehensive approach considering the specifics of the national economic environment remains key to success. Thus, the analysis of the regulatory environment of the Russian Federation, with its unique features, such as the influence of tax legislation, is of particular interest for the development of an adaptive and balanced model to regulate the platform economy.

2. Experience in Regulating E-commerce in Russia: Analysis of the Legal Framework During the Three Stages

An examination of foreign experience in regulating the platform economy has revealed three distinct stages in the evolution of legal rules. The following analysis will focus on legislative solutions adopted in Russia during the periods comparable to the identified stages of foreign regulatory development. The purpose of this section is to identify directions for further regulatory development, considering both compliance with global trends, taking into account the specific aspects of taxation, and the need to address current and potential issues in the platform economy.

Particular attention will be paid to analyzing the compliance of Russian regulations with the second and third stages of regulation development, which correspond to data protection issues and the implementation of a comprehensive approach to regulation. This focus will allow to identify areas requiring improvement and to formulate practical recommendations for legislative development.

At the first stage the primary consumer protection measures were taken, the key event was the introduction of the concept of distance selling of goods in Article 26.1 of the Law of the Russian Federation No. 2300-1 of February 7, 1992, "On Protection of Consumer Rights" (hereinafter– the Consumer Rights Protection Law). This measure has allowed for the adaptation of consumer protection mechanisms to the specifics of online commerce, like establishing requirements for information about product, the right to refuse goods, and quality guarantees. At the same time, this step has laid the foundation for regulating relations in the digital environment but focused primarily on the classic "seller-consumer" model, which implies only the fact of purchase and sale and the interaction of the seller and consumer without digital inter mediation.

The second stage was characterized by expanding the scope of regulation to digital aggregators. Federal Law No. 112-FZ of May 5, 2014 has enshrined the freedom to choose the form of payment in the Consumer Rights Protection Law, and Federal Law No. 266-FZ of July 1, 2021 has established safe legal framework for the collection and analysis of consumer data. An important innovation was the definition of "Owner of an aggregator platform for goods (services) information" which made it possible to extend consumer guarantees to digital platforms. Moreover, Federal Laws No. 250-FZ of July 29, 2018, and No. 135-FZ of May 1, 2022 have specified the provisions of the Consumer Rights Protection Law regarding the liability of aggregator owners for their misconduct. Thus, the consolidation of the status of aggregator was an important step in adapting legislation to the realities of the platform economy, which blur the traditional boundary between a seller and an intermediary. However, there remains a need for further detailing the responsibility of aggregators and platform operators, as well as distinguishing their functions. The proximity of the concept of "aggregator owner" to the definition of "digital platform owners" (Zap'yantsev A.A., 2024: 57–60) indicates the potential for unifying terminology and developing regulation.

At the second stage of the development of platform economy regulation in the Russian Federation, an approach was formed to the distribution of rights and obligations of sellers, aggregator owners, and consumers, it aimed at ensuring guarantees of consumer rights in digital commerce and creating conditions for business development.

The analysis of the digital commerce market structure in Russia (Figure 2) reveals relatively balanced competition, characterized by the presence of several major players. The Republic of Korea has a similar situation, with a specific regulatory approach including the elements of

self-regulation being developed. The competitive environment in Russia, unlike the USA, EU, or China, may be due to the smaller economic scale of digital markets, as well as due to regulatory policies to maintain stable relationships between market participants. It should be noted that the data collected in Figure 2 requires additional studies to identify the specific factors influencing the market structure and to assess the utility of regulatory measures.

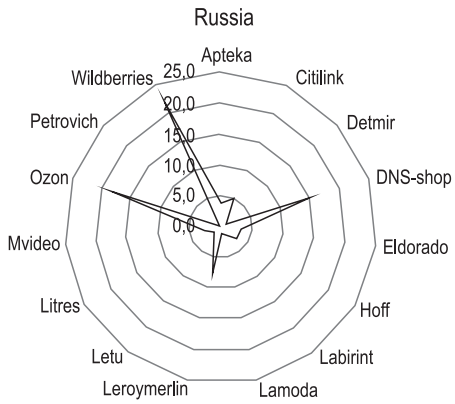

Figure 2. Distribution of major digital market players in Russia

Source: Statista.

In line with the foreign trends in legislative development, Federal Law No. 301-FZ of July 10, 2023, «On Amendments to the Federal Law «On Protection of Competition»» (the fifth antimonopoly package), was passed in Russia as part of the third development stage of regulatory policy regarding digital commerce. The law aims to strengthen control over the activities of large digital platforms and to prevent abuses of dominant market positions. Law No. 298-FZ introduces amendments to antimonopoly regulation adapted to the specifics of digital platforms, expanding the criteria for determining a dominant position, strengthening liability for anti-competitive agreements (including those using algorithms), requiring transparency of algorithms and data, and expanding the powers of the Federal Antimonopoly Service. The passing of the fifth antimonopoly package demonstrates the Russian legislator's desire to follow foreign trends in regulating the digital economy and countering anti-competitive practices of large digital platforms.

However, the analysis of legislation reveals the lack of unified terminological apparatus, in particular, the absence of a clear definition

of «digital platform.» It is assumed that the absence of clear definitions may lead to ambiguous interpretations of the law, problems in qualifying entities falling under its scope, and, as a result, difficulties in law enforcement and the potential violation of the principle of equality before the law. For better implementation of the provisions of the fifth antimonopoly package, advancement of legislative and regulatory acts is required, as is the formation of judicial practices that reflect the unique aspects of digital markets.

Overall, Russia still lacks comprehensive regulation of digital platforms, and the conceptual apparatus of Law No. 298, although it appeared to be a major step forward, was not designed to cover all intersector regulation. The fragmentation of definitions in the field of digital commerce creates risks for businesses and requires significant resources to ensure compliance with numerous acts. Despite the general regulatory impact of the Civil Code of the Russian Federation (hereinafter — the Civil Code) and the Tax Code of the Russian Federation (hereinafter — the Tax Code) on civil and tax relations involving digital platforms, there is no dedicated regulation that considers the specifics of the platform economy. They do not constitute regulation of the platform economy as such. Instead, they regulate general civil and tax relations digital platforms are involved in. The Civil Code governs the relationship between the platform and user-contractors (agency, commission, performance of services) and between user-contractors and customers (conclusion of contracts, purchase and sale, contract, performance of services).

Furthermore, unlike trading aggregators with transparent payments, information platforms (classifieds websites and social networks) create risks of incomplete reflection of users income, requiring increased attention from tax authorities. At the same time, the Tax Code establishes general rules for paying personal income tax and corporate income tax. However, unlike trading digital aggregators (marketplaces), where payment for transactions is transparent, information platforms (classifieds, social networks), providing only information services without recording contacts and payments, create a situation in which users may not account for income, despite the generally applicable norms of tax legislation. There is no specific regulation of classifieds activities in this respect.

It is worth noting that measures have been taken within the framework of the third stage of regulation to specialize tax regimes for digital platforms. Article 208 of the Tax Code has introduced a rule for taxable income to include remuneration for work, services, and intellectual property rights

provided/granted via the Internet using the Russian domain zone or hardware located in the Russian Federation. This rule also applies to individual entrepreneurs, platform owners providing intermediary services. This provision aims to increase tax collection on income received through Russian digital platforms and to establish equal taxation conditions for various types of activities. However, the force of this rule will be determined by its application and the ability of tax authorities to identify cases where income from activities on digital platforms is not fully accounted for.

The latest amendments to the Tax Code (Article 284) address the taxation of income from the provision of services on the Internet placing advertisements and offers, but marketplaces, taxi and food ordering services qualify for exemption from tax. This exemption, along with the terminological heterogeneity inherent in the regulation of digital platforms, may indicate the need for further systematization of approaches. It is assumed that the general norms of the Tax Code on taxation of corporate profits apply to these three types of platforms. Federal Law No. 422-FZ of November 27, 2018, «On Conducting an Experiment to Establish a Special Tax Regime «Tax on Professional Income»,» offers a clearer definition of digital platform operators, covering a wide range of participants in the platform economy, including large platforms.

Thus, the Federal Law mentioned, in the context of taxation, appears to be a more consistent and universal instrument for digital platform regulation compared to individual provisions of the Tax Code, that, nevertheless, require further interpretation and harmonization with other regulatory acts.

It should be noted that, within the framework of the third stage of e-commerce regulation development, the Federal Tax Service (hereinafter — FTS) organized information exchange with 70 operators of digital platforms, including systemically important digital platforms of various types (Wildberries, You Do, Yandex Taxi, etc.), to simplify the tax payment procedure for users. Despite this initiative, which demonstrates progress in regulating platform activities, the overall extent of regulatory impact on the platform business remains uneven one. The discrepancy between the criteria for classifying entities as «operators of digital platforms» for the purposes of interaction with the FTS, on the one hand, and the definitions of digital platforms used in other sector laws, on the other, may reduce the predictability of legal relations with other regulatory authorities (e.g., Federal Service on Consumer Rights Protection and Human Well-Being, “Rospotrebnadzor”).

The above shows that, despite timely initiatives, the regulatory legal framework for digital platforms still may be described as fragmented and conflicting; that does not allow for the full realization of the digital commerce potential with the greatest benefits for the private sector (businesses and consumers) and with the least risks for the state. Thus, to strengthen and expand the progress achieved, it is necessary to move to the next stage of regulatory development, to unify and comprehensively systematize the regulatory legal framework which will be able to provide a clear system of rules for the digital market and to coordinate activities of various government bodies to ensure the stability and predictability of market relations.

The analysis of legal terminology has revealed an uneven coverage of various types of digital platforms, that factor raises, in some cases, the question of their eligibility for tax exemptions, as regulatory gaps can be profitably used by unscrupulous participants in digital commerce. The analysis is presented in detail in Table 1.

Table 1. Basic concepts used in Russian legislation as of 2024

Definition	Market-place	Classifides Websites	Online shop	Service intermediary
1. Law of the Russian Federation No. 2300-I of February 7, 1992, «On Protection of Consumer Rights», Preamble: **Owner of an aggregator platform for goods (services) information**	■		■	■
2. Federal Law No. 422-FZ of November 27, 2018, «On the Implementation of an Experiment to Establish a Special Tax Regime 'Tax on Professional Income'»: **Operator of an electronic platform**	■	■		■
3. Federal Law No. 135-FZ of July 26, 2006, «On Protection of Competition»:	■			■

Definition	Market-place	Classifides Websites	Online shop	Service in-termediary
Digital platform (The scope includes only the category termed "Programs for electronic computing machines,", e.g. apps, whereas websites and information systems fall outside its scope)				
4. Federal Law No. 289-FZ of August 3, 2018 «On Customs Regulation in the Russian Federation and on Amendments to Certain Legislative Acts of the Russian Federation»: **Trade platform (website)** (The scope includes only website, whereas "Programs for electronic computing machines" fall outside its scope)				
5. Art. 147 Tax Code: **Electronic trade platform**				
6. Federal Law No. 149-FZ of July 27, 2006, «On Information, Information Technologies, and Information Protection»: **Audiovisual service**				
Federal Law No. 149-FZ of July 27, 2006, «On Information, Information Technologies, and Information Protection»: **Classified ads service**				

Thus, the fragmented and inconsistent legislative regulation of digital platforms in the Russian Federation creates cognitive and transaction

costs for economic entities and consumers. These costs are due to the absence of a unified concept and clear criteria to distinguish between different types of platforms that simultaneously enter several types of legal relations. Moreover, such phenomenon as niche regulation provokes conflicts in law enforcement and reduces regulatory efficiency. As a result, the same platform de facto may fall under divergent regulatory regimes. This directly reduces the degree of legal certainty and transparency of requirements: for example, the definition of a social network in Article 10.6 of the Law on Information manifests incongruence with existing models of digital platforms, although nowadays remote commerce takes place through these networks very actively. The application of traditional legal constructs to digital platforms without considering their specifics is fraught with regulatory imbalance, since, for example, it is not always fair to assign responsibility for the quality of goods to the platform rather than to the seller.

Finally, the absence of legal rules that explicitly count the influence of algorithms hinders the implementation of tort liability for breaches of consumer rights. The lack of adequate regulation of digital platforms also creates risks for contractors (Silberman M.S., Harmon E., 2018: 911]; [Stewart A., Stanford J., 2017: 425), potentially leading to economic instability and increased social inequality (Drahokoupil J., Jepsen M., 2017: 103]; Healy J. et al., 2017: 232–245]; [Lehdonvirta V., 2018: 19–29). In the long term, this may force platforms to take excessive preventive quality control measures, which would negatively affect user-contractors and the development of small and medium-sized businesses, as well as platform pricing policies, and it would ultimately negate the positive impact of the platform economy.

To preserve balanced development of the digital economy and to protect consumer rights, it is necessary to develop a *modus operandi* for interaction between participants in the platform economy. At the same time, it is advisable to establish a *praesumptio* of responsibility for the owner/operator of the platform for control of proper functioning of the algorithms and their correctness, for example, in cases when the digital elements of the platform, algorithms, violate consumer rights. It is also important to ensure compliance with the principles of *bona fide* by both platform owners and users, to strive for *status quo ante* in cases of violations of consumer rights, and to take into account the principle of *pacta sunt servanda* when elaborating regulation for contractual relations in the platform environment.

Additionally, it should be stated that the application of traditional legal norms disregarding the intermediary nature of platforms is fraught with unjustified imposition of liability. On the other hand, excessive regulation, coupled with high bureaucratic costs, is able to influence negatively small and medium-sized businesses and limit consumer choice.

Considering the above, it is important to emphasize that the regulation of the platform economy requires the creation of a systematic and uniform legislative framework that takes into consideration the specifics of digital platform activities and the interests of all stakeholders. Key to this is the recognition of platform intermediary function, the formation of a unified terminological apparatus, and the distribution of liability accordingly.

3. Legal Glossary as a Priority Task for Regulating the Platform Economy

As demonstrated above, such regulation—with its significant divisions—creates risks for digital platforms, even in the case of gradual (evolutionary) development. Different requirements prescribed by classic branches of civil law vis-à-vis consumer protection provisions lead to contradictions in transactions. Differences in information law and personal data protection requirements hinder cross-border activities and confidentiality. Inconsistencies in antitrust regulation weaken the fight against unfair competition. As a result, the lack of a unified approach produces legal uncertainty and increases costs for platforms. This hinders their innovative development and the formation of a predictable environment for business.

Friedrich Hayek among others emphasized that economic success is based on the predictability of the legal reaction to the actions of economic agents (Hayek F., 1944). In this regard, to maintain system and uniformity of legal impact, it is necessary to harmonize the terminological apparatus in various branches of law applicable to digital platforms. This is particularly relevant for legal systems rooted in the continental tradition, which are built on a «structural» approach and heavily depend on a precise and coherent system of legal concepts enshrined in legislative texts. As A. Ortolani recently has noted aptly, the “tendency to organize knowledge in a well-ordered and cross-referenced system is a distinctive trait of the civil law tradition which continues today” [Ortolani A., 2024: 211–234].

The lack of a unified approach to regulation, manifested in the uncertainty of criteria for classifying entities as digital platforms, in the selective application of sector requirements to individual platforms, and in the absence of a general vector of applicable legislation interpretation, causes inevitably legal uncertainty. In addition, there is a dual paradox: the specialized regulation of several aspects of platform activity is characterized by a high degree of detail and technocracy, while there is no normative typology of platforms that distinguishes them by the specifics of their economic activity at the most general level even though such a typology is highly demanded. It is necessary for the adequate application of regulatory measures to various types of platforms.

A unified conceptual and terminological framework for application to the activities of digital platforms will be, logically, a necessary basis for the development of high-quality regulatory legal acts governing a specific product or a specific type of digital platform activity, since this is the only way on that the range of subjects and objects that legal norms affected may be clearly defined. As A. Strowel and J. Vergote have noted rightly [Strowel A., Vergote J., 2017], when developing a legal framework for regulating a platform market, it is most appropriate to first form a general (inter-sector) structure for regulating the digital economy (including the principles of regulating the platform economy and the terminological apparatus), and only after that to move on to more specific issues relating separately to various aspects of digital platform activities.

The comprehensive approach, which is overdue at the third stage of regulatory evolution, like any consolidation of law, will produce a positive influence on the development of relevant public relations, due to increased certainty, elimination of legal conflicts, and the construction of a clear system of interaction between citizens, businesses, and the state [Zhukov V.N., Frolova E.A., 2024]. Therefore, while finding the optimal legal regulation of the platform economy, it is necessary, first to define the concepts, as the adherents of classical legal positivism argued (Nersesyants V.S., 2003).

Thus, the primary task is to develop a precise terminological framework that adequately reflects both the general and specific characteristics of digital platform activities. Without a harmonized terminological understanding of phenomena, it is impossible to begin conceptualizing the principles of legal regulation, since these principles should be targeted at specific subjects and objects (digital platforms, intermediary digital services, independent remote employment or work, etc.) that have not yet been legally defined.

In this regard, the concepts projected defining the platform economy should contain the main features that distinguish the defined phenomenon from similar ones. Overburdening definitions with elements of legal regulation is not advisable. First, it is necessary to introduce a basic concept of «digital platform» and to define the types of digital platforms (marketplace, classified, etc.), considering the characteristics of each type when differentiating sector regulation to minimize negative effects and unjustified dominance of individual types. The diversity of digital platform types must be taken into account within the framework of further legal regulation.

When developing legal definitions within the Russian legal system, it is essential to ensure their alignment with the terminology of the Law on Information, the Civil Code, the Tax Code, the Customs Code of the Eurasian Economic Union, and the Consumer Rights Protection Law, which has become common in legal practice, in order to avoid large-scale changes in legislation. Otherwise, massive changes in the legislation of the Russian Federation and the EAEU will be inevitable.

Thus, considering the terminological constructions used in these legislative acts, it is proposed to understand a digital platform as a website and (or) a page of a website on the Internet, and (or) an information system, and (or) a computer program intended, and (or) used for the purpose of selling goods or works or services, which provides users with the opportunity to receive full information about such goods or works or services and about related offers to conclude a contracts of sale (including an agreements for performance of work or an agreements for performance of services), and, if applicable, that also allows to remotely conclude these contracts and make a down payment for these goods or works or services.

To define the burden of fulfilling obligations and to delimit responsibility for fulfilling the requirements that are imposed on digital platforms, it is necessary to understand adequately the difference between the owner and the operator of a digital platform. The owner of a digital platform is a natural person, including an individual entrepreneur, or a juridical person who has a legal title to the digital platform and is responsible for its strategic management and development. In turn, the operator of a digital platform is the owner of a digital platform (if the owner retains this functionality), or a person authorized by him, who administers the digital platform and ensures its functioning, including interaction with users of the digital platform.

Clarifying the definition of digital platforms in relation to the activities of trading digital aggregators — marketplaces — it is necessary to remember that they are entities regulated by most detailed legislation of the Russian Federation in comparison with other types of digital platforms. This is explained not only by their greatest popularity due to their direct daily work with consumer goods in demand (which affects the total consumer demand), but also by their business model, which assumes: establishing a direct contact between users (seller/contractor and buyer/customer), formalizing the relations through the conclusion of a public contract (Art. 426 of the Civil Code), payment for goods/services either directly on the digital platform or under the control of its operator (charging a fee by a partner on behalf, for example, at a point of issue or by a courier). Even more obvious is the presence of a conscious, purposeful contact between the operator of the digital platform and user contractors, who conclude one of the forms of inter mediation transaction to organize interaction between themselves and between contractors and customers to make a profit.

Thus, the business model of marketplaces has all the signs of the emergence of civil law relations and typical factual patterns that are amenable to generally accepted methods of legal regulation with the establishment of appropriate exceptions (features), where necessary. Given these considerations, it is not surprising that the Consumer Rights Protection Law — a key act regulating the procedure for the sale of goods, works and services to every person in everyday life — was one of the first to be extended to the activities of marketplaces. Moreover, the influence of this Law was extended to digital platforms without deconstructing its underlying concept and structure — the triad that has been in force for more than 30 years: «General Provisions — Consumer Rights in the Sale of Goods — Consumer Rights in the Performance of Work or Services.»

It should be stated the current legislation does not differentiate between services and goods in relation to digital platforms organized according to the marketplace business model. Consequently, digital platforms through which both goods and services are sold (for example, food ordering services are also considered aggregators and are covered by the concept of a marketplace from the point of view of the Consumer Rights Protection Law) fall under the concept of a marketplace ("aggregator owner" in the strict wording of the Law). It is also important that, in addition to the previously studied definition of the "aggregator owner", the Law selectively incorporated this new participant in the consumer market into the current rules for selling goods, works and services to

consumers. At the same time, not all rules were extended to aggregators, but only the most significant from the point of view of protecting consumer rights and considering both the specifics of digital commerce and the need for its development in the future.

However, the concept of «aggregator owner» stems mainly from the needs of consumer legal relations and does not fully describe the specifics of information legal regulations are essential for the platform economy in the context of current rates of technological development. In view of the above, to formulate the definition of a trading digital aggregator (marketplace), it is proposed to combine approaches derived from civil and information law. Taking into account the previously identified turning points in both Russian and foreign practice, such an aggregator should be understood as a digital platform through the online storefront of which the platform operator and/or user-contractors direct a public offer to an indefinite number of persons (place offers on the Internet) regarding goods sold, works performed, and/or services provided by them, enabling contact with user-customers and/or the remote conclusion of contracts for sale, compensated work, or compensated services, as well as the possibility of making advance payments for goods, works, or services. At the same time, the online showcase is an audiovisual element of the digital platform that allows the user customer to search for goods, works, services, to familiarize themselves with information about goods, works, services to continue their correct selection. This clarification will also be useful for regulating the requirements for information about goods, works and services.

Ordinary online stores, well-known to the majority of consumers, should obviously not fall under this definition, in accordance with the goals of potential legal regulation. This is one of the fundamental issues, whose solution is a terminological innovation, since the current legislative understanding of the «aggregator owner» does not differentiate between ordinary online stores and the entire variety of digital platforms. At the same time, not only the legal nature of their activities, but also their business models themselves are strikingly different. An online store is, in fact, only a website of a specific real seller who uses this remote method of selling their goods on a par with the traditional method (offline). Marketplaces, as it has been shown in this study, are inherently built for intermediary activities, their business model is to combine many sellers on their digital platform and to create competition for their offers. Therefore it is necessary to clarify that an online store constitutes a specific type of digital platform, whose online storefront provides an

indefinite number of persons with information about the goods offered by the digital platform operator and/or related parties. This platform enables the user-customer to familiarize themselves with the seller's offer to conclude a contract of sale for such goods, to enter into the contract, and to make payment (including prepayment) through applicable forms of non-cash transactions.

At the same time, the most «transitional» form of the platform economy, distinct in its economic nature, is the classified. Unlike trading digital aggregators (marketplaces), the current legislation does not contain a concept and does not regulate the activities of information digital aggregators, which classifieds are. The regulation provided by the Law on Information practically does not address to the issues of civil circulation with the help of classifieds like registration and regulation of relations between digital platform operators and user contractors, between user customers and user contractors.

It is important to note three features of the current regulatory norms established by the Law on Information. First, only the platform, access to which is more than one hundred thousand Internet users per day, is recognized as an ad posting service. Consequently, the classifieds of smaller scale, even with 90,000 users per day, do not fall under regulation at all, although even if a tenth of that number of users concludes a transaction in the amount of ~2,000 rubles, the turnover will be ~18,000,000 rubles per day, i.e. ~540,000,000 rubles in 30 days. It turns out that the quality of products sold in this way, as well as the issues of shadow employment and legalization of such amounts of money remain outside the purview of the state.

Secondly, the Law on Information establishes a requirement for the owner of the ad posting service — he must only be a citizen of the Russian Federation who does not have citizenship of another state, or a Russian legal entity. This approach differs significantly from the approach to regulating marketplaces. At the same time, since the concept of an ad posting service is constructed through reference to the language of the posted advertisements, it may be assumed that foreign services where advertisements are posted in foreign languages do not fall under the requirements of this Law, including the requirement for citizenship of the classified owner. However, with this approach, some legal collision is noticeable in the relation of these criteria.

Thirdly, the aforementioned Law provides the Government of the Russian Federation with the opportunity to impose requirements on

"ad posting service" operators to ensure the integration and interaction of the service with Unified Identification and Authentication System and the Federal State Information System «Unified Portal of State and Municipal Services (Functions)» (transliterations: ESIA and FGIS EPGU). Therefore, it is theoretically permissible to develop this regulation in order to ensure proper recording of agreements concluded between customers and contractors through classifieds (preventing «going» into the gray area and concluding a transaction without recording on the digital platform).

In view of the above, there is an acute need for clear conceptual delimitation of classifieds specifically for the purposes of regulating the platform economy and in accordance with its inherent features, not only for information policy considerations. Then, an information digital aggregator (classified) should be understood as a digital platform on which digital platform users independently post information about offered or requested goods, works, services and which allows user contractors and user customers to establish contact for the purpose of concluding a contract, and (or) conclude a contract, and (or) pay (prepay) under the concluded contract.

The online store, marketplace and classified have been discussed above, but there is also a larger-scale phenomenon, which needs definition especially in connection with the consolidation of players: digital giants tend to combine several diverse digital platforms under a «single cloud,» offering cross-referrals to complementary services (in order to increase referrals and profits) and encouraging users for such behavior. The practice of large companies combining platforms for ordering goods/food/products/medicines, providing educational/telemedicine services, audiovisual services, courier services, etc. under their influence is well known. There arises a digital ecosystem — a set of digital platforms united by belonging to one person (one group of interdependent persons), through the joint and (or) interdependent functioning of which (including the organization of a unified system of authorization and authentication, the establishment of a coordinated system of discounts, increasing the convenience and (or) profitability of accessing several such digital platforms at the same time) the person (group of interdependent persons) attracts increased interest of the user customer, motivates them to make additional purchases, order additional services from the person (group of persons), forms additional consumer value of accessing these digital platforms. Separately, it is worth considering complex digital platforms that combine features of individual types and (or) types of digital platforms.

Thanks to the above conceptual series, we are capable to solve the priority problem of not only identifying the key actors in relations in the platform market for the purposes of law, but also of meaningfully delimiting the nature of their activities (including the services provided), which is, of course, intermediary in essence. However, even with these definitions, the conceptual model cannot yet look logically complete and systematic. In addition, an accurate legal description of the participants in legal relations «on the opposite side,» is required i.e. considering those using digital platforms to enter the trading process (on both the demand and supply side). For this purpose, it is required to name the digital platform user as such, as well as their individual varieties — as it has already been partly shown in the previous definitions, these are user contractors and user customers.

The concept of a digital platform user is generic one. Users are individuals or entities utilizing the platform for, or intending to utilize it for, transactions involving goods, works, or services. In turn, the differentiation of this concept should occur according to the nature of the relationship between the user and the operator of the digital platform, i.e. based on the purposes of its entry into legal relations and depending on «which side» it joins the platform. Consequently, a user-contractor should be recognized as such a digital platform user who is a seller, contractor or "platform worker" and places, on the basis of a remunerated contract concluded with the digital platform operator, publicly available information about the goods they wish to sell or the works (services) they wish to perform, as well as about the offers for the purchase and sale of goods, the performance of works (services), and for the conclusion of the pertinent contracts with user-customers.

At the same time, there are ample grounds to refer to the concepts of seller and service provider in their traditional meanings as established by the Consumer Rights Protection Law, that, among other things, will maintain continuity between innovative legal regulation and the well-established, time-tested, and proven legal framework. In turn, this definition itself, as can be seen, dichotomously assumes that a user-customer is a digital platform user who intends to purchase goods or works or services based on an offer posted on the digital platform by a user-contractor, or who posts a relevant request for goods or works or services on the digital platform.

It appears to be that comprehensive, systematic, empirical, and practice-oriented elaboration of legal terminology is key to enabling inte-

grated regulation of the pressing issues facing society and the state in this field. In this regard, the glossary developed in this study is based on an analysis of the successful foreign and Russian experience in regulating electronic and digital commerce. At the same time, the features of the Russian legal system and the possibility of applying pertinent definitions in related areas of law were duly considered. This glossary, therefore, reflects the key principles of legal regulation in different countries, primarily from an economic point of view (what the type of platform is as an economic question) and is adapted to the Russian legislation. The next step will be to develop a regulatory framework that will combine the proposed definitions and world experience into specific rules and regulations applicable to the platform economy in Russia. This framework will create more clear and useful regulation in all the features of this area.

4. Key Principles for Regulating Platform Economy: *Primum Non Nocere*

To date, the regulatory framework governing economic and commercial activities on or through the Internet can be described as fragmented, unsystematic, and sometimes contradictory one. On the one hand, a whole layer of tax, antitrust, consumer, and information relations in the sale of goods (works, services) through the Internet is clearly regulated by legislative acts. On the other hand, each of these sectors operates with a different terminological apparatus, which defines and classifies online trading tools according to different characteristics and properties. Consequently, if civil, antitrust, and tax legislation act uniformly with respect to classic forms of trade, applying equally to each economic entity or consumer, then when relations are complicated by an «online element,» the very same legal relations are regulated differently depending on how the relevant law defines a digital platform and whether the specific online tool under consideration falls under this definition. This not only hinders the implementation of the major principles of the market economy — equality, freedom of trade, and competition — but also allows the exploitation of regulatory loopholes to evade government oversight. Other branches of law do not operate with special terminology at all, regulating platform trading on a case-by-case basis using casuistic regulatory prescriptions.

The most obvious solution to this problem is the development of a specialized federal legislative act that would consolidate and bring to

a common terminological denominator all the principles, norms, and institutions relating to the basic issues of regulating the digital platform market. This should be a law on the foundations of legal regulation of digital platform activities, aimed at systematizing legislative approaches to regulating digital commerce and at ensuring comprehensive streamlining of relations not only in this market segment, but also in the market in general within the idea about demonstrated above trends of global economic influence of digital platforms.

The regulatory core of a full-fledged legal institution is general principles of law [Frolova E.A., 2023: 200–202], and the law of digital platforms will be no exception to this pattern, since with the help of such principles individual legal rules acquire a normative value-goal-setting relationship, necessary both for their adequate joint impact on public relations and for the qualitative interpretation of these rules in law enforcement. The principles of a specific legal institution should be identified as based on the adaptation of the general principles of law to the particularities of relations arising in a specific field (in the case under research, the adaptation of the general principles of civil and commercial law to the present-day outcomes of the experience gained from the implementation of digital platforms in market relations). Clarification of these principles is necessary, since each individual norm that will be included in the consolidated legislative act must originate from and comply with them. Otherwise, it will be extremely problematic to achieve the necessary systemic regulatory effect. Considering the analysis carried out in the present study, the principles of digital commerce and platform work may include: transparency and legality of the digital platform market; equality of participants in the digital platform market; recognition and protection of consumer rights; protection of competition also; a combination of state principles of regulation and self-regulation of digital commerce and platform work; the development of the platform economy within a overall structure of the national economy.

Since the key issue in this article is precisely the regulation of digital platform activities, it is necessary to focus immediately on the designated *primum non nocere*. The latter implies the creation of a regulatory system in which state regulation is combined with self-regulation to be able to both restrain platforms from opportunism and not hinder business development. Where are these boundaries?

With regard to the state part of regulation, the key role in this case should rather relate to the powers of federal state authorities (systematic

interpretation of clauses «e,» «j,» and «o» of Article 71 of the Constitution of the Russian Federation), since the digital platform market, especially large platforms, is common for all the regions of Russia and there can be no a priori regional specifics that would affect the core regulation of the relevant relations in different regions. For example, the powers in question may include establishing minimum technical requirements for digital platforms and introducing rules for identifying digital platform users in order to maintain the stability and fairness of civil turnover as well as the reliability and validity of transactions (this also solves the tasks of the legislator in the field of tax compliance and information security). In addition, within the framework of designing the powers of state authorities, there can be considered the need to create a consolidated register of digital platforms, a state information system of digital platforms to promote reliable and safe interaction between digital platform owners (operators) and the state, on the one hand, and operators, the state and the user contractors, on the other.

Self-regulation can be expressed inversely as the calculation of the degree of state intervention that is optimal for a particular national market to achieve the previously identified goals (stability, fair trade, security, and so on), beyond which not the state, but the market institutions themselves begin to act (within the framework of their self-regulation system). This issue arises most sharply in the sphere of social public relations in the digital platform market most associated with state intervention and the restriction of the boundless desire of business for profit — in the sphere of antitrust regulation. In addition to the described above experience of South Korea, whose competitive situation in the platform market is similar to that in Russia, a few facts must be considered. Platforms providing wide access to the customer base for suppliers and sellers create both opportunities and risks of their business dependence on these market participants. This relationship — or exactly, its potential risks — often explains the state's rigid antitrust position. However, in accordance with the principle of *primum non nocere*, it is necessary to find a balance between antitrust regulation, which restrains abuses, and opening a door of opportunities for business development.

The studies on the subject rightly emphasize the complete absence of a specialized regulatory framework to suppress anti-competitive practices of digital platforms is as detrimental to the market situation as the presence of a gap or an unsystematic legal institution, since this contributes to abuse of a dominant market position on the part of digital platforms [Egorova M.A., Petrov A.A. et al., 2022: 329–343]. Such an anti-

competitive situation in the market may be characterized by the rapid growth of some entities with the absorption of others, which leads to the concentration of market power in the hands of one or a few large platforms and, accordingly, the emergence of digital «giants.» Digital platform operators may, for theirs and often self-serving purposes, contribute to the creation of unfair competition in relation to any mass segment of companies, and competitors may collude with each other, which in the end may lead to global instability of the digital market [Strowel A., Vergote J., 2017].

Nevertheless, a rigid antitrust position does not always contribute to an adequate response to the real market situation, taking into consideration all relevant circumstances. Thus, self-preferencing behavior/practices are restricted or prohibited by the antitrust legislation of a number of states, nevertheless, the impact of such behavior on the consumer market and the competitive environment is not unequivocally negative (Cheng Y., Deng F., 2023: 20–27). In particular, self-preferencing on a platform may involve competition between the platform itself and the sellers represented on it who offer similar goods or services, in this case competitive pressure is manifested in lower prices, designed to attract consumer flow and increase sales.

The policy of limiting the amount of platform commission fees is also ambiguous in its nature. A study by two specialists [Li Z., Wang G., 2024], based on data from the three largest delivery platforms Door Dash, Grubhub and Uber Eats in combination with some additional data, shows: the policy of reducing commission fees for independent restaurants, despite the declared goal of stimulating small businesses and competition, led to a decrease in orders and revenues of such restaurants compared to chain establishments. The result is due to the strategic response of delivery platforms to the regulation of commission fees. There was a decrease in the frequency of recommendations for independent restaurants by the platform and simultaneous active promotion of chain restaurants, which pay higher commission fees. In addition, due to reduced commission fees, platforms increased delivery charges in the cities where the fee limitations were in force.

As an example of unfair competition by platform operators one of the most illustrative cases in China may be cited. Thus, in 2008, Alibaba has blocked the Baidu, Google and Yahoo search engines and did not allow Baidu to display the internal pages of its Taobao platform [Fei L., 2023: 1–11]. Later, in 2013, Alibaba has suspended third-party applica-

tions associated with We Chat, disabled all its data transfer interfaces, and prohibited Taobao sellers from posting We Chat QR codes. However, in 2021, the Ministry of Industry and Information Technology of the People's Republic of China has demanded that Internet platforms unblock external links. In the same year, Alibaba was fined $2.8 billion (approximately 4% of the company's annual revenue) by the antitrust regulator of China for abusing its dominant position over competitors. This antitrust initiative was a response to a series of blockades by Alibaba aimed at third-party applications, as well as external links to these applications. Thus, on the one hand, in China platform companies are required to develop their own operating rules [Afina Y. et al., 2024], but, on the other hand, some of their rules may attract unfavorable attention from the government.

The analysis of the situation with digital platforms in China reveals regional asymmetry in the degree of state regulation. According to the data available (Yang G. et al., 2022), the Western regions of China have a higher degree of state intervention than the Eastern ones. From the view of competition theory, this may lead to increased price rivalry between large players in the West, but an increased likelihood of monopolization in the East and abuse of large players. Therefore, it is advisable for the eastern regions to strengthen supervision over large platforms, while for the western and central regions the goal may be to create more favorable conditions for the development of platform economy and self-regulation.

These examples clearly show the initially stated dichotomy of state intervention and self-regulation, whose optimal boundary is extremely difficult to find even in the most advanced digital economies. Excessively rigid regulation is able to contribute into a reduction in the market share of a certain platform and a decrease in the quality of its functioning. At the same time, with a single manifestation of the regulatory weakness on the part of the state, businesses, by virtue of their very nature, will immediately use their opportunity to extract more profit within the framework of not formally prohibited, but unfair and anti-competitive practices.

Consequently, the question of self-regulation of the digital platform market is synonymous with the theme of their qualitative self-development, which is also beneficial to society, since only this development guarantees the absence of stagnation, the multiplication of benefits and the overall prosperity of the economy. But the system of such self-reg-

ulation must be carefully thought out and consciously introduced into the normative concept of the law of digital platforms. The state should clearly indicate its interest in the development of the platform economy, with minimal intervention in the development of digital platforms, provided that the established necessary requirements and guarantees are observed, and platforms should be given the opportunity to independently establish the basic rules of e-commerce through self-regulation, beyond the subject of state regulation. In other words, as the analysis of the three stages of the evolution of e-commerce regulation, the main principle of regulation should not be fanatical, all-pervasive technocracy, but the principle of *primum non nocere*, suggesting degree of regulatory impact should be sufficient, but so that a sufficiently large field of opportunities for constructive market development remains outside its scope. Proceeding from this general value principle, all other conceptual points of regulation of digital platform activities can be considered.

Considering the mentioned, to exclude excessive state intervention and to allow the market to actively develop in step with the rapid development of digital technologies and related business techniques, but to secure the openness and transparency of the processes in this market according to the rules established by the state, it is advisable to legalize a system of self-regulation for market participants, delegating to this system a number of significant functions.

The advanced expertise in specific branches of the real sector will allow participants in a self-regulatory institution to develop additional regulatory models suitable for a specific market, considering the area specifics (for example, a community of digital platforms in the pharmaceutical industry or in the field of remote medical services). In addition, one of the most important functions of the self-regulatory institution should be the resolution of disputes between owners, operators and user-contractors of digital platforms. The market participants themselves will be able to develop a fair mechanism for resolving disputes and ways to ensure claims satisfaction without resorting to judicial remedies. This is in line with the general trend towards the development of a system of mediation and alternative dispute resolution, encouraging amicable settlement of disputes in pre-trial proceedings. In addition, the Ministry of Justice of Russia has recently directly indicated the relevance of introducing alternative mechanisms for remote dispute resolution in the field of online commerce and online services and a corresponding bill has been drafted.

The most obvious model of self-regulation may be an institution like a digital platform council — a non-profit organization based on the membership of digital platform owners (operators) supporting representation of user contractors and user customer associations. There are institutions of this kind in foreign jurisdictions; in some cases, the domestic legislator has already delegated the development of relevant norms to non-governmental institutions, for example, in the sphere of innovative research and technological centers.

The next fundamental question that requires primary conceptual understanding before developing targeted regulatory prescriptions is platform employment. If we consider digital platforms as a way of employment, it is obvious that they often act as the main source of income for freelancers, who today could freely offer their services on the market through the global network, often with the help of several platforms at once. Nevertheless, due to the gaps in the law in this area, people find themselves facing the risk of unpaid work, when the contractor may not receive remuneration from the customer for their activity on the platform. At the same time, looking at the opposite side of the issue, there is no clear or effective procedure to ensure the tax burden on the part of freelancers and to guarantee the quality of work or services performed by them.

In light of the above, it is interesting to mention the most disorganized form of platform employment is the activity of information digital aggregators (classifieds), thus, while developing future special regulation, the greatest attention should be paid to the principles and rules of employment through their mediation.

Transactions through classifieds fall mainly under the general provisions of Civil Law on the sale of goods, works and services. However, the effect of these norms in relation to the participants in the classified market is not guaranteed. Formally, many users sell goods (works, services) according to the «personal contact» model, although in reality the search for counterparties and preparation for the transaction takes place through the digital platform using the advantages they offer, affecting both the volume and the pace of sales. In fact, classifieds become for many individuals a source of regular earnings and a form of main occupation, providing a platform for targeted, continuous and «smart» search for counterparties in real time. However, the regulation of classifieds does not correspond to the nature of their activities, since the absence of a clear system for identifying users and recording contacts and

stages of their interaction allow them to exchange benefits as if the relationship between the participants in the turnover arose randomly on a one-time basis. Among other things, this violates the principle of equality, since the users of classifieds have unjustified advantages compared to the subjects of the real sector of work and services, who conscientiously work according to the traditional «face-to-face» scheme, paying taxes and answering in rubles for the quality of their work and services.

To solve this problem, when consolidating the legislative regulation of platform work, it is necessary to provide for a more detailed regulation of trade activities and platform work through classifieds, including mandatory procedures for identifying users (using the Unified Identification and Authentication System, bank identifiers or a mobile phone number), as well as the scheme for their remote interaction in order to conclude a transaction. It has a sense to develop specific measures to regulate relations according to the «customer — contractor» model in the framework of interaction on classifieds. The absence of formal employment relations between the customer and the contractor on the digital platform must be viewed from the point of view of the growth of informal economy. At the macro level, this may negatively affect economic growth indicators, such as GDP per capita and labor productivity.

In addition, as researchers rightly point out, digital platforms that have not received a regulatory framework in this area may put their contractors in a deliberately disadvantageous position due to the lack of rules for protecting labor and workers rights (Silberman M.S., Harmon E., 2018: 950]; [Stewart A., Stanford J., 2017). This, in turn, causes instability in the platform economy itself, resulting in increased income gaps in the population (Drahokoupi J., Jepsen M., 2017]; [Healy J. et al., 2017: 246–248]; [Lehdonvirta V., 2018: 24, 26).

Speaking about specific measures, the following conclusions may be suggested to help in solving the identified problems. Platform workers who make a living by providing goods, services and works via digital platforms should be recognized as individuals registered as: individual entrepreneurs, payers of tax on professional income (self-employed) (see point 2 in the Table 1 above), at last as platform workers. Registration (initial identification and authentication) of individuals employed on platforms (user-contractors) should be carried out in ways that promise the reliability of information about them for the purpose of further relations with consumers. This may be registration using a mobile phone number belonging to an individual, as well as, at the choice of the

user-contractor, through: (1) the Unified Identification and Authentication System or (2) other technical means that the appropriate federal executive agency shall determine. Subsequent access of an individual to the platform can be carried out using a mobile phone number belonging to the individual.

Registration of platform employed and maintaining a unified register of platform employed, and providing digital platform operators with access to it may be entrusted to the Federal Tax Service of Russia, to consolidate all information about similar taxpayers (it is also responsible for keeping the Unified State Register of Legal Entities and the Unified State Register of Individual Entrepreneurs, the register of the self-employed). The fee for registration as a platform employed may be paid on principles similar to the patent taxation system, i.e. the amount of the fee for registration as a platform employed is differentiated according to the type of activities. At the same time, the platform fees should be lower than the fees under the patent taxation system, to stimulate participants in this new sphere of the market. Even if concluding an agreement on the use of another digital platform, re-registration of a platform employed who has already been included in the unified register is not required, otherwise the excessive administrative barrier may hinder market development. The possibility should be ensured to verify data-personal identifiers-taxpayer identity through the exchange of information and digital interaction of the Federal Tax Service with digital platforms, which is fully consistent with the previously designated general vector of *primum non nocere*.

For the efficacy of public administration processes, a balance of public and private interests has to be constantly maintained (tax interests of the state, quasi-labor interests of the platform employed, consumer interests of citizens). Thus, digital platforms should report on all transactions made by user-contractors, with a frequency established by law, exchanging information with tax authorities through digital services.

Finally, a question similar in nature to issues related to antitrust regulation is the need to moderate the content on digital platforms, i.e., preventing and suppressing the monopoly on the dissemination of information. As far as content is concerned, one can have in mind both moderation by the state and the possibility of moderation by the digital platforms themselves. For example, in the United States platforms can independently establish their internal regulations and therefore bear minimal legal responsibility for the actions of users and the information

they post. In addition, websites and online services are not responsible for third-party content and their decisions on content filtering.

Conclusion

The diachronic analysis of Russian and foreign experiences reveals that the integration of the platform economy into present-day legal frameworks presents a number of unresolved challenges. Addressing these issues requires continued and coordinated efforts of the state, society, the researchers and expert communities, businesses, and consumer groups. Only through collaborative engagement may be balanced legal solutions be developed for the evolving realities of the platform economy. The vector proposed in this article stems from the need to achieve the goals of stability, transparency, security and permanent development of the platform economy as a new way (form) of organizing the market, taking into consideration a number of core patterns found in Russia and abroad. This vector does not claim absolute accuracy and infallible truth, but its direction is unequivocally characterized by the desire to find a general balance for the common good: a balance of public and private interests, a balance of state intervention and self-regulation, a balance of conservative security guarantees and the legitimate pursuit of progress.

Due to the prevailing legal uncertainty in the regulation of platform-based economic activities and the increasing significance of digital platforms, this article proposes a conceptual framework for the legal regulation of the platform economy. This framework is grounded in a typology of key definitions and the systematic analysis of foreign experiences, adapted to context of the Russian legal system. The peculiarity of the approach lies in the comprehensive approach to defining and classifying various types of platforms, and in the consideration the interests of all stakeholders. Recognizing the platform economy's exponential growth, this legal approach offers a flexible framework and clear rules for all participants, without getting bogged down in sector-specific details.

The negative consequences of rigid regulation of digital platforms initiated at the current stage of legislation development, as well as the negative effects of the absence of regulation in some areas of the platform market, still have to be assessed economically. Nevertheless, one thing is already clear — a legislative solution to consolidate comprehensive regulation within the institution of digital platform law is urgently needed, it is in line with the historic trends in the development of the platform economy, and such a solution is a matter of time.

The conducted research offers a basis for discussion on the basic principles that are likely to have positive fruits and therefore can be used as a conceptual basis for further legislative developments in this area. First, without a legally correct and uniform terminological understanding of the phenomena in the platform market, it is impossible to begin conceptualizing potential principles of legal regulation, since these principles assume a focus specifically on the subjects (platforms and their types) and objects (services, employment, etc.) that have not yet been legally defined in the Russian legal system. Secondly, the principles that will be laid down in the conceptual basis of consolidated regulation must correspond to the main tracks of the analysis of the experience accumulated to date in the development of e-commerce and the platform economy: protection of competition, protection of consumer rights, protection of personal data, and platform employment.

In the present study, a conceptual approach to the legal regulation of the platform economy is proposed; it is based on the key terms typology (digital platform, marketplace, classified, platform operator, etc.) and the systematization of foreign experience (for example, the EU experience in regulating digital services and digital markets), adapted to Russian conditions.

Further efforts should focus on the development of regulatory mechanisms — mandating transparency in ranking algorithms, introducing platform liability for the dissemination of inaccurate information, and establishing robust dispute resolution processes between platform participants. This study aims to provide foundation for such efforts, facilitating the formation of a holistic and positive system of legal regulation of the platform economy in Russia, considering both current and future challenges and opportunities associated with the development of digital technologies.

References

1. Afina Y., Buscher M. et al. (2024) Towards a global approach to digital platform regulation: preserving openness amid the push for Internet sovereignty. Research Paper. London: Royal Institute of International Affairs, 63 p. DOI: https://doi.org/10.55317/9781784135935

2. Cheng Y., Deng F. (2023) Enhancing antitrust analysis of digital platforms: what can we learn from recent economic research? *Antitrust*, vol. 37, no. 3, pp. 20–27. URL: https://www.americanbar.org/content/dam/aba/publications/antitrust/magazine/2023/vol-37-issue-3/enhancing-antitrust-analysis-digital-platforms.pdf

3. Deldjoo Y., Jannach D. et al. (2024) Fairness in recommender systems: research landscape and future directions. *User Model User-Adap Inter*, no. 34, pp. 59–108. DOI: https://doi.org/10.1007/s11257-023-09364-z

4. Drahokoupil J., Jepsen M. (2017) The digital economy and its implications for labor. *European Review of Labor and Research*, vol. 23, no. 2, pp. 103–107. DOI: https://doi.org/10.1177/1024258917701380

5. Egorova M.A., Petrov A.A., Kozhevina O.V. (2022) The impact of digitalization on the implementation of antimonopoly regulation and control over economic concentration in the high-tech sector. *Zhurnal Sankt-Peterburgskogo Universiteta*. *Pravo*=Saint Petersburg State University Journal. Law, vol. 13, no. 2, pp. 327–343. DOI: https://doi.org/10.21638/spbu14.2022.203 (in Russ.)

6. Fei L. (2023) Regulation under administrative guidance: The case of China's forcing interoperability on digital platforms. *Computer Law & Security Review*, no. 48, pp. 1–11. DOI: https://doi.org/10.1016/j.clsr.2022.105786

7. Frolova E.A. (2023) Principles of law: term and implementation. *Gosudarstvo i pravo*=State and Law, no. 1, pp. 200–202 (in Russ.)

8. Haggard S., Tiede L. (2011) The rule of law and economic growth: where are we? *World Development,* vol. 39, no. 5, pp. 673–685.

9. Hayek F.A. (1944) The road to serfdom. London: George Routledge & Sons, 296 p.

10. Healy J., Nicholson D. et al. (2017) Should we take the gig economy seriously? *Labor and Industry*, vol. 27, no. 3, pp. 232–248. DOI: https://doi.org/10.1080/10301763.2017.1377048

11. Heimburg V., Wiesche M. (2023) Digital platform regulation: opportunities for information systems research. *Internet Research*, vol. 33, no. 7, pp. 72–85. DOI: https://doi.org/10.1108/INTR-05-2022-0321

12. Hossain M.B. et al. (2022) Exploring key success factor for sustainable E-commerce adoption. *Polish Journal of Management Studies*, vol. 25, no. 1, pp. 162–178. DOI: https://doi.org/10.17512/pjms.2022.25.1.10

13. Ivanov A.Yu. (2019) Digital economy and antitrust law: unity and conflict of opposites. *Pravovedenie*=Studies in Law, vol. 63, no. 4, pp. 486–521 (in Russ.)

14. Kenney M. et al. (2016) The rise of the platform economy. *Issues in Science and Technology*, vol. 32, no. 3, pp. 61–69.

15. Lafuente E. et al. (2024) Analysis of the digital platform economy around the world: a network DEA model for identifying policy priorities. *Journal of Small Business Management*, vol. 62, no. 2, pp. 847–891. DOI: 10.1080/00472778.2022.210089

16. Lehdonvirta V. (2018) Flexibility in the gig economy: managing time on three online piecework platforms. *New Technology, Work and Employment,* vol. 33, no. 1, pp. 16–29. DOI: https://doi.org/10.1111/ntwe.12102

17. Li Z., Wang G. (2024) On-Demand Delivery Platforms and Restaurant Sales. *Management Science*. https://doi.org/10.1287/mnsc.2021.01010

18. Lofstedt R.E. (2023) The precautionary principle: risk, regulation and politics. *Process Safety and Environmental Protection*, vol. 81, no. 1, pp. 36–43. https://doi.org/10.1205/095758203762851976

19. Nersesyants V.S. (2003) History of political and legal theories. Textbook. Moscow: Norma, 944 p. (in Russ.)

20. North D.C. (1990) *Institutions, institutional change and economic performance. Political economy of institutions and decisions.* Cambridge: University Press, 153 p.

21. Ortolani A. (2024) The Civil law. In: M. Siems, P. Yap (eds.) The Cambridge Handbook of Comparative Law. Cambridge: University Press, pp. 211–234.

22. Paun C., Ivascu C. et al. (2024) The main drivers of e-commerce adoption: global panel data analysis. *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 19, issue 3, pp. 2198–2217. DOI: https://doi.org/10.3390/jtaer19030107

23. Rylova M.A. (2014) The precautionary principle as harm prevention and (or) protection of economic interests: the European legal experience. *Vestnik Moskobskogo gosudarstvennogo universiteta*. *Pravo*=Moscow State University Bulletin. Law, pp. 30–42 (in Russ.)

24. Schulte-Nölke H., Rüffer I. et al. (2020) The legal framework for e-commerce in the internal market. Brussels: European Parliament Printing, 43 p. URL: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652707/IPOL_STU(2020)652707_EN.pdf

25. Shelepov A.V., Kolmar O.I. (2024) Regulation of digital platforms in Russia. *Vestnik mezhdunarodnukh organizatciy*=International Organizations Bulletin, vol. 19, no. 2, pp. 110–126 (in Russ.)

26. Silberman M.S., Harmon E. (2018) Rating working conditions on digital labor platforms. *Computer* Supported *Cooperative Work*, vol. 28, issue 5, pp. 911–960. DOI: https://doi.org/10.1007/s10606-018-9313-5

27. Stewart A., Stanford J. (2017) Regulating work in the gig economy. *The Economics and Labor Relations Review*, vol. 28, issue 3, pp. 420–437. DOI: https://doi.org/10.1177/10353046177224

28. Strowel A., Vergote W. (2017) Digital platforms: to regulate or not to regulate? Fix the economics first, then focus on the right regulation. *Research Observatory on Sharing Economy, Law and Society*, 16 p. URL: http://www.rosels.eu/2017/11/15/digital-platforms-to-regulate-or-not-to-regulate-message-to-regulators-fix-the-economics-first-then-focus-on-the-right-regulation/

29. Yang G., Deng F. et al. (2022) Digital paradox: platform economy and high-quality economic development—new evidence from provincial panel data in China. *Sustainability*, vol. 14, no. 4, pp. 22–25. DOI: https://doi.org/10.3390/su14042225

30. Zap'yantsev A.A. (2024) Creation and regulation of digital platforms in China. *Aziya i Afrika segodnia*=Asia and Africa Today, no. 4, pp. 55–60. DOI: http://doi.org/10.31857/S032150750027789-9 (in Russ.)

31. Zhai K. (2021) Alibaba hit with record $2.8 billion antitrust fine in China. *The Wall Street Journal*. April 10.

32. Zhukov V.N., Frolova E.A. et al. (2024) Theory of state and law. Textbook. Moscow: Prospekt, 640 p. (in Russ.)

Information about the authors:

A.S. Koshel — Doctor of Sciences (Law), Associate Professor, Professor of Department of Public Law, Higher School of Economics (HSE University).

Ya.I. Kuzminov — Candidate of Sciences (Economics), Associate Professor, Academic Supervisor of the Higher School of Economics (HSE University).

E.V. Kruchinskaia–Lecturer, Department of Politics and Governance, Higher School of Economics (HSE University).

B.V. Lesiv–Candidate of Sciences (Law), Associate Professor, Department of Theory of Law and Comparative Law, Higher School of Economics (HSE University).

The article was submitted to editorial office 10.04.2025; approved after reviewing 12.05.2025; accepted for publication 15.05.2025.

Legal Issues in the Digital Age. 2025. Vol. 6, no.2.
Вопросы права в цифровую эпоху. 2025. Том 6. № 2.

Artificial Intelligence and Law

Research article
JEL: K00
UDK: 34
DOI:10.17323/2713-2749.2025.2.50.68

Shaping Artificial Intelligence Regulatory Model: International and Domestic Experience

Vladimir O. Buryaga[1], Veronika V. Djuzhoma[2], Egor A. Artemenko[3]

[1, 2, 3] National Research University Higher School of Economics, 20 Myasnitskaya Str., Moscow 101000, Russia,

[1] buriaga@mail.ru, ORCID 0009-0005-4796-2797

[2] vdzhuzhoma@mail.ru, ORCID 0004-0008-7446-3557

[3] artemenkoea@gmail.com, ORCID 0000-0002-6874-620X

Abstract

The article contains an analysis of AI regulatory models in Russia and other countries. The authors discuss key regulatory trends, principles and mechanisms with a special focus on balancing the incentives for technological development and the minimization of AI-related risks. The attention is centered on three principal approaches: "soft law", experimental legal regimes (ELR) and technical regulation. The methodology of research covers a comparative legal analysis of AI-related strategic documents and legislative initiatives such as the national strategies approved by the U.S., China, India, United Kingdom, Germany and Canada, as well as regulations and codes of conduct. The authors also explore domestic experience including the 2030 National AI Development Strategy and the AI Code of Conduct as well as the use of ELR under the Federal Law "On Experimental Legal Regimes for Digital Innovation in the Russian Federation". The main conclusions can be summed up as follows. A vast majority of countries including Russian Federation has opted for "soft law" (codes of conduct, declarations) that provides a flexible regulation by avoiding excessive administrative barriers. Experimental legal regimes are crucial for validating AI applications by allowing to test technologies in a controlled environment. In Russia ELR are widely used in transportation, health and logistics. Technical regulation including

standardization is helpful to foster security and confidence in AI. The article notes widespread development of national and international standards in this field. Special regulation (along the lines of the European Union AI Act) still has not become widespread. A draft law based on the risk-oriented approach is currently discussed in Russia. The authors of the article argue for the gradual, iterative development of legal framework for AI to avoid rigid regulatory barriers emerging too prematurely. They also note the importance of international cooperation and adaptation of the best practices to shape an efficient regulatory system.

Keywords

artificial intelligence; technology; principles; statutory regulation; strategy; experimental legal regime; soft law; technical regulation.

For citation: Buriaga V.O., Djuzhoma V.V., Artemenko E.A. (2025) Shaping an Artificial Intelligence Regulatory Model: International and Domestic Experience. *Legal Issues in the Digital Age,* vol. 6, no. 2, pp. 50–68. DOI:10.17323/ 2713-2749.2025.2.50.68

Background

The development of artificial intelligence (hereinafter AI) for use in various spheres of social life is a major factor of modern economic progress. The tasks to introduce and promote AI technologies have been defined in strategic governmental documents in many countries including Russia.

Rapid development and penetration of AI in different spheres of government and society have not only positive effects but also downsides. An important issue in this regard is to provide adequate legal mechanisms to regulate social relations associated with AI design, its development and implementation [Bourcier D., 2001: 853]; [Talapina E.V., 2020: 27].

On one hand, countries should provide the environment and incentives for AI development and, on other hand, minimize risks associated with the use of these technologies. Thus, there is a need for regulatory balance.

Moreover, regulation should be responsive to rapid AI progress and envisage tools for integrating new technologies into community life swiftly and seamlessly.

Since regulation governing AI is still emerging in a majority of countries, AI development strategies and plans prevail. It is of interest to analyze their content in comparison with domestic regulation.

1. Domestic and International Approached to AI Regulation

In 2016, the United States has approved a National AI R&D Strategic Plan for human-AI collaboration and long-term investments to ensure security and to address ethical, legal and societal implications[1]. In 2023, the plan was updated with a focus on AI-related R&D.

In 2017, China has passed a New Generation AI Development Plan to regulate AI introduction[2], with AI recognized as crucial for the development of national research and technology. The plan contains strategic objectives for introducing AI in health care, smart cities, national defense and agriculture, with China poised to achieve global leadership in AI by 2030.

In 2018, India has approved an AI Development Strategy in prioritizing five key areas for AI introduction: health care, education, agriculture, infrastructure (including smart cities) and transportation.

The 10-year National AI Strategy in force in the United Kingdom (passed in December 2022)[3] describes key actions to assess long-term risks associated with AI. The strategy aims to unlock AI power for innovative economy, create more jobs, improve the infrastructure and business environment. While being of general nature, the document outlines AI development vectors.

In 2018, Germany has approved a federal level AI Strategy[4] to boost the national competitiveness in this field and make sure that AI is used in the interest of society by observing statutory provisions, ethical standards and cultural values. Currently, the relevant strategies have been passed at the regional level in 5 out of 16 federal lands (states)[5]. The

[1] National AI R&D Strategic Plan: 2023 Update //Available at: https://bidenwhitehouse.archives.gov/wp-content/uploads/2023/05/National-Artificial-Intelligence-Research-and-Development-Strategic-Plan-2023-Update.pdf (accessed: 29.04.2025)

[2] Available at: https://flia.org/wp-content/uploads/2017/07/A-New-Generation-of-Artificial-Intelligence-Development-Plan-1.pdf (accessed: 29.04.2025)

[3] Available at: https://www.gov.uk/government/publications/national-ai-strategy/national-ai-strategy-html-version (accessed: 04.04.2025)

[4] Artificial Intelligence Strategy of the German Federal Government // Avai lable at: https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf. (accessed: 04.04.2025)

[5] OECD Artificial Intelligence Review of Germany.

German AI standardization roadmap approved in 2020 specifies the stages of AI standardization to promote more competitive research and to create an enabling environment for AI innovations. The roadmap's effective version[6] covers both the main sectors, first of all are health care, transport, energy, environment, financial services, industrial automation, and fundamental issues including AI classification, security, certification, socio-technical systems and ethics. Moreover, the document defines the main AI-related terms, covers a total of 116 standardization needs and contains 5 regulatory priorities: uniform normative approach to AI standardization including the adoption of a framework regulation; harmonizing the national legislation with European law; AI security requirements; flexible AI regulation; and minimizing the risks of AI misuse.

In 2017, Canada has passed a Pan-Canadian AI Strategy. To implement it, the Federal Government has appointed in May 2019 the AI Advisory Board to include the representatives of high-tech businesses and AI application developers.

Thus, the above countries regulate AI by establishing uniform principles for AI implementation, prioritizing specific sectors and specifying national objectives for the development of promising technologies.

In Russia, the main high-level document to identify AI development vectors and parameters is the 2030 National AI Development Strategy approved by Presidential Decree No. 490 of 10 October 2019 (hereinafter "Strategy"). At the strategic level, a comprehensive, decentralized regulatory system is envisaged as a logical step since no single federal framework law could currently regulate multiple AI technologies while artificial drafting of such a law would hold back technological development.

The Strategy also sets the task of creating favorable regulatory environment for AI design, development and use, something that requires to maximize informal, flexible and generally accepted regulation. Another crucial vector is avoidance of administrative barriers and a focus on the best international regulatory practices. The introduction of ethical standards for AI is also in focus. Such a comprehensive regulatory approach will generally put the principle of collaboration between man and AI technologies at the heart of regulation.

Balanced regulation is supposed to maintain a balance between protection of human rights and liberties, personal and national security, on

[6] Available at: https://www.dke.de/resource/blob/2008010/11faae856dd4332e5a5c62f3447fd06f/nr-ki-deutsch---download-data.pdf (accessed: 05.04.2025)

the one hand, and AI development incentives, on the other hand; regulation should not slow down the pace of development and implementation of new technologies.

The abundance of provisions (principles) and high-level regulatory focus do not mean that the Strategy is devoid of crucial practical importance: provisions are regularly updated while the reference to specific AI technologies means explicit recognition of their status by the government, something that allows their authors to qualify for public support, preferential tax regimes, etc.

In furtherance of the Strategy, the Federal Government has approved the 2024 Regulatory Development Concept Note for AI and Robotics, Resolution No. 2129-r of 19 August 2020 (hereinafter Concept Note). While both the Concept Note and the Strategy serve a general purpose, the former is more focused on security of AI applications, the need to harness AI for higher economic growth, security and living standards. As a crucial conceptual aspect, the Concept Note argues for an incentivizing regulatory regime and non-acceptability of using AI for regulatory restrictions in the future.

Thus, the discussed documents assume cautious and consistent application of rules and provisions, and "cascading" regulation via interrelated instruments updatable on a regular basis.

Overall, Russia's current AI regulatory system exhibits the following trends:

"soft law" used as a regulatory mechanism for the institution in question;

expanded use of experimental legal regimes;

progressing technical regulation of artificial intelligence.

2. Ethics and "Soft Regulation" of Artificial Intelligence

Alternatives to statutory regulation become crucial for striking a regulatory balance to avoid excessive government intervention into AI usage scenarios.

One such alternative appeared to be ethical standards that regulate the relations between human person and AI. Moreover, ethical standards should prioritize human-centric approach, with human security, wellbeing and avoidance of harm at its core.

Internationally, ethical standards are a major component of AI regulation. In 2021, China has issued ethical guidelines for AI use in China that require researchers to make sure that AI technologies are consistent with universal human values, are under human control and do not put public security at risk[7]. The UK's Centre for Data Ethics and Innovation drafts recommendations for safe, ethical and innovative implementation of AI applications[8].

Unlike statutory provisions, ethical standards are advisory and make part of the so-called "soft law" [Kashkin S. Yu., 2021: 193].

Using "soft law" in AI is objectively necessary at this stage [Antonova L.I., Korneva K.A., 2022: 37], because it allows the government to identify the overall development vector that organizations can use to establish rules and requirements through their bylaws.

It is worth noting the basis for soft law in artificial intelligence in Russia was laid back in 2021 with the passing of the AI Code of Conduct[9].

The Code is an advisory document, and its provisions could be updated and complemented for specific AI fields, actors, etc[10].

The Code has six core principles of AI development and implementation:

1. Human rights are the main AI development priority. Human rights and liberties should constitute a supreme and undisputed value, with AI to consistently observe the humanistic approach and contribute to human development. AI cannot challenge human will, deprive man of a choice, contribute to negative implications for man. AI actors should be aware of and consistently abide by AI regulation. Discrimination of any kind is prohibited in respect of AI use, with the risks of human right violation to be assessed before AI is introduced.

2. Responsible AI introduction meaning, in particular, the introduction of an AI-related risk management system (incorporating relevant evaluation standards and methodologies), the possibility to forecast and

[7] Available at: https://cset.georgetown.edu/ (accessed: 20.03.2025)

[8] Available at: https://www.gov.uk/government/organisations/centre-for-data-ethics-and-innovation. (accessed: 05.04.2025)

[9] Available at: URL: https://ethics.a-ai.ru/ (accessed: 20.03.2025)

[10] As defined by the Code, AI actors mean parties to relations associated with AI (AI system developers, manufacturers, operators, experts, customers, persons associated with regulatory impact on the field etc.).

avoid negative AI implications, rule out any intentional harm through AI use, ensure transparency and openness of AI applications (man should be always aware of the contact with AI technologies). Responsible AI introduction equally assumes information security and protection, voluntary certification, and a possibility to identify and timely highlight AI that has evolved into "strong" systems.

3. Since man is always responsible for AI implications, AI should be put under control and man held liable (AI should never make any managerial decisions or moral choices).

4. AI mechanisms should be harnessed for maximum public benefit.

5. AI development should avoid unfair competition, maintaim transparency of information on AI technologies (including a uniform system of measurements), improve skills and collaboration of AI developers for higher security, quality and availability of technologies.

6. A crucial principle is provision of credible and open information on AI technologies being introduced including their security, potential and AI-related risks that may arise.

While the Code is voluntary to abide by, its adoption can result in normative benefits (if it is a precondition for government support for AI development and introduction). Moreover, AI actors can use specific provisions of the Code to draft bylaws and documents, and to shape the conditions for cooperation with different counterparties.

The Implementation Commission, established in 2022, monitors the Code's performance and compliance with its provisions. The Code is currently adopted by 335 business entities, by 21 federal and regional level agencies, and by 50 international parties[11].

In addition to the Code, two declarations were approved in Russia under "soft" regulation: on responsible generative AI[12] and on responsible exports of AI technologies and associated software[13]. Both ones contain ethical principles and standards of conduct with regard to AI development and use.

[11] Available at: URL: https://ethics.a-ai.ru/#actors (accessed: 05.04.2025)

[12] Available at: URL: https://ethics.a-ai.ru/assets/ethics_files/2024/03/13/GenAi_Declaration_Ai_Alliance_Russia_FpNJ2Lc_82yB8pD.pdf (accessed: 05.04.2025)

[13] Available at: URL: https://ethics.a-ai.ru/assets/ethics_files/2024/04/24/Декларация_об_ответственном_экспорте_ИИ_cH11Lzg.pdf (accessed: 05.04.2025)

Also, the guidelines for general purpose robots were drafted to provide general ethical principles and recommendations aimed at ensuring compliance throughout the process of developing general purpose robots and associated technologies[14].

A soft regulatory approach is also observed internationally. Thus, in March 2023, the United Kingdom has published a White paper entitled "A pro-innovation approach to AI regulation"[15] that contained the general regulatory principles while providing considerable room for the respective regulators to adapt these principles to specific fields such as transportation or financial markets.

The United States have approved an order on maintaining American leadership in AI[16] whereby the Office of Science and Technology Policy has published a draft memo on AI applications[17] that contained a list of conditions for public agencies to decide how to regulate AI, if at all (principles of openness, transparency, engagement, regulatory flexibility etc.). The memo assumes that not all aspects of AI usage are subject to regulation. Upon review of a specific AI application, a public agency can determine that the present-day rules are adequate, or that the benefits of new regulation do not justify its costs now or in the foreseeable future. In such a case, the agency may want to abstain from action or, alternatively, to come up with non-regulatory approaches that can be feasible to address the risk inherent in AI applications.

It is worth noting that in an AI regulation system "ethics has a potential of a full-fledged regulator of social relationships, along with standardization and law" [Ibragimov R.S., Suragina E.D., Churilova D.Y., 2021: 89]. This observation is especially relevant at this stage since ethics, in regulating social relationships, can underpin a regulatory framework while at the same time acting as a regulator in its own right to exclude the risk of excessive government intervention and to avoid barriers to technological development.

---

[14] Available at: URL: https://ethics.a-ai.ru/assets/ethics_files/2024/12/12/2_Руководящие_принципы_в_сфере_роботов_общего_назначения.pdf (accessed: 05.04.2025)

[15] Available at: https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper (accessed: 05.04.2025)

[16] Available at: https://trumpwhitehouse.archives.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/ (accessed: 05.04.2025)

[17] Available at: https://niso.org/niso-io/2020/02/memo-drafted-federal-omb-innovation-and-use-ai (accessed: 10.04.2025)

Meanwhile, one has to accept an argument advanced by some authors that "the government, by adopting the AI Code of Conduct, has already opted for a "soft" regulatory model in concert with AI companies, large corporations, major universities and the banking sector" [Arzamasov Yu. G., 2023: 138] as traditional regulation is emerging in addition to the "soft" model, only to set the stage for the use of AI technologies without explicitly mentioning them.

Thus, Article 10.2-2 of the Federal Law No. 149-FZ "On Information, Information Technologies and Information Security" of 27 July 2006[18] stipulates the conditions for collection and provision of data to analyze Internet users' preferences where AI technologies can be harnessed to perform such analysis.

According to experts, "traditional regulation can contain some components [of the soft law] where ethical principles are incorporated into the legal language and thus made binding through governmental enforcement" [Popova A.V., 2021: 91], as seen in experimental legal regimes.

3. AI Validation Mechanisms as a Basis of New Regulation

Experimental legal regimes (ELR) serve as a mechanism for validating AI-enabled products to facilitate their introduction in Russia under Federal Law No. 258-FZ "On Experimental Legal Regimes for Digital Innovations in the Russian Federation" of 31 July 2020 ("Law No. 258").

It is worth noting that the ELR mechanism is used across many jurisdictions: while in the U.S. regulatory sandboxes are observed in specific states, Canada uses them to introduce AI in health care and securities markets; India in the financial sector (processing payments and credit requests, and fighting financial fraud) and health care (health insurance); while China has 16 pilot AI development zones[19] for validating

[18] Has obtained force by Federal Law No. 408-FZ "On Amending the Federal Law "On Information, Information Technologies and Information Protection" of 31 July 2023 // SPS Consultant Plus.

[19] Notice of the Ministry of Science and Technology on the Issuance of the Guidelines for the Construction of National New Generation Artificial Intelligence Innovation and Development Pilot Zone (Revised Version) // Available at: URL: https://www.most.gov.cn/xxgk/xinxifenlei/fdzdgknr/fgzc/gfxwj/gfxwj2020/202012/t20201224_171987.html. (accessed: 21.04.2025)

institutional regulatory decisions before they are subsequently upscaled. The United Kingdom applies a flexible approach, with AI regulatory sandboxes launched by the agencies themselves. The country has experimental platforms to support specific projects of introducing AI for various purposes: financial literacy, psychiatric aid, etc. Other experimental legal regimes apply to AI used in air travel and transportation. A focus on promoting regulatory sandboxes for AI applications is also made in a draft framework law on AI implementation currently in progress in the UK.

In Russia, a list of mechanisms and technologies subject to ELR legislation is established by Federal Government Resolution No. 1750 of 28 October 2020 and includes AI and neural technologies (machine learning, computer vision, language models and speech recognition, neural prosthetics etc.), big data processing, quantum computing and manufacturing technologies, robotics, augmented reality, distributed ledger systems etc., for a total of 10 types and about 50 sub-types of AI-related technologies.

Of 16 ELRs under way in Russia, 13 ones concern unmanned vehicles including the use of highly automated vehicles[20] (HAV), in particular, as part of the Unmanned Logistical Corridor initiative in the Neva Highway (M-11)[21] and federal regions[22]; and technologies for collecting data on individual diagnoses and health as part of the Personal Health Assistant socioeconomic initiative[23].

---

[20] Government Resolution No. 309 "On Introducing Experimental Legal Regime for Digital Innovations and Approving the Experimental Legal Regime Programme for Digital Innovations in Highly Automated Vehicles" of 09 March 2022 // SPS Consultant Plus.

[21] Government Resolution No. 1849 "On Introducing Experimental Legal Regime for Digital Innovations and Approving the Experimental Legal Regime Programme for Digital Innovations in Highly Automated Vehicles: Unmanned Logistical Corridor Initiative in the Neva General Purpose Federal Highway (M-11)" of 17 October 2022 //SPS Consultant Plus.

[22] Government Resolution No. 2495 "On Introducing Experimental Legal Regime for Digital Innovations and Approving the Experimental Legal Regime Programme for Digital Innovations in Highly Automated Vehicles in the Territory of Specific Federal Regions" of 29 December 2022 (as amended on 28 April 2023) // SPS Consultant Plus.

[23] Government Resolution No. 2276 "On Introducing Experimental Legal Regime for Digital Innovations and Approving the Experimental Legal Regime Programme for Digital Innovations in Health Care: Use of Technologies for Collecting and Processing Data on Individual Diagnoses and Health in Implementing the Personal Health Assistant Initiative for Socioeconomic Development of Russia" of 09 December 2022 // SPS Consultant Plus.

Federal Law No. 123-FZ of 24 April 2020 envisages AI-related ELR for the City of Moscow. This region-specific approach is found with other federal nations. Thus, specific regulatory sandboxes for AI apply to Arizona[24], Utah and Wyoming in the United States. There are also examples of regulatory sandboxes applicable to specific sectors: a Digital Health Sandbox Program is underway in Massachusetts to harness AI for making clinical simulations, collecting data and improving safety of surgical interventions[25].

Promoting ELR as an institution is crucial in the context of AI technologies introduced under "soft law" as the primary regulatory model since ELR allows to disregard specific provisions standing in the way of innovations and thus to avoid the risk of non-compliance and legal liability.

Over the last few years, Law No. 258 has absorbed important legal novelties that allowed to expand the use of ELR in the field of AI technologies, improve safety of the parties involved, reduce the risks and assess ELR performance.

The law also applies to intellectual property assets ("IPA") produced through the use of AI, as well as to liability insurance of natural and legal persons for the harm resulting from ELR (Federal Law No. 169-FZ of 08 July 2024). Pursuant to Article 14, Law No. 258, a party under ELR must now maintain a register of IPA including assets created through the use of AI.

A major innovation in Law No. 258 is new Article 18.1 providing for a procedure to investigate the harm caused by AI to persons and entities under ELR. In particular, it is envisaged to set up a commission to look into the circumstances that caused such harm.

The procedure for the commission to set up, proceed and issue its opinions is approved by Ministry of Economic Development Order No. 752 of 26 November 2024.

Federal Law No. 331-FZ "On Amending Specific Regulations of the Russian Federation Following the Adoption of the Federal Law "On Experimental Legal Regimes for Digital Innovations" has added part 8 to Article 36.1, Federal Law No. 323-FZ "On the Principles of Protect-

[24] House Bill 2434 // Available at: URL: https://www.azleg.gov/legtext/53leg/2R/laws/0044.pdf (accessed: 21.04.2025)

[25] Available at: URL: https://hitconsultant.net/2019/04/25/digital-health-sandbox-program/ (accessed: 29.04.2025)

ing Public Health in the Russian Federation" whereby the requirements to the ethics committee and the expert council set up by a public agency shall not apply in case of AI-assisted health care under ELR.

Further expansion of ELR is necessary to promote AI regulation including to remove legal and administrative barriers since harnessing AI as a substitute for conventional technologies involves a high degree of innovation, only to result in possible negative implications in absence of a balanced position, comprehensive risk assessment and validated options.

A number of legal novelties in Russia address the issues of AI used as part of specific ELR. In particular, Federal Law No. 152-FZ "On Personal Data" of 27 July 2006 has come to include Article 13.1[26] that details anonymized personal data processing in identifying and providing access to specific data structures.

4. Technical Regulation for AI

Standardization is "a crucial factor of Russia's modernization, technological and socioeconomic development, including the capabilities of its national defense".

One can accept a view that "a whole range of issues related to harnessing AI technologies and marketing AI-enabled outcomes (or the associated rights arising with specific agents) may be addressed by standards" [Kharitonova Y. S., Savina V. S., 2020: 537, 542].

Technical regulation holds a special place among regulatory tools for AI, with Russia having adopted and implemented a number of state standards for AI despite the technology's relative novelty. These include both individual standards to address both specific aspects of AI use across sectors (such as GOST R 70250-2022. AI systems for road vehicles), and also generalized, universally applicable standards (for instance, GOST R 59276-2020. AI systems. Credential assurance methods. General provisions)[27].

[26] Went into force by Federal Law No. 233-FZ "On Amending the Federal Law "On Personal Data" and the Federal Law "On Experimental Regulation to Enable the Development and Introduction of AI Technologies in a Federal Territory (Federal City of Moscow) and on Amending Articles 6 and 10 of the Federal Law "On Personal Data" of 08 August 2024 // SPS Consultant Plus.

[27] SPS Consultant Plus.

Since the associated standardization system is only emerging, some national standards are tentative and time-bound, with a whole set of tentative standards being adopted, for instance, in civil aviation (PNST 783-2022. AI for navigation systems of civil aircraft. General requirements; PNST 787-2022. AI for navigation systems of civil aircraft, etc.). Upon expiry of a three-year effective term and once validated for practical use, the said tentative national standards are likely to be updated to the level of permanent standards.

As for the latest national standards, the following standards approved in 2024 by Rosstandard orders to take effect in 2025 and deserve be noted: GOST R ISO/IEEC 20547-3-2024. Information technologies. Big data reference architecture. Part 3. Reference architecture (approved by Rosstandard order No. 1541-st of 28 October 2024); GOST R 71562-2024. AI-enabled measuring tools. Metrological support. General requirements (approved by Rosstandard order No. 1526-st of 28 October 2024); GOST R ISO/IEEC 24029-2-2024. AI. Neural network robustness evaluation. Part 2. Methodology of formal methods (approved by Rosstandard order No. 1542-st of 28 October 2024); GOST R ISO/IEEC 42001-2024. Information technologies. AI. Control system (approved by Rosstandard order No. 1549-st of 28 October 2024), GOST R 71750-2024. AI-enabled technologies for road construction equipment. Terms and definitions (approved by Rosstandard order No. 1546-st of 28 October 2024); GOST R 71751-2024. AI-enabled technologies for road construction equipment. Usage scenarios (approved by Rosstandard order No 1547-st of 28 October 2024), etc.

These standards give an idea how the standardization system for AI is taking shape. Thus, GOST R ISO/IEEC 20547-3-2024 provides a generalized reference structure for big data to describe the relevant components, processes and systems for standardized design. The standard relies on and takes into account international standards indicating that the national AI standardization system generally follows in the wake of international practice. Such approach is important in the sense that the Russian standardization system is often used as a reference for standards designed by the EEU and other intergovernmental associations involving Russia and, therefore, indirectly impacts AI institutional development elsewhere. Overall, both domestic experience and international sources are used to draft and adopt sector-specific standards: for instance, the aforementioned GOST R 71750-2024, in describing terms and definitions for implementing AI in road construction equipment, relies on domestic experience while GOST R 71751-2024 accounts for international sources

and the experience of harnessing AI to control traffic of road construction equipment. GOST R ISO/IEEC 24029-2-2024 is essentially based on an adapted translation of ISO/IEEC 24029-2:2023 international standard. Thus, in borrowing, adapting and building on what is available elsewhere, national standards allow to upgrade AI's current regulatory regime in such a way that national efforts are in step with the best international practices.

Undoubtedly, the national AI standardization system is ever improving. With ongoing standardization of AI-enabled technologies, international collaboration for developing relevant standards will allow in future to put in place a comprehensive system to regulate the development and introduction of AI applications and their legal effects. To establish a universal standardization system, a technical committee for AI standardization was set up (Rosstandard order No. 1732 of 25 July 2019).

5. Special Regulation for AI

In 2024, the UN has passed resolutions on safe and trustworthy AI systems for sustainable development[28] and on promoting international cooperation to enhance AI potential[29] (aiming to reduce barriers for AI development and provide access to technologies and knowledge), while the Council of Europe has adopted the Framework Convention on AI, human rights, democracy and the rule of law"[30] (containing the basic principles, risk-oriented approach to AI implementation and remedies against AI-related implications including possible moratoria on specific AI applications). In 2024, the CIS developed and conceptually approved a draft framework law for harnessing AI to improve living standards and security and to boost socioeconomic development.

With the exception of the EU's supranational regulation, most countries have no specific regulation for AI at the moment.

The EU's Artificial Intelligence Act[31] passed by the European Parliament on 14 June 2023 is to be gradually applied to different AI system

[28] Available at: URL: https://digitallibrary.un.org/record/4040897?ln=ru&v=pdf (accessed: 21.04.2025)

[29] Available at: URL: https://digitallibrary.un.org/record/4053245?v=pdf&ln=en (accessed: 21.04.2025)

[30] Available at: URL: https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence (accessed: 11.04.2025)

[31] Available at: URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206 (accessed: 21.04.2025)

types: 6 months for prohibited systems, 12 months for general systems, 24 and 36 months for high-risk systems depending on the risk level. In regulating the terms of AI marketing and operation, the AI Act aims to ensure security and legal certainty, and provides for AI control rules and support policies for AI application developers. A major feature of the AI Act is risk-oriented approach to AI systems whereby the extent of regulatory rigor will depend on potential risk and the criteria of unacceptable risk where an AI application will be prohibited and cannot be used (for instance, social scoring, behavior manipulation, etc.). Along with the right to self-regulation for low-risk systems, the AI Act provides for at least one experimental legal regime ("regulatory sandbox") for AI in each of the EU's national jurisdictions in order to ensure more pre-marketing control and testing for AI systems.

While the United States currently have no AI framework act, there are draft laws to regulate AI including machine learning, prohibition of face recognition, etc.

A number of drafts on AI and data (AIDA)[32], protection of personal data and confidentiality are tabled and under discussion in Canada now. The AIDA act purports to establish AI-enabled regulation of international and domestic trade and envisages measures to avoid illegal use of AI technologies, reduce the underlying risks and provide for liability.

There is no special regulation of AI in Russia. Providing for such regulation at the current stage is a matter of academic debate. A number of authors argue that it is crucial to establish the overarching principles of AI implementation now, with the requirements to technologies to be established through bylaws [Sucharev A.N., 2021: 18]; [Minbaleev A.V., 2022: 1098].

An affirmative answer to this question brings forth the following dilemma: "will amendments to the effective regulations suffice or will special law applicable to specific AI aspects be needed or else a codification instrument to govern digital technologies, AI, technological innovations etc.?" [Popova A.V., 2021: 90].

It is worth noting a draft AI regulation is currently discussed in Russia to provide for a risk-oriented approach and introduce new rules for AI developers and operators. The draft was developed within the framework of the 2030 National AI Development Strategy by a working

[32] Available at: URL: https://www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading (accessed: 21.04.2025)

group that included legal scholars, experts and representatives of the IT market[33].

In particular, the draft law is to define a number of concepts such as artificial intelligence, AI technology, AI system; introduce AI system marking requirements; and to classify AI systems by the level of potential risk. This risk-oriented model is supposed to prohibit the development and operation of AI systems associated with unacceptable risk — that is, creating a security threat for individuals, society and government — and violating fundamental human and civil rights and liberties.

Moreover, the draft law defines liability for the harm caused to life, health or property of those involved in the development and operation of AI systems, as well as mandatory liability insurance for operators of high-risk AI systems.

Specific solutions are also proposed with regard to copyright associated with AI-created intellectual property assets. To identify the holder of copyright to such IP assets, it is needed to determine to what extent human creative contribution was essential. Where human contribution was essential, the exclusive right should go to the person in question, otherwise the exclusive right will be held by the AI system operator.

Conclusion

Thus, AI regulation in Russia currently follows in the wake of international trends. The prevailing documents are AI development plans and strategies which determine the main vectors of progress of both technologies themselves and the underlying regulation.

Moreover, Russia, like most countries of the world community, does not have any special regulation of relationships involved in AI design, development and application. In this regard, one cannot accept a point of view that a new institution [Kosykh A.A., 2021: 161]; [Polyakova T.A., Kamalova G.G., 2021: 135] or a new branch of law is already emerging in Russia for artificial intelligence [Mishina N.V., 2020: 64]. Such assessment would be a premature one.

As a general trend, it has been accepted internationally that there is no need for statutory regulation of all aspects of AI use since this would create unnecessary barriers.

[33] Available at: URL: https://www.rbc.ru/technology_and_media/11/04/2025/67f7dc399a79477fdd97bf30?ysclid=ma1a3t4rm5885753225 (accessed: 21.04.2025)

This makes the case for the so-called "soft law" as a substitute for traditional statutory regulation. The same model is used in Russia where a number of advisory documents were approved including the AI Code of Conduct; the latter continues to be approved by both businesses and public agencies.

Apart from "soft law", Russia is pursuing standardization including for safe introduction of AI technologies.

An equally important regulatory mechanism is the emergence of ELR which allow to test AI development and introduction outside the originally established administrative barriers and cumbersome requirements provided that the necessary level of security is observed.

Since AI technologies rapidly permeate many spheres of life, the state as a regulator should respond accordingly.

In this regard, countries make a special emphasis on risks and safety of AI applications, as well as on the resulting liability. Russia is no exception, with a risk-oriented approach underpinning draft regulations that are currently discussed.

Parameters and development stages of regulation are a matter of discussion in doctrinal literature, with a search for balance between statutory, technical and ethical aspects accepted as an optimal condition. It would be also potentially useful to establish a procedure for self-regulation of AI technologies with monitoring as a follow-up [Ibragimov R.S., Suragina E.D., Churilova D.Y., 2021: 91].

The authors of the article presented believe that while public regulation for AI should be introduced in a phased, iterational way, there is no need to artificially fast-track an institution or branch of law.

In addition, one cannot design a system for statutory regulation of "things whose operating principles are not fully understood" [Baturin Yu.M., Polubinskaya S.V., 2022: 152]: regulatory mechanisms will not be strong and purposeful ones unless the potential of AI technologies is made clear.

The approving of relevant regulations should be justified, with provisions of the AI Strategy and the Regulatory Development Concept Note on the avoidance of excessive legal regulation to be adhered to.

The problem of AI regulation continues to be a challenging one from the point of view of methodology, legal technique and practice. Where a legal framework is required to account for security and ethical risks,

regulation should be flexible enough for the adopted approach to keep pace with the current trends in view of rapid progress of AI technologies because a failure to do so may negatively impact upon the technological, information and innovative development of different business segments given the role of AI for national security, technological sovereignty and leadership in the field.

References

1. Antonova L.I., Korneva K.A. (2022) On Soft Law and Self-Regulation in the Codes of Ethics for Artificial Intelligence. *Knowledge*, vol. 4, pp. 34–38 (in Russ.)

2. Arzamasov Yu. G. (2023) Optimal Model of Legal Regulation in the Field of Artificial Intelligence. *Vestnik Voronezhskogo gosydarstvennogo universiteta. Pravo*=Bulletin of Voronezh State University. Law, no. 2, pp. 133–148 (in Russ.)

3. Baturin Yu. M., Polubinskaya S.V. (2022) Artificial Intelligence: Legal Status or Legal Regime? *Gosudarstvo i pravo=State and Law*, no. 10, pp. 141–154. DOI: 10.31857/S102694520022606-7 (in Russ.)

4. Bourcier D. (2001) De L'intelligence Artificielle a la Personne Virtuelle: Émérgence D'une Entité Juridique? *Droit et société*, vol. 3, pp. 847–871.

5. Ibragimov R.S., Suragina E.D., Churilova D.Y. (2021) Ethics and AI Regulation. *Zakon*=Law, vol. 8, pp. 85–95 (in Russ.)

6. Kashkin S. Y. (2021) *The Philosophical, Moral and Ethical Foundations of Soft Law as The Initial Stage of Legal Regulation of Artificial Intelligence.* Moscow: Prospekt, pp. 184–198 (in Russ.)

7. Khabrieva T.Y., Chernogor N.N. (2018) The Law in the Conditions of Digital Reality. *Zhurnal rossiyskogo prava*=Journal of Russian Law, no. 1, pp. 85–102 (in Russ.)

8. Kharitonova Y.S., Savina V.S. (2020) Artificial Intelligence Technology and Law: Challenges of Our Time. *Vestnik Permskogo gosudarstvnnogo universiteta. Juridicheskie nauki*=Perm State University Herald. Juridical Sciences, vol. 49, pp. 524–549. DOI: 10.17072/1995-4190-2020-49-524-549 (in Russ.)

9. Kosykh A.A. (2021) The Artificial Intelligence Law in Law System: Branch of Law or Law Institution? *Vestnik Vladimirskogo gosudarstvennogo juridicheskogo instituta*=Bulletin of Vladimir State Law Institute, vol. 1, pp. 159–164 (in Russ.)

10. Minbaleev A.V. (2022) The Concept of "Artificial Intelligence" in Law. *Vestnik Udmurtskogo gosudarstvennogo universiteta*. *Seria Ekonomika i pravo*=Bulletin of Udmurt State University. Series Economics and Law, vol. 6, pp. 1094–1099 (in Russ.)

11. Mishina N.V. (2020) Artificial Intelligence Law as a New Area of Legal Regulation or a New Branch of Law. *Evraziyskiy souiz ychenykh*=Eurasian Union of Scientists, vol. 5, pp. 62–65 (in Russ.)

12. Polyakova T.A., Kamalova G.G. (2021) «Law of Artificial Intelligence» and its Place in the System of Information Law. *Pravovoe gosudarstvo*: *theoria i praktika*=The Rule of Law State: Theory and Practice, no. 33, pp. 133–145 (in Russ.)

13. Popova A.V. (2021) Soft Law as a Structural Part of the Complex Branch of Russian Legislation in the Field of Legal Regulation of Artificial Intelligence. *Zhurnal Moskovsogo gosudarstvennyogo pedagogisheskogo unuversiteta. Ekonomiseskie issledovania*=Moscow State Pedagogical University Journal of Economic Studies, no. 4, pp. 86–93 (in Russ.)

14. Sucharev A.N. (2021) Prospects of Legal Regulation of the Use of Artificial Intelligence in the Russian Federation. *Vestnik Tverskogo gosudarstvennogo universiteta. Yuridicheskie nauki*=Tver State University Herald. Juridical Sciences, no. 3, pp. 13–21 (in Russ.)

15. Talapina E.V. (2020) Algorithms and Artificial Intelligence in the Human Rights Context. *Zhurnal rossiyskogo prava*=Journal of Russian Law, no. 10, pp. 25–39. DOI: 10.12737/jrl.2020.118 (in Russ.)

Information about the authors:

V. O. Buriaga — Candidate of Sciences (Law), Chief Analytic,

V. V. Djuzhoma — Candidate of Sciences (Law), Chief Expert.

E. A. Artemenko — Researcher.

The article was submitted to editorial office 04.04.2025; approved after reviewing 12.05.2025; accepted for publication 19.05.2025.

Legal Issues in the Digital Age. 2025. Vol. 6, no. 2.
Вопросы права в цифровую эпоху. 2025. Том 6. № 2.

Research article
JEL: K00
UDK: 340
DOI:10.17323/2713-2749.2025.2.69.86

Trust in Artificial Intelligence: Regulatory Challenges and Prospects

Svetlana S. Vashurina

National Research University Higher School of Economics, 20 Myasnitskaya Str., Moscow 101000, Russia,

svashurina@hse.ru, ORCID: 0000-0002-4892-1971.

Abstract

The last few years have witnessed a rapid penetration of artificial intelligence (AI) into different walks of life including medicine, judicial system, public governance and other important activities. Despite multiple benefits of these technologies, their widespread dissemination raises serious concerns as to whether they are trustworthy. The article provides an analysis of the key factors behind public mistrust in AI while discussing ways to build confidence. To understand the reasons of mistrust, the author invokes the historical context, social study findings as well as judicial practices. A special focus is made on the security of AI use, AI visibility to users and on decision-making responsibility. The author also discusses the current regulatory models in this area including the development of universally applicable legal framework, regulatory sandboxes and self-regulation mechanisms for the sector, with multidisciplinary collaboration and adaptation of the effective legal system to become a key factor of this process. Only this approach will producer a balanced development and use of AI systems in the interest of all stakeholders, from their vendors to end users. For a more exhaustive coverage of this subject, the following general methods are proposed: analysis, synthesis and systematization; special legal (comparative legal and historic legal) research methods. In analyzing the available data, the author argues for a comprehensive approach to make AI trustworthy. The following hypothesis is proposed based on the study's findings. Trust in AI is a cornerstone of efficient regulation of AI development and use in various areas. The author is convinced that, with AI made transparent, safe and reliable one, provided with human oversight through adequate regulation, the government will maintain purposeful collaboration between man and technologies thus setting the stage for AI use in critical infrastructures affecting life, health and basic rights and interests of individuals.

Keywords

artificial intelligence; trust in AI systems; system transparency; system visibility; system security; system reliability; regulatory model; regulatory sandbox; self-regulation for AI system development.

For citation: Vashurina S.S. (2025) Trust in Artificial Intelligence: Regulatory Challenges and Prospects. *Legal Issues in the Digital Age,* vol. 6, no. 2, pp. 69–86. DOI:10.17323/ 2713-2749.2025.2.69.86

Background

According to statistics, the Russian public perceives AI mostly in a neutral positive light, a fact confirmed, in particular, by the popular belief that AI would never get out of human control[1]. A survey by Pegasystems showed that only 24% of all those polled in North America, Europe, Near East and Africa, and the Asian Pacific region believed in AI getting out of human control while almost 40% did not agree that AI could handle customer service better than man[2]. Thus, trust in AI cannot be judged as high. However, one has to agree that confidence in AI systems is a key factor of further technological revolution [Leshkevich T.G., 2023: 36]. AI applications can have sizeable impact on people, up to legally binding implications [Vinogradov V.A., 2023: 164]. Obviously, the general criticism of the algorithms based on machine learning comes from their dependance on data quality. Once the source data is biased, the software will generate biased results [O'Neil C., 2016: 87].

Ubiquitous introduction of AI systems raises a critical regulatory issue, that of human trust in AI. In this context, one has to agree with professor Vinogradov that AI systems should be visible and comprehensible to users [Vinogradov V.A., 2023: 157–166]. In this study, author attempts to formulate problem of trust in technologies and its impact on legal regulation of AI. The study primarily purports to discuss what causes mistrust in AI and how to overcome it.

Making AI trustworthy is a prerequisite of regulatory regime that will make AI more intelligible and transparent to users and reduce the risks

[1] Available at: URL: https://ai.gov.ru/knowledgebase/etika-i-bezopasnost-ii/202_ncrii/?ysclid=lvt627n4mj432190293 (accessed: 23.04.2024). Trust in AI: URL: https://wciom.ru/analytical-reviews/analiticheskii-obzor/doverie-k-ii (accessed: 25.04.2025)

[2] Available at: URL: https://www.pega.com/ai-survey (accessed: 23.04.2024)

of violation of human rights. Thus, the challenge is twofold: firstly, to identify what causes public mistrust in AI and, secondly, to discuss regulatory models adopted worldwide and, based on the available regulatory experience, propose ways to offset the causes of mistrust. It is worth noting this study is multidisciplinary with a focus on a comprehensive issue, thus requiring not only to invoke purely legal arguments and assertions but also to apply social study findings and those from related fields of knowledge. In particular, the article refers to examples from history to illustrate socioeconomic implications of high levels of mistrust and human concerns raised by the emergence of new technologies, as well as causes of mistrust and ways to overcome it.

The article provides an analysis of different aspects of social relationships to be regulated amidst complications brought by AI, in particular, those dealing with AI development and introduction, ethical aspects of designing, using and ensuring oversight of AI, human trust in AI, as well as adapting legal regulation of social relationships to the emergence of new technologies.

1. Mistrust in the Emerging Technologies and its Causes

Discussion about human trust in technologies requires a focus on psychological and sociological studies since it is human attitude to innovations that largely foreshadows provisions to regulate a certain area of social relations. While regulation cannot (nor should) anticipate the development of socioeconomic relations, legal provisions, in responding to social conflicts that have taken place, can become an relevant way to address them.

In psychological studies, trust is defined as "emotional attitude, optimistic perception of a thing" [Jones K., 1996: 5], or "psychological attitude consisting of the emotional, cognitive and behavioral components" [Kupreichenko A.B., 2008: 571]. Trust is a critical element of social collaboration expressed in various forms such as trust in government, public agencies, laws. Interestingly, S. Stepkin views trust as relying, among other things, on a balance of individual rights and duties, a reasonably commensurable balance of private and public interests, stability and predictability, openness of government agencies, independence and impartiality of judicial authorities, reliability and consistency of official information [Stepkin S.P., 2023: 32]. It is important to invoke A.N. Kokotov' view whereby the relations built on trust or mistrust define the essence of law and its meaningful functional and formal manifestations

[Kokotov A.N., 2020: 42]. Psychological attitude to a phenomenon will be thus reflected in a legal content.

Indeed, trust in technologies critically depends, in our view, not only on human response to innovations but also on what these technologies are capable of. In discussing this question, it is necessary to identify at least three aspects that clearly illustrate the problem of human trust in technologies:

Changes to the nature of work from AI used in production;

AI safety and reliability;

AI visibility and transparency.

Analysis of the key challenges related to mistrust in technologies will allow to make practical proposals for better regulation of this area.

1.1. Changes to the Nature of Work from AI Used in Production

As a result of the 19th century industrial revolution, machines stepped out as a partial replacement of human functional duties and physical capabilities, with less qualified workers put in charge of automated processes largely to control the equipment. This trend led to gradual ousting of the skilled factory workforce from economic relations associated with production of goods. The introduction of novel and improved capital goods was caused by a desire to make manufacturing better, faster and cheaper. Despite the clearly positive changes for society from automated equipment in different production sectors, these new technologies met with fierce opposition[3]. With a transition from manual to machine work, automation changed the nature of work, only to impact socioeconomic relations.

Mankind is now approaching the fourth industrial revolution caused by AI and big data systems. It is fair to say that current technologies can be a substitute for not just physical but also intellectual human capabilities, being able to process large quantities of data within minimum time, propose graphical or text solutions, create works of art. However, AI use in many areas is not regulated and can potentially become a key issue leading to human rights violations.

[3] In Lancashire automatic equipment ousted manual work in cotton spinning, only to cause violent riots in 1768 and 1779. Available at: URL: https://historyofinformation.com/detail.php?id=443 (accessed: 20.11.2024). In 1866, Belgian workers on strike demolished a glass factory following the introduction of glass melting furnaces. See: G. Deneckere. 1900 België op het breukvlak van twee eeuwen. Tielt, 2006, pp. 70–71.

Bloomberg Intelligence is expecting a 30-fold growth of the generative AI market up to USD 1.3 trillion by 2032[4] as generative AI-enabled solutions will constantly transform industrial operations over the next decades[5]. As of late 2024, generative AI had a major impact on the existing labor market with considerable competitive pressures on different walks of life. Thus, according to a study of the freelance market in Russia, the generative AI — in particular, the rise in popularity of ChatGPT — hit the text processing segment of translators, copyrighters and editors[6]. Meanwhile, the International Labor Organization (ILO) believes that AI will help create more jobs despite that a majority of current occupations will be fully or partially automated[7].

However, the ongoing automation of jobs and partial or full replacement of man in production processes does not always accord well with law, only to cause a negative response by trades. Thus, the United States have become a focal point of strike action, with the Writers Guild of America protesting against the Producers' Alliance for Cinema and Television practices of using AI to write and rerecord any material, and using screen writers' output for machine learning[8]. Meanwhile, the WGA also made proposals to regulate AI use across the industry in the first ever attempt to prohibit using AI as a substitute for workers. The Screen Actors Guild held a no less important strike in the U.S. against video game publishers over a concern that generative AI could be trained to reproduce voice, only to push actors out of work[9].

---

[4] ChatGPT to Fuel $1.3 Trillion AI Market by 2032, New Report Says. Available at: URL: https://www.bloomberg.com/news/articles/2023-06-01/chatgpt-to-fuel-1-3-trillion-ai-market-by-2032-bi-report-says (accessed: 20.11.2024)

[5] Labor market 30 years after: neural networks as the core tool. Available at: URL: https://trends.rbc.ru/trends/education/64ee043f9a79472565f6efde?from=copy (accessed: 20.11.2024)

[6] Labor market impact of artificial intelligence. Availablde at: URL: https://www.tadviser.ru/index.php/%D1%F2%E0%F2%FC%FF:%C2%EB%E8%FF%ED%E8%E5_%E8%F1%EA%F3%F1%F1%F2%E2%E5%ED%ED%EE%E3%EE_%E8%ED%F2%E5%EB%EB%E5%EA%F2%E0_%ED%E0_%F0%FB%ED%EE%EA_%F2%F0%F3%E4%E0 (accessed: 20.11.2024)

[7] Available at: URL: https://rg.ru/2023/08/29/chisto-avtomaticheski.html (accessed: 20.11.2024)

[8] AI can't replace humans yet — but if the WGA writers don't win, it might not matter. Available at: URL: https://www.polygon.com/23742770/ai-writers-strike-chat-gpt-explained (accessed: 20.11.2024)

[9] Video game actors to go on strike over AI // URL: https://www.gamefile.news/p/video-game-actors-strike-sag-aftra, see also: Actors say Hollywood studios want their AI replicas — for free, forever. Available at: URL: https://www.theverge.

However, that it is not only strike action but also trials that dramatically exemplify the rejection of new technologies. In many instances, content providers accused one or more companies of stealing intellectual assets to train large language models[10]. The matter of dispute unambiguously points out that society represented by professional communities is still fearful of losing jobs or incomes. Obviously, those involved in creative occupations, routine work and text processing (translators, copywriters, editors) are all at risk. However, the changing nature of work will generate new jobs required to service AI (like cyber-security specialists, prompt engineers, AI system trainers etc.).

1.2. Security and Reliability of Technologies

A critical aspect of trust in technologies is their security and reliability from a human perspective. The emergence of new technological solutions impacting social relations gives rise to relevant provisions to make technologies trustworthy. With AI systems gradually penetrating all human activities across the board — from leisure to contacts with public authorities — the success and efficiency of their use in areas critical for individual life and rights depend on a high level of security and reliability. In a number of such areas, AI is already around[11].

How should AI safety and reliability be manifested? First of all, AI systems should be resistant to external exposure as a key aspect of cybersecurity. The issue of AI security and reliability is largely related to the stable operation of the system itself, predictability of its behavior and possibility to maintain human oversight. No secure and reliable use of AI in critical infrastructures is possible unless there is an assurance that the system is under control of its owner and/or developer and is able to resist outside attacks and to operate correctly in an uncertain environment. Above all, AI security and reliability criteria come from technical

---

com/2023/7/13/23794224/sag-aftra-actors-strike-ai-image-rights (accessed: 20.11.2024)

[10] Available at: URL: https://www.fastcompany.com/91179905/openai-anthropic-and-meta-tracking-the-lawsuits-filed-against-the-major-ai-companies (accessed: 20.11.2024)

[11] In particular, to analyze medical images in health care; personalize web searches and recommendations; improve road traffic and accessibility of public transport; make public governance more efficient and less costly; provide for maximum comfort in the delivery of public services; ensure face recognition in fighting crime; facilitate and automate routine processes at court, for instance, in predictive administration of justice, etc.

documents and standards regulating the development, introduction and use of this technology but strategic AI regulations should envisage, in our opinion, mandatory drafting and, possibly, harmonization of security and reliability criteria depending on where AI is to be used.

Notably, legal regulation of technologies should meet individual interests, in particular, via the requirements of security and reliability, while, on the other hand, avoid arresting or retarding technological development. The experience of legal regulation of technologies in the 19th century Britain vividly demonstrates provisions meant for safe use of technological achievements can put obstacles to industrial development[12], as evidenced by the automotive sector. This example demonstrates the legislator's strife to enhance other parties' trust in self-propelled vehicles via mandatory traffic hazard warning but the chosen mechanism proved to be inefficient, only to result in provisions that significantly obstructed the sector's development.

1.3. AI Visibility and Transparency to Users

The issue of making AI systems trustworthy is also hinged on AI visibility to users and possibility of authentication and verification of information that AI can generate and disseminate.

As noted above, AI is increasingly harnessed to serve daily needs prompting the widespread use of many technologies. In this regard, it has to be admitted that "the simplicity of using and creating basic products, the emergence of applications for a wide range of users have resulted in the risk of misuse and threats of illicit behavior enabled by technology" [Vinogradov V.A., Kuznetsova D.V., 2024: 218].

Deepfake, a technology harnessed not only to create entertaining content but also to achieve critical business objectives (in cinema, advertising etc.) exemplifies the problem of AI visibility. Meanwhile, this

[12] Under the British Locomotive Act (also known as the Red Flag Act) passed in the second half of the 19th century (1865), the speed of horseless vehicles was limited to 2 miles/hour in urban and 4 miles/hour in rural areas (1 mile/hour=1.61 km/hour). Under the Act, each vehicle was to have three drivers — two in the vehicle and one walking in front with a red flag to warn others of a self-propelled vehicle on the road. Such way of regulating the emerging technologies was clearly contrary to the interests of sectoral development. (see: The Locomotives Act 1865 (Victoriae Reginae 28&29, p. 83 — legislation.gov.uk. Available at: URL: https://www.legislation.gov.uk/ukpga/Vict/28-29/83/pdfs/ukpga_18650083_en.pdf (accessed: 23.04.2024); The Red Flag Act. Available at: URL: https://law-school.open.ac.uk/blog/red-flag-act; Available at: URL: https://Red_Flag_Act_Locomotive_1865_Cars_Speed_Limits_Man_Running_Carrying_A.htm (accessed: 23.04.2024)

technology can be used both for good and evil purposes since it assumes employing AI to manipulate audio, photo and video materials to make them look like original images, videos or sound tracks. In illicitly using a deep fake, the wrongdoer attempts to produce and disseminate AI-generated information that is false and misleading, an equivalent of intentional deception and breach of trust. As a result, this technology is used to commit a crime for personal gain.

However, it is not only deep fake technologies that can lead to a breach of trust and misinformation. With AI capability for self-learning and data generation giving rise to chat bot technologies, a popular AI-enabled chat bot generated false allegation of sexual harassment against a George Washington University professor involving a female student[13]. The chatbot generated on its own a Washington Post article with false information about the crime and would produce upon request a quotation from this article as if it were real. Following this story, Jonathan Turley, US lawyer and legal analyst, called for cautious use of AI stressing the threat of misinformation that this technology can disseminate.

Meanwhile, algorithms are used not only in routine situations but also in human contacts with public authorities, with examples of mistrust also found in the area of justice. Notably, relief in court is inalienable human right to be observed, guaranteed and enforced by the government, so that decision-making algorithms are to be regulated and made visible and comprehensible to trial parties. Because a court decision has an enormous impact on individual rights, especially in criminal proceedings, there should be a mechanism to make sure that algorithmic decision-making is never unfair or inaccurate[14].

AI COMPAS, a system used in the United States for administration of justice, is often subject to criticism. In an important precedent in 2013 involving a certain Mr. Loomis detained in the State of Wisconsin, software (AI COMPAS) was used for risk assessment. The defense argued that this software was used in violation of the right to due process since the accused could not challenge either the evidence for or the accuracy of the text behind the system's decision. Notably, in delivering the sentence, the judge took into account the person's prior criminal history as well as the assessment produced by AI COMPAS.

[13] Available at: URL: https://www.foxnews.com/media/chatgpt-falsely-accuses-jonathan-turley-sexual-harassment-concocts-fake-wapo-story-support-allegation (accessed: 20.08.2024)

[14] Available at: URL: https://towardsdatascience.com/bias-in-the-ai-court-decision-making-spot-it-before-you-fight-it-52acf8903b11 (accessed: 24.04.2024)

AI COMPAS is based on a patented algorithm that takes in account some of the answers to a questionnaire. The algorithm is proprietary and not disclosable to an indefinite range of persons. Under this criminal case, AI COMPAS has identified the accused as being subject to a high risk of relapse, with Loomis convicted to six years in prison. Responding to an appeal, the Supreme Court of the State of Wisconsin has ruled that algorithmic risk assessment used by the first instance court in delivering the sentence did not violate the accused person's right to due process despite this assessment was not disclosed either to the court or the accused[15]. As follows from the above example, the judge has delivered the sentence with reliance on algorithmic decision of a software which was neither transparent nor intelligible to the trial participants. Thus, the guilty sentence relied on a decision generated by the machine has analyzed input data through mathematical calculation.

Thus, whether judicial decisions are fair and correct depends exclusively on the quality of data the developer uploaded to the software. Once introduced not only to the judicial system, but also that of public governance, decision-making algorithms predicting human behavior will eventually result in a technocratic and bureaucratic governance and declining percentage of human decisions [Janssen M., Kuk G., 2016: 371–377], with final decision-making guided by conclusions of an automatic system with minimum human control, only to aggravate the problem of algorithmic responsibility. In this context, building trust in AI will become crucial since a dramatic decline of trust in AI and related algorithmic systems may lead to a still graver crisis of trust in social institutions such as government, businesses and community organizations [Jian J.-Y., Bisantz A.M., Drury C.G., 2000: 53–55].

Algorithmic complexity is a major cause of non-transparency. Widespread use of AI in critically important areas of social life is only feasible if an algorithm as a possible substitute for human decision-making is able to make a decision at least as fair and justified as human person would. In this regard, it is argued that AI systems could be trustworthy if they are legitimate, ethical and reliable[16]. In this context, it is crucial to understand that the issues of legitimacy and ethics cannot be addressed

[15] Loomis v. Wisconsin, 881 N.W.2d 749 (Wis. 2016), cert. denied, 137 S.Ct. 2290 2017. Available at: URL: https://harvardlawreview.org/print/vol-130/state-v-loomis/ (accessed: 24.04.2024)

[16] Ethics guidelines for trustworthy AI // European Commission. Available at: URL: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai (accessed: 24.04.2024)

without involving the regulators that adopt regulations governing AI development and use.

2. Regulatory Models for AI

Regulation has a particular task of making AI visible to users and of creating conditions for more trustworthy AI [Vinogradov V.A. et al., 2023: 157–166]. As for the need to create a legal framework regulating AI development, implementation and use, it is worth noting that it is important for the legislator to identify a balanced approach to regulation of technologies. In social relations complicated by the use of technologies in a digital environment, regulatory challenges come from the fact regulation must not hold back the technological change. Notably, these value-based reference points are contradictory [Barfield W., Pagallo U., 2018: 53]. Thus, one needs to strike a balance between regulation based on constitutional principles and an enabling technological environment. The study of international regulatory experience with regard to AI suggests that neither legal system has so far drafted and adopted a comprehensive instrument addressing all challenges in this area. Meanwhile, several models are worth considering to deal with this task:

Adoption of an overarching regulation;

Adoption of sandbox regulations applicable to AI and other technologies;

Self-regulation of the sector.

Notably, this study, while not considering all regulations approved and came in force in different jurisdictions, is focused only at those that vividly demonstrate regulatory models and purport to enhance trust in technologies.

2.1. Overarching Regulation (Exemplified by the European Union)

In March 2024 the European Union has passed Artificial Intelligence Act (AI Act) to establish an overarching legal framework for AI use. It h`s took force on 1 August 2024 with provisions to be applied gradually over the next 6 to 36 months. The Act's declared purposes were: better functioning of the internal market and promoting the uptake of human-centric and trustworthy AI[17]. It is important AI Act implements a risk-

[17] Art. 1 AI ACT. Available at: URL: https: //eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021 (accessed: 24.04.2024)

oriented approach based primarily on guarantees of human rights and trustworthy AI systems.

The starting point of regulatory approach to AI in the European Union was the White Paper on AI[18] identifying not only risks from the use of AI systems, but also a priority task of making them trustworthy. Thus, the risk-oriented approach enshrined in the Act identifies four risk categories that AI systems could be attributed to. The method of regulation varies considerably depending on the said categories. For instance, AI systems posing a clear threat to security, livelihoods and human rights are to be prohibited.

AI systems classified as high risk will be subject to tougher requirements. Thus, high-risk AI systems include critical infrastructures likely to put at risk the life and health of individuals; educational and vocational trainings determining job access or admission; administration of justice and democratic processes; critical private and public services etc. Notably, the key requirements applicable to high-risk AI systems are visibility and transparency to users[19], human oversight[20] and also high quality of databases for AI learning[21] that would allow to minimize risks for users and generate non-discriminatory outcomes. Moreover, users of limited risk systems subject to only specific transparency requirements[22] will be advised that they deal with an AI system with an option of either continue or reject further use. Obviously, AI system transparency is crucial for the regulator for establishing legal regulation in this area.

In addition, the AI Act provides for a multi-level governance system and support for innovations in the AI sector. On the one hand, this system is expected to ensure efficient oversight over the development, deployment and use of AI across sectors while, on the other hand, to support R&D and law enforcement practices of member states at the national level, with public agencies such as the European Commission on AI, European AI Office, Advisory Forum and Panel of Independent

[18] Available at: URL: https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en (accessed: 24.04.2024)

[19] Art. 13 of AI ACT. Available at: URL: https: //eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021 (accessed: 24.04.2024)

[20] Ibid. Art. 14.

[21] Ibid. Art. 10.

[22] Ibid. Art. 50.

Experts to be set up. Moreover, the Act obliges EU member states to establish AI regulatory sandboxes at the national level[23]. It is worth noting that the national-level regulation in the form of sandboxes offers a considerable potential as it allows to assess the effectiveness of provisions that regulate social relations in this domain.

2.2. Regulatory Sandboxes

While many countries have opted for regulations establishing regulatory sandboxes, it is necessary to identify the benefits of this regime as a whole before discussing how it is used across countries. The institution of regulatory sandboxes is renowned in jurisdictions and advised by international organizations such as the OECD and the International Telecommunication Union [Efremov A.A., 2019: 21-23]. Essentially, a regulatory sandbox allows regulators to establish a special legal regime for innovative businesses in sectors such as IT, finance[24], transportation[25], health[26], public and municipal services.

The institution of regulatory sandboxes allows to drop certain requirements that often hold back the technological development. Using regulatory sandboxes, the companies involved in the development of innovative products and services can test them on practice without running the risk of non-compliance. As regards AI, such sandboxes are used in Germany[27], Russia[28], Canada[29] and other countries.

---

[23] Ibid. Art. 57.

[24] See:CSA Regulatory Sandbox. Available at: URL: https://www.securities-administrators.ca/industry_resources.aspx?id=1588 (accessed: 24.04.2024)

[25] Available at: URL: https://www. timesofisrael.com/new-legislation-paves-path-for-trial-of-driverless-autonomous-taxis-in-israel/; 2020 Autonomous Vehicles Readiness Index. Available at: URL: https://assets.kpmg/content/dam/kpmg/es/pdf/2020/07/2020_KPMG_Autonomous_Vehicles_Readiness_Index.pdf; Self-driving vehicles. Available at: URL: https://www.government.nl/topics/mobility-public-transport-and-road-safety/ self-driving-vehicles (accessed: 24.04.2024)

[26] Health and Biosciences: Targeted Regulatory Review—Regulatory Roadmap. Available at: URL: https://www.canada.ca/en/health-canada/corporate/about-health-canada/legislation-guidelines/acts-regulations/targetedregulatory-reviews/health-biosciences-sector-regulatory-review/roadmap.html (accessed: 15.07.2022)

[27] Making space for innovation vehicles. Available at: URL: https://www.bmwi.de/Redaktion/EN/Publikationen/Digitale-Welt/handbook-regulatory-sandboxes.html (accessed: 24.04.2024)

[28] Federal Law No. 258-FZ “On the Experimental Legal Regimes for Digital Innovations in Russia” of 31.07.2020 (as amended). Available at: URL: http://www.consultant.ru/document/cons_doc_LAW_358738/ (accessed: 24.04.2024)

[29] CSA Regulatory Sandbox...

Once established, regulatory sandboxes allow to have "smart" regulation of information technologies to account for the needs of IT system vendors and users, once the experiment's outcomes are validated. Among the benefits of regulatory sandboxes are lower information asymmetry and regulatory costs, higher capital commitments of companies involved in the experiment, and better understanding of technological innovations by control and supervisory bodies.

For the study of sandbox law provisions for more trustworthy AI, one needs to refer to Federal Law No. 258-FZ "On Experimental Legal Regimes for Digital Innovations in Russia" of 31 July 2020. In particular, as was noted above, the issues of security and regulation of responsibility are crucial here. Thus, FZ No. 258 was amended in 2024 to include a procedure for processing claims on harm to life, health or assets of natural or legal persons from solutions developed by AI under the experimental legal regime. This procedure provides for setting up a special commission to clarify circumstances of the harm being caused. It is worth noting its members may represent not only the regulator but also other stakeholders: sandbox participants, business community, expert community etc. Another equally positive aspect is the principle of open deliberations of such commissions[30] that, in our view, also serves to enhance trust both in AI systems themselves and their vendors.

Meanwhile, there is an issue of assessing the extent of sandbox success. In this regard, one has to accept A.A. Efremov's approach that "a successful experiment may be both the one that yielded positive outcomes and the one whose outcomes cannot be deemed successful for further large-scale application" [Efremov A.A., 2022: 21]. In our view, the main advantage of sandboxes is that the regulator can, via a legal experiment, identify the best approach to effective regulation that strikes a balance between the law-protected values and the imperatives of technological change.

2.3. Sector Self-Regulation

The regulatory models for R&D in artificial intelligence, discussed above, are to be established by public regulators. Meanwhile, they will be less effective where AI vendors are not interested in elaborating shared approaches and principles of AI development, deployment and use at the level of self-regulation. Importantly, it is self-regulation instruments

[30] Para 5, Article 18.1, Federal Law No. 258-FZ // SPS Consultant Plus.

that laid down the early principles and defined business values of major IT companies in this market in what came to be called "codes of good conduct".

The first code of this kind was developed in the United States by Google, a renowned IT giant, in 2018[31]and contained important principles of AI development including data security and privacy.

In Russia and China, self-regulation of the AI sector is also widespread, with large membership associations such as Russia's AI Alliance bringing together IT market leaders (like Sber, Yandex, VK, Uralkhim or Rusagro), and in China — web companies (like Baidu or Tencet), telecom (Huawei) or financial companies (Ping An).

In the People's Republic of China, the focus on self-regulation is made at the level of strategic documents approved by the authorities, with the Next Generation AI Plan stressing the importance of self-regulation at the corporate level, and the White Paper on AI Governance considering AI companies as key entities for future regulation of the sector[32]. Moreover, the interim measures to regulate generative AI services taken in summer of 2023 encouraged collaboration between businesses, universities, research institutions and public agencies in the AI sector, as well as participation of Chinese representatives in the development of international rules for generative AI[33]. For lack of regulation over a long time, several entities (mostly Internet companies) set up in-house AI governance systems and collaborated with other businesses to design a framework for self-regulation and promote the guiding principles for the sector.

Self-regulation sector-by-sector is based on the fundamental principle of bona fide conduct by the parties to legal relationships. Ethical standards for more severe requirements to the development, introduction and use of AI systems are crucial for a balanced regulation of technologies. Undoubtedly, trust between the government and society is not possible without *bona fide* conduct on both sides in the widest sense[34].

In Russia, the parties to the AI Alliance have endorsed in 2021 an AI Code of Ethics as a starting point for self-regulation in developing,

[31] AI at Google: our principles //Available at: URL: https://blog.google/technology/ai/ai-principles/ (accessed: 09.11.2024)

[32] Global Atlas of AI Regulation / Ed. by A.V. Neznamov. Moscow, 2023.

[33] Available at: URL: https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm (accessed: 24.04.2024).

[34] Ethics and Law: Correlation and Mechanisms of Reciprocal Impact / Ed. by V.A. Vinogradov. Moscow, 2023.

introducing and using AI at all stages of its lifecycle not regulated by law and/or technical standards[35]. In guiding the development of technologies in Russia, this document is also expected to build confidence in AI on the part of users, society and government. With 363 companies endorsing the AI Code of Ethics as official signatories[36], these come not only from Russia, but also other countries like Nigeria, Zambia, Cyprus, Senegal, Uganda, Kenya, Uzbekistan, Cuba, etc.

In 2024, the parties to the AI Alliance signed a declaration on responsible development and use of generative AI (Declaration) to establish ethical principles and recommendations for responsible treatment of AI not only for vendors but also users of neural network services[37]. The Declaration builds on advisory provisions of the AI Code of Ethics since "the parties have agreed on the principles of security and transparency, ethical treatment of sensitive issues, measures to prevent abuse and misinformation, as well as on promoting user awareness of the capabilities of new technologies"[38]. To achieve the purposes of the Code, a national Commission for Implementation of the AI Code of Ethics was set up as a body in charge of the implementation of its provisions and related performance monitoring of AI actors; collaboration and exchange of the best practices of AI ethics; drafting proposals on AI development priorities related to ethical aspects. Apparently, such practices can make codes of ethics very efficient as a method of the so-called soft regulation. A controlling authority in place will engage more parties into self-regulation and help develop standardized approaches in this area. Notably, the rules of self-regulation also strive to make AI transparent and intelligible to users which is indicative of a general trend shared both by regulators and businesses themselves.

Conclusions

To sum up the findings of this study, it appears necessary to formulate the following points.

---

[35] Available at: URL: https://ethics.a-ai.ru/assets/ethics_files/2023/05/12/%D0%9A%D0%BE%D0%B4%D0%B5%D0%BA%D1%81_%D1%8D%D1%82%D0%B8%D0%BA%D0%B8_20_10_1.pdf (accessed: 24.04.2024)

[36] The signatories of the AI Code of Ethics. Available at: URL: https://ethics.a-ai.ru/ (accessed: 24.04.2024)

[37] Available at: URL: https://ai.gov.ru/mediacenter/uchastniki-alyansa-v-sfere-ii-podpisali-deklaratsiyu-ob-otvetstvennoy-razrabotke-i-ispolzovanii-gene/ (accessed: 24.04.2024)

[38] Available at: URL: https://tass.ru/ekonomika/20221995 (accessed: 24.04.2024)

Firstly, making AI trustworthy both through regulation and self-regulation (by companies which develop and introduce AI) is a priority task in a number of jurisdictions. In drafting regulatory provisions for transparency and intelligibility of AI actions, it is necessary to strike the right balance between individual rights and liberties associated with AI use and the interests of the sector since excessive administrative procedures behind complicated bureaucratic processes can become a major obstacle to technological development.

Secondly, the problem of trustworthy AI largely depends on its security and reliability as well as on its visibility and transparency to users. As demonstrated by international regulatory experience, the issues of AI visibility and transparency to human users could be addressed, in particular, by mandatory marking AI systems and advising users accordingly. The AI visibility challenge is pending for all legal systems as this criteria will enhance trust in AI systems and AI-enabled decision-making. Trustworthy AI will allow to overcome the digital divide caused not so much by technologically ill-equipped territories as by psychological perception of AI systems by different categories of individuals. Moreover, the experience of the People's Republic of China to step up the liability for AI-enabled misinformation appears useful and promising[39]. As was noted above, massive use of technologies has resulted in illicit ways to harness them. Introducing criminal liability for using AI to deceive or mislead the parties to legal relationships will enhance the society's trust in technologies.

The issue of AI security and reliability is largely related to the system's sustainable, predictable operation and a possibility to maintain human oversight. However, this issue depends, in particular, on security of person and law-protected data behind AI training. The legislator must provide for mechanisms to protect this data. Thus, strike action by creative trades to protest against making copyrighted material or biometric data of celebrities (such as voice) available for AI training is a clear demonstration of the professional communities' rejection of such training practices. It would appear that only the legislator is well-placed to settle the arising controversies.

Thirdly, this discussion of different regulatory models suggests that effective regulation of AI development and introduction in various walks of life requires as a crucial and promising aspect both comprehensive

[39] China seeks to root out fake news and deep fakes with new online content rules //Available at: URL: https://www.reuters.com/article/us-china-technology/china-seeks-to-root-out-fake-news-and-deepfakes-with-new-online-content-rules-idUSKBN1Y30VU/ (accessed: 20.11.2024)

regulation by competent public authorities and self-regulation by the key market players. In this regard, it is worth noting regulatory sandboxes appear to be a shrewd way to proceed since this specific arrangement will facilitate an experiment based on the envisaged purposes, objectives and key indicators of success or failure. Let author of article to believe such sandboxes will allow the regulator to strike a necessary balance for effective regulation of this sector.

Thus, trustworthy AI is currently crucial and trendsetting for further progress in regulating AI development and use. Addressing this challenge will contribute to efficient introduction of these systems into critical spheres of social life including justice, electoral process and other democratic procedures, health, public security, transport accessibility. Apparently, using regulation to build trust in AI is the main vector for legal systems both domestically and internationally wherever one aspires to become a global AI leader.

References

1. Barfield W., Pagallo U. (2018) Research Handbook on the Law of Artificial Intelligence. Northampton: Edward Elgar, 736 p.

2. Efremov A.A. (2019) Experimental Legal Regimes for Digital Innovations: International Experience and Domestic Prospects. *Informatcionnoe pravo*=Information Law, no. 3, pp. 21–23 (in Russ.)

3. Efremov A.A. (2022) Experiments in Public Administration: Aspects of Delivery and Efficiency. *Gosudarstvennaya sluzhba*=Public Service, vol. 24, no. 1, pp. 19–28 (in Russ.)

4. Ethics and Law. Relationship and Mechanisms of Mutual Influence (2023). Monograph. Ed. by V.A. Vinogradov. Moscow: Prospekt, 272 p. (in Russ.)

5. Global Atlas of AI Regulation (2023) Ed. by A.V. Neznamov. Moscow: no publisher, 308 p. (in Russ.)

6. Janssen M., Kuk G. (2016) Challenges and Limits of Big Data Algorithms in Technocratic Governance. *Government Information Quarterly*, vol. 33, pp. 371–377.

7. Jian J.-Y., Bisantz A. M., Drury C. G. (2000) Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics*, vol. 4, no. 1, pp. 53–71.

8. Jones K. (1996) Trust as Affective Attitude. *Ethics*, vol. 107, no. 1, pp. 4–25.

9. Kokotov A.N. (2020) Trust. Mistrust. Law. Moscow: Norma, 192 p. (in Russ.)

10. Kupreychenko A.B. (2008) *Psychology of Trust and Mistrust.* Moscow: Kogito-Center, 739 p. (in Russ.)

11. Leshkevich T.G. (2023) The Paradox of Trust in AI and its Justification. *Filosofiya nauki* i *tekhniki*=Philosophy of Science and Technology, vol. 28, no. 1, pp. 37–47 (in Russ.)

12. O'Neil C. (2016) Weapons of Math Destruction: how Big Data Increases Inequality and Threatens Democracy. New York: Crown Publishers, 209 p.

13. Stepkin S.P. (2023) The Concept of "Trust" in Modern Constitutional Law: Emergence, Prospects of Development and Assessment. *Aktualnye problemy rossiyskogo prava*=Urgent Issues of Russian Law, vol. 18, no. 10, pp. 30–44 (in Russ.)

14. Vinogradov V.A. (2023) Legal Aspects of Development of AI Systems. *Zakon*=Pravo, no. 12, pp. 157–166 (in Russ.)

15. Vinogradov V.A., Kuznetsova D.V. (2024) Deep Fake Technology: International Regulatory Experience. *Pravo*. *Zhurnal Vysshey shkoly ekonomiki*=Law. Journal of the Higher School of Economics, vol. 17, no. 2, pp. 215–240 (in Russ.)

Information about the author:

S.S. Vashurina — Postgraduate Student, Lecturer.

The article was submitted to editorial office 19.03.2025; approved after reviewing 24.04.2025; accepted for publication 14.05.2025.

Legal Issues in the Digital Age. 2025. Vol. 6, no. 2.
Вопросы права в цифровую эпоху. 2025. Том 6. № 2.

IT. Law. Human Rights

Research article
JEL: K1
UDK: 342.7
DOI:10.17323/2713-2749.2025.2.87.117

Informational Privacy in the Age of Artificial Intelligence: A Critical Analysis of India's DPDP Act, 2023

Usha Tandon[1], Neeraj Kumar Gupta[2]

[1] Dr. Rajendra Prasad National Law University, Prayagraj, Uttar Pradesh 211013, India,

vc@rpnlup.ac.in;utandon26@gmail.com, https://www.rpnlup.ac.in/

[2] Institute of Law, Nirma University, Ahmedabad, Gujarat 382481, India,

neeraj_6336700@yahoo.co.in, https://law.nirmauni.ac.in/ India

Abstract

Informational privacy, often referred as data privacy or data protection, is about an individual's right to control how their personal information is collected, used and shared. Recent AI developments around the world have engulfed the world in its charm. Indian population, as well, is living under the cyber-revolution. India is gradually becoming dependent on technology for majority of the services obtained in daily life. Use of internet and Internet of Things leave traces of digital footprints which generate big data. This data can be personal as well as non-personal in nature. Such data about individuals can be utilised for understanding the socio-economic profile, culture, life-style, and personal information, like love life, health, well-being, sexual preferences, sexual orientation and various other types of individual traits. Issues like data breach, however, have also exposed users of information and technology to various types of risks such as cyber-crimes and other fraudulent practices. This article critically analysis recently enacted Digital Personal Data Protection Act, 2023 (DPDP) in the light of following questions: How it tackles with the issues of informational privacy and data processing? What measures have been envisaged under the DPDP Act, for the protection of informational privacy? How individual rights with respect to data protection

are balanced against the legitimate state interest in ensuring safety and security of the nation? Whether this right is available only against the State or against the non-State actors as well? etc. Having critically analysed DPDP Act, the article calls for further refinement of DPDP Act in various areas, more specifically, suggesting that, it is imperative that DPDP Act requires critical decisions based on personal data to undergo human review, ensuring they are not solely the result of automated data processing.

Keywords

privacy; data; technology, governance; DPDP Act; AI.

For citation: Tandon U., Gupta N.K. (2025) Informational Privacy in the Age of Artificial Intelligence: Critical Analysis of India's DPDP Act 2023. *Legal Issues in the Digital Age,* vol. 6, no. 2, pp. 87–117. DOI:10.17323/ 2713-2749.2025.2.87.117

Introduction

India is the largest country in terms of population, and if it is compared with some of the European countries, it may accommodate many such countries. India is also the biggest democracy thriving in the world working towards the material and spiritual well-being of its citizens. Indian population and its rate of consumption has been the hallmark of India's growth in last few decades. Therefore, businesses, and corporations see India as one of the biggest markets. The rate of consumption in every sector has been unprecedented, especially the mobile and internet usage. Today, a large population is using mobile connections as well as Internet services.[1] The volume of Internet data being consumed and number of mobile and internet users reveal there is an increasing trend towards digitalisation.[2]

It is estimated that India's E-commerce industry is worth 125 billion US$ and it is expected to reach 345 billion US$ by financial year 2030. Another estimation provides by the end of 2025 India will have 200 million e-commerce consumers. Further, India's digital banking revolu-

[1] It is estimated that around 102 billing mobile connections were active in the year 2024, 806 million individualsare using the internet. Along with it, there are around 491 million social media users in the country. See Data Reportal, "Digital 2025: India" Feb 25, 2025, available at: https://datareportal.com/reports/digital-2025-india (accessed: 14 April 2025)

[2] Ray Le Maistre, "India now has 1.15 billion mobile connections", *Access Evolution,* Jan 12, 2024, available at: https://www.telecomtv.com/content/access-evolution/india-now-has-1-15-billion-mobile-connections-49371/ (accessed: 19 May 2025)

tion along with UPI is being accessed by 350 million users. A large network of interconnected network of 550 banks is working in the country with the help of 77 mobile applications. Around 2.19 trillion dollars' worth transactions were carried out in India with the help of UPI.[3] It is also to be noted India has world's largest Unique Identification System (UIDAI) where biometric identity in the form of fingerprints and iris scan of 1.38 billion is captured and stored in digital form.[4]

These numbers are sufficient to indicate that India is living in the era of digital revolution. However, the picture narrated above is just the half of the story. Use of digital technology and digital processes have posed various challenges in recent past. India has faced many instances of data breach where data of individuals stood compromised. Some of the major examples of data breach include breach of credit and debit card user's data,[5] LPG consumer's data,[6] *AADHAR* data.[7] Further, data breach in the State Bank of India,[8] and Kudankulam nuclear power plant's data breach,[9] and many more instances highlight that data breach may be a

[3] Ritesh Shukla, "UPI: revolutionising real-time digital payments in India" June 26, 2024, available at: https://www.europeanpaymentscouncil.eu/news-insights/insight/upi-revolutionising-real-time-digital-payments-india#:~:text=How%20many%20users%20and%20payment,in%20a%20seamless%20digital%20manner (accessed: 19 May 2025)

[4] Unique Identification Authority of India. Government of India. About UIDAI, available at: https://uidai.gov.in/en/about-uidai/unique-identification-authority-of-india.html#:~:text=About%20UIDAI&text=The%20UID%20had%20to%20be,to%20the%20residents%20of%20India (accessed: 19 May 2025)

[5] Anshika Kayastha, ICICI Bank blocks 17,000 credit cards after data breach. The Hindu Business Line, April 26, 2024, available at: https://www.thehindubusinessline.com/money-and-banking/icici-bank-blocks-17000-credit-cards-after-data-breach/article68109673.ece, (accessed: 19 May 2025)

[6] Business Standard, "Top LPG supplier leaked millions of Aadhaar data: Security researcher", Feb 19, 2019, available at: https://www.business-standard.com/article/news-ians/indane-leaked-millions-of-aadhaar-numbers-french-security-researcher-119021900172_1.html, (accessed: 19 May 2025)

[7] Nabeel Ahmed, "How the personal data of 815 million Indians got breached | Explained" *The Hindu*, November 07, 2023, available at: https://www.thehindu.com/sci-tech/technology/how-the-personal-data-of-815-million-indians-got-breached-explained/article67505760.ece, (accessed: 19 May 2025)

[8] Udit Verma, "SBI data leak: What happened? What can you do? All you need to know" *Business Today*, available at: https://www.businesstoday.in/technology/story/sbi-data-leak-what-happened-sbi-data-breach-financial-data-168220-2019-02-01, (accessed: 19 May 2025)

[9] Nirmal John, "Breach at Kudankulam nuclear plant may have gone undetected for over six months: Group-IB", *Economic Times*, Nov 25, 2020, available at:

national security concern. Furthermore, usage of mobile and internet services for banking purposes has led to rise in cases of digital financial frauds. Number of such cases have increased massively in last decade.[10] Such shocking instances severely impact the lives of the victims of such frauds.[11] Additionally, recent issues of deepfake images and voice cloning with the help of Artificial Intelligence (hereinafter as AI) have led to various types of frauds and embarrassing situation in many cases.[12]

In background, the article contain critical analysis recently enacted Digital Personal Data Protection Act, 2023 in the light of following questions: How it tackles with the issues of data privacy and data processing in the era of AI? How right to privacy, especially, informational privacy, may be protected in the technological era? Whether such right is available only against the State or against the non-State actors as well? What measures have been envisaged under the DPDP Act, for the protection of personal data? How individual rights with respect to data protection are balanced against the legitimate State interest in ensuring safety and security of the nation? etc.

The article is divided into six parts including Introduction and Conclusions. Dealing with the evolution of the right to informational privacy in India, it analyses the judgment of *Puttaswamy* case[13]. It proceeds to discusses the relevant provisions on informational privacy from the IT Act, 2000. The next pages contain the critical analysis of recently enacted data protection law *viz*. the Digital Personal Data Protection Act, 2023 (DPDP Act) in the context of AI. Last part deals with conclu-

---

https://economictimes.indiatimes.com/news/politics-and-nation/breach-at-kudankulam-nuclear-plant-may-have-gone-undetected-for-over-six-months-group-ib/articleshow/79412969.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst (accessed: 19 May 2025)

[10] India loses 107 crore to cyber fraud in the first three quarters of this fiscal, https://www.cnbctv18.com/business/finance/india-cyber-fraud-digital-payments-losses-rs-107-crore-fy25-19571280.html (accessed: 20 April 2025)

[11] Pavneet Singh Chadha, "A reclusive couple and a double suicide — Karnataka village wakes up to fallout of digital fraud", *The Indian Express,* April 11, 2025, available at: https://indianexpress.com/article/long-reads/a-reclusive-couple-and-a-double-suicide-karnataka-village-wakes-up-to-fallout-of-digital-fraud-9937402/ (accessed: 19 May 2025)

[12] Pankaj Mishra, "AI Scams Surge: Voice Cloning and Deepfake Threats Sweep India", *NDTV AI*, Oct 10, 2024, available at: https://www.ndtv.com/ai/ai-scams-surge-voice-cloning-and-deepfake-threats-sweep-india-6759260 (accessed: 19 May 2025)

[13] *Justice K. S. Puttaswamy (Retired.) And Anr.* v. *Union of India and Ors.* (2017) 10 SCC 1.

sion and suggestions calling for the further enhancement of DPDP Act, with special focus on the suggestion that DPDP Act must incorporate provisions mandating that consequential decisions derived from data analytics be subject to human oversight, rather than relying exclusively on algorithmic outputs.

1. Evolution of the Right to Informational Privacy in India: *Puttaswamy* Judgment

In simple words informational privacy, that's an emerging phenomenon and often referred as data privacy or data protection is about an individual's right to control how their personal information is collected, used and shared. The concept of informational privacy stems from the right to privacy. In India the questions relating to right to privacy have been the matter of concern since its independence. Concerns for privacy were raised in the Constitutional Assembly Debates. It was argued that privacy of correspondence must be included expressly in the Constitution of India.[14] Also, it was proposed that there should be express provision recognizing protection from the unwarranted and intrusive searches and seizure by the State as provided in the American Constitution.[15] However, the final text of the Constitution of India[16] did not contain any express provision with respect to right to privacy.

Issues of unreasonable searches, seizure and State surveillance by the State came to be argued in Supreme Court in 1954[17] and 1964.[18] These cases held that searches and seizure by State are not protected by right to privacy as the same is not expressly recognised under the Indian Constitution. It is interesting to note: the case of *Kharak Singh* regarded the sanctity of home and privacy as a facet of liberty but ironically has failed to recognise right to privacy as a fundamental right.[19] Later on, there

[14] Centre for Law and Policy Research Trust. Constitution of India/ Debates, available at: https://www.constitutionofindia.net/debates/30-apr-1947/ (accessed: 19 May 2025)

[15] The United States Constitution, Fourth Amendment, available at: https://constitutioncenter.org/the-constitution/full-text (accessed: 19 May 2025)

[16] The Constitution of India, 1950. Gazette of India Extra. No. CA/83/ Cons./49. 26th Nov. 1949.

[17] *M P Sharma* v. *Satish Chandra, District Magistrate, Delhi* [(1954) SCR 1077)].

[18] *Kharak Singh* v. *State of Uttar Pradesh*, (1964) 1 SCR 332.

[19] See the observation of the Supreme Court in the *Justice K S Puttaswamy (Retd.), And Anr.* v. *Union of India and Ors.* (2017) 10 SCC 1, p. 352, para 15.

were other judgments by Apex Court declaring there is a right to privacy by highlighting various facets of right to privacy such as wiretapping, narco-analysis, gender based identity, medical information, informational autonomy and other manifestations of privacy.[20]

Finally in 2017, in *Puttaswamy* case[21], nine judges of Constitutional Bench unanimously have decided and settled legal issues revolving around right to privacy, especially the informational privacy. The facts of the case are simple. In 2009, the Indian government introduced one scheme known as *Aadhaar* scheme, that provided a unique 12-digit identification number to every resident of India. It was projected to enable easier access to government services and welfare programs. The *Aadhar* scheme required individuals to provide their biometric data, including fingerprints and iris scan for enrolment. This data was then stored in a centralizes database. The storage and accessibility of a vast amount of biometric data raised concerns about the government's potential for mass surveillance. In 2012, Justice K.S. Puttaswamy, a retired judge, has filed a Public Interest Litigation (PIL) in the Supreme Court of India challenging the constitutionality of the *Aadhaar* scheme arguing that it violates right to privacy due to the mandatory collection of biometric data without adequate safeguards and the potential for surveillance.

The judgment finally has declared that right to privacy is a right on which other rights, as recognised under the Constitution, derive their sustenance. The Court has declared that right to privacy is natural, primordial, basic, inherent and inalienable right. It is the base of liberty and dignity and directly related to it for meaningful exercise of liberties. Mere absence of express provision cannot be the reason to deny such right. Right to privacy is omnipresent and natural right of the individuals as well as group of individuals. The Court has highlighted three important components[22] of right to privacy—spatial control, decisional autonomy and informational control. It was held that the content of right to privacy can be positive as well as negative depending on the facts and circumstances of the case at hand.[23] It was held that right to privacy belongs to physical as well as mental aspects of life. Concerns of cognitive freedoms are dependent on privacy. It was highlighted that dignity and liberty at individual level are inextricably linked and privacy is a subset

[20] Ibid. pp. 400–401, para 102.

[21] Ibid.10 SCC 1.

[22] Ibid. p. 509, para 325.

[23] Ibid. p. 509, para 326.

of liberties. The right to privacy was held to be the inarticulate major premise of the Part III of the Constitution of India and not merely a derivative right.

Among the various facets of privacy discussed in the judgment, informational privacy received prominent attention. The Court, through six concurring judgments, has elaborated on the concept of informational privacy. It said that the interconnectedness of devices and computer sources create large amount of data. These data if seen in silos, may not make sense, but it becomes capable identifying the individuals, if the same is aggregated and then analysed.[24] Further, these data are capable of drawing inferences about personal characteristics and attributes of individuals.

As the Court pointed out, today the usage of Internet has made it difficult to ensure informational privacy. It has observed informational privacy relates to the person's right to determine when, how and to what extent information about him or her is to be communicated to others. It is a right to control personal information. Information which can lead to identification of individual if the same is accessed, used or disclosed. Supreme Court has highlighted: informational privacy requires that if personal information is provided by an individual to a third party, such parting of the information carries with it a reasonable expectation that the same will be utilised only for the specified purposes. The Court however, recognised the exception of legitimate interests of the State.[25] The Court has pointed out that prevention and investigation of crime, protection of revenue and good governance are some of the legitimate State interest for collection of personal information.[26]

The Supreme Court was of the view that the Parliament should enact laws to protect informational privacy. Such law should create a balance between the legitimate use of data by the State as well as non-state actors. The position of the Court was that any such law has to comply with three-fold requirements. Firstly, there has to be express legislation for curtailment of right to privacy, which must be substantive as well as procedurally fair law. Secondly, the law, even if it is for legitimate purpose, must be based on reasonableness as expected under Article 14 of the Constitution, and thirdly, the law has to be proportional. Curtailment

[24] Ibid. p. 500, para 300.

[25] Ibid. p. 501, para 301.

[26] Ibid. p. 505, para 312.

of right to privacy must be only when necessary and only to the extent which is necessary.

The observations of Supreme Court in *Puttaswamy* case about informational privacy in the times of internet and technology provides succinct insights on the need to have a robust legal framework for data protection.

2. Laying the Legislative Foundation: IT Act, 2000

Twenty-five years ago, in 2000, the Indian Government has enacted The Information Technology Act, 2000 (IT Act, 2000),[27] mainly to recognize electronic transactions and facilitates electronic commerce and address cybercrimes. This Act along with IT Rules[28] and Amendments[29] laid down foundational principles for informational privacy. Until the recent enactment of DPDP Act of 2023[30], the IT Act of 2000 was the primary legal framework for data protection and informational privacy. Though new specific law has been enacted in 2023, the provisions of IT Act 2000 still remain relevant for understanding the evolution of informational privacy law in India.

Before the enactment of IT Act the legal status of electronic data was ambiguous. Though the main objective of IT Act, 2000 was to provide legal recognition to electronic records and transactions, it inherently had privacy implications. By bringing digital data under the legal framework, the Act allowed the possibility of regulating the handling of digital data, thus protecting individual information.

One of the most important provisions in the IT Act 2000, relating to informational privacy is 43A[31], was added to the Act through an amend-

[27] Act No. 21 of 2000.

[28] Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011.

[29] Particularly Information Technology (Amendment) Act, 2008.

[30] Act No. 22 of 2023. The DPDP Act hasn't yet come into force as it needs supporting Rules and Regulations, which are currently being developed by the Ministry of Electronics and Information Technology. These rules are crucial for outlining the operational framework and specifics of how the DPDP Act will be implemented and enforced. While the Act itself was passed and notified, the details needed to make it fully operational are still being finalized. The Ministry of Electronics and Information Technology (MeitY) has recently released draft rules are currently open for public feedback. he Act is likely to come into force in a phased manner, with specific provisions being notified by the government as the rules are finalized.

[31] Section 43A of the IT Act 2000 will be repealed once the Digital Personal Data Protection Act (DPDP Act) comes into force. See section 44 (2) (a) DPDP Act.

ment in 2008. It states that if a "body corporate" (any company, firm, sole proprietorship, or association of individuals engaged in commercial or professional activities) possessing, dealing with, or handling "sensitive personal data or information" in a computer resource is negligent in implementing and maintaining "reasonable security practices and procedures," and thereby causes wrongful loss or wrongful gain to any person, such body corporate shall be liable to pay damages by way of compensation to the affected person. The IT Rules 2011, notified under section 43 A, explained "Sensitive Personal Data or Information" (SPDI)[32] includes passwords, financial information (bank account, credit/debit card, other payment instrument details), physical, physiological and mental health conditions, sexual orientation, medical records and history biometric information and any other information received by a body corporate for processing, stored, or processed under a lawful contract or otherwise, which falls under the above categories. It means that the information freely available in public domain or under the Right to Information Act, 2005 cannot be considered as SPDI.

Further, the IT Rules, 2011, also have defined what constitutes "reasonable security practices and procedures" to mean those security practices and procedures that are designed to protect information from unauthorized access, damage, use, modification, disclosure, or impairment. It also specified that compliance with the international standard[33] would be considered compliance with reasonable security practices.

The IT Rules, 2011 provided more specific details regarding data protection obligations. These Rules mandated several key practices for body corporates handling personal information and SPDI. It required body corporates to publish clear and easily accessible Privacy Policy on their websites.[34] This Private Policy must include the type of information col-

[32] The DPDP Act, 2023 on its enforcement, will omit Section 43A and the Sensitive Personal Data or Information Rules (SPDI Rules) under Section 43 A of IT Act 2000. See section 44 (2) (a) DPDP Act.

[33] IS/ISO/IEC 27001 (Information Technology–Security Techniques–Information Security Management System — Requirements).

[34] The Information Technology (Reasonable security practices and procedures and sensitive personal data or information) Rules, Published by Ministry of Communications and Information Technology, G.S.R. 313(E), April 11, 2011, available at: https://www.indiacode.nic.in/handle/123456789/1362/simple-search?query=The%20Information%20Technology%20(Reasonable%20Security%20Practices%20and%20Procedures%20and%20Sensitive%20Personal%20Data%20or%20Information)%20Rules,%202011.&searchradio=rules (accessed: 19 May 2025). See Rule 4.

lected, the purpose of collection, who the information will be disclosed to, and the security practices employed. While laying emphasis on the consent of the Information Provider, the Rules required explicit consent for the collection and disclosure of SPDI.[35] The information provider must be given the option to opt out of providing such information and to withdraw their consent at any time.[36] It states the collection of data should be minimized to the actual necessity for required purpose.[37]

Moreover, the personal information can only be collected and used for the specific purpose for which it was initially collected[38] and should not be retained for longer period than required.[39] Most significantly, the Rules mandates that for disclosure of SPDI to a third party, prior permission from the information provider is required unless it's necessary for compliance with a legal obligation or agreed upon in a contract.[40] The third party receiving the data is also prohibited from further disclosing it.[41] To deal effectively with the grievances, the Rules require that body corporates must appoint a Grievance Officer and the details of the Grievance Officer must be published on their website.[42] This officer is responsible for redressing grievances of information providers within a stipulated timeframe of one month.[43]

Information providers have been given the right to review the information provided and request corrections for inaccuracies, if any.[44] Section 72A of IT Act, 2000 is another significant provision that provides for punishment for disclosure of information in breach of lawful contracts. It stipulates that 'any person, including an intermediary, who, while providing services under a lawful contract, secures access to personal information about another person with the intention of causing wrongful loss or wrongful gain, or discloses such information without the consent of the person concerned or in breach of a lawful contract, can be punished with imprisonment for a term up to three years, a fine

[35] Ibid. Rule 5(7).

[36] Ibid.

[37] Ibid.Rule 5(1)(b).

[38] Ibid.Rule 5(5).

[39] Ibid.Rule 5(4).

[40] Ibid.Rule 6(1).

[41] Ibid.Rule 6(4).

[42] Ibid.Rule 5(9).

[43] Ibid.

[44] Ibid. Rule 5(6).

up to five lakh rupees, or both.' It must be said this provision directly aims at protecting informational privacy by taking unauthorised disclosure of personal information seriously and penalizing unauthorized sharing of data obtained under a contractual obligation.

Some other provisions of the IT Act, 2000, though not directly focused on informational privacy, have indirect implications by criminalizing various cybercrimes. For instance, Section 43 provides penalty for unauthorized access, computer damage, and data theft. This helps protect the integrity and confidentiality of data, which is fundamental to informational m privacy. Section 66 punishes various cybercrimes like hacking, identity theft, and cyber fraud, often involving the unauthorized access or misuse of personal information. Section 69 that gives authority to the government to intercept, monitor, and decrypt information raises privacy concerns, it is intended to address national security issues due to cyber threats. Recognizing the importance of vital data system, Section 70 deals with the protection of critical information infrastructure

Despite these progressive provisions, the IT Act, 2000, had several limitations in safeguarding informational privacy. It primarily focused on cybercrimes and electronic transactions and not on data protection or informational privacy. It was applicable only to 'body corporates' and 'sensitive personal data', leaving other entities and types of personal information less protected. Unlike in many other jurisdictions, it did not provide for independent data protection authority to oversee compliance and enforcement. Though it incorporated provisions on consent and review, it lacked to provide certain upcoming rights to the information provider like the right to erasure (right to be forgotten) or data portability.

Thus, the IT Act, 2000, served as the foundational legal framework proving grounding for addressing the issues of informational privacy in India. Through Section 43A and the IT Rules, 2011, it introduced important concepts such as "sensitive personal data," "reasonable security practices," and the requirement for consent and a privacy policy. Section 72A further strengthened privacy by penalizing unauthorized disclosure. However, rapidly evolving global data protection regime and the constraints of IT Act, led to a more comprehensive and dedicated law, the Digital Personal Data Protection Act, 2023. This new Act aims to address the shortcomings of the IT Act, 2000, by providing a more robust framework for individual data rights, stronger obligations for data fiduciaries, and a dedicated regulatory body. Nevertheless, the IT Act,

2000, played a crucial role in laying the groundwork for recognising and protecting the informational privacy in India.

3. A Comprehensive Legislative Framework: DPDP Act, 2023

The Digital Personal Data Protection Act, 2023[45] was enacted by the Parliament that recognize right to informational privacy, providing a legal mechanism for processing of digital personal data. It provides a comprehensive framework, well explained by lots of Illustrations[46] attached to various provisions of the Act. It provides for responsible data handling, empowers individuals with greater control over their data, and ensures accountability for Data Fiduciaries (hereinafter DFs). The DPDP Act, 2023 is intended to provide for rights of Data Principals (hereinafter DPs) over their personal data.[47] The preamble of the law provides that DPDP Act, 2023 is intended to provide a balancing of interest between the protection of personal data and recognizing of digital data processing for lawful purposes.[48] The competency to enact this legislation by the Parliament can be traced to the Residuary clauses of the Constitution. The Constitution does not contain the word 'data' anywhere in the text or the Seventh Schedule, therefore, the Parliament has exercised its residuary power while enacting this legislation as provided in Article 248 of the Constitution of India.[49]

---

[45] The Digital Personal Data Protection Act, 2023 (hereinafter as the DPDP Act, 2023). It received the assent of the President on 11th August, 2023. The law is yet to be enforced as the commencement date of the same is not yet notified.

[46] For instance DPDP Act, See Sections 5-8.

[47] The DPDP Act comprises of forty-four Sections and a schedule. These forty-four sections are divided in nine chapters. First chapter of the legislation (Ss. 1–3) deals with preliminary matters, such as short title, commencement and the definition of words and phrases as used throughout the legislation. Chapter two of the legislation (Ss. 4–10) deals with obligations of data fiduciaries. Chapter three (Ss. 11–15) deals with rights and duties of data principals. Chapter four titled as 'Special Provisions' contain two sections i.e. section 16 and 17. Chapter five (Ss.18–26) is concerned with matters connected to establishment of Data Protection Board of India. Chapter six (Ss-27 & 28) deals with powers and functions of the Board. Chapter seven (Ss. 29–34) deals with appellate jurisdiction. Chapter eight (Ss. 33 & 34) contains provisions relating to penalties and adjudication. The last chapter of the law (Ss. 35–44) deals with miscellaneous matters and the only schedule attached to the DPDP Act, 2023 contains a list where quantum of penalties has been specified against breach of various provisions under the DPDP Act, 2023.

[48] See the Preamble of the Act.

[49] The Constitution of India. Article 248. (1) Parliament has exclusive power to make any law with respect to any matter not enumerated in the Concurrent List

The following pages provide a detailed account as to how DPDP Act addresses informational privacy.

3.1. The Commencement of DPDP Act

The DPDP Act was passed by the Indian Parliament in 2023 and received the assent of the President of India on August 11, 2023. It was published in the Official Gazette on the same day, thereby becoming law. However, the DPDP Act has not come into force so far at the time of writing this article.

The provision on commencement of the Act provides that the law will come into force as per the notification by the central government and the central government may provide different dates of commencement for specific provisions.[50] This law is yet to be enforced as the commencement date of the same by the central government is not yet notified. For its effective, implementation Act needs supporting Rules which are being developed by the concerned Ministry.[51] The Rules are required to provide clarity on the processes for obtaining consent, rights of Data Principals, grievance redressal mechanisms, technical and organizational safeguards, etc. functions and powers of the Data Protection Board of India (DPBI) etc.

These rules are crucial for outlining the operational framework and specifics of how the DPDP Act will be implemented and enforced. The Ministry of Electronics and Information Technology (MeitY) has recently released Draft Rules,[52] and made them open for public feedback. The government is likely to adopt a phased implementation approach, giving Data Fiduciaries, especially small and medium entities, time to build the necessary compliance infrastructure.

3.2. Applicability and Scope of the Act

Section 3 of the Act provides the scope of applicability of the legislation. It is provided that the Act is applicable in all those cases where pro-

---

or State List. (2) Such power shall include the power of making any law imposing a tax not mentioned in either of those Lists. Read with Entry 97 of the Seventh Schedule.

[50] The DPDP Act. Section 1(2).

[51] Ministry of Electronics and Information Technology (MeitY).

[52] See the Draft the Digital Personal Data Protection Rules, 2025, Ministry of Electronics and Information Technology Notification, G.S.R. 02(E). Jan. 03, 2025.

cessing of digital personal data takes place within the territory of India irrespective of the fact, whether such data was collected in digital form or non-digital form, once the data has been digitized.[53] The Act is applicable in those situations as well, where data is being processed outside the territory of India but the purpose of such processing relates to offering of goods or services in India to DPs located in India.[54]

The same provision also deals with non-applicability of the Act. It is provided that the Act will not apply in two situations. These are (i) when data is processed by an individual for any personal or domestic purpose; and (ii) Such personal data has been made publicly available by the DP herself [55] or the data was made available publicly by any other person who is under a legal obligation to make such data public.[56]

The Act however, does not define the meaning of 'personal' as well as 'domestic' purposes. Concerns have been raised that this may lead to problems.[57] For example, what if a person A sends a courier to person B with the help of a company C. on the one hand, use of data by A may be considered personal but the processing of data by C may not be covered within the exception as provided. Similarly, processing of data by an individual for research will be personal purpose or academic purpose poses a question of concern, what if the research is conducted under the grants received by a funding agency, and research carried out for academic degree purposes? How the distinction is to be drawn? Similar questions may arise about the domestic and non-domestic use.

The Draft Rules provides the application of the Act is exempted when data processing is necessary for research, archiving or statistical

[53] The DPDP Act. Section 3(a) (ii).

[54] Ibid. Section 3 (b).

[55] Ibid. Section 3(c). The clause appears to have used words which are rendered redundant.

[56] Interpretation of Section 3(c) may pose problems. Use of the word 'and' in between clauses (i) and (ii) may become bone of contention. The word 'and' is generally used as a conjunctive word and not disjunctive, which means that when 'and' is used, both the conditions must be fulfilled. Although, the word 'and' in the present case preceded by a semicolon, which is generally understood as ending the clause which denotes that a new and independent clause begins. Therefore, there is a possibility of argument that both the clauses should be read conjunctively. It is submitted that these two clauses do not appear to be related as such and there is no common denominator between these two clauses hence, they should be read disjunctively.

[57] Meghna Bal, "Data Wrapped in Red Tape" *The Indian Express*, April 11, 2025, available at: https://indianexpress.com/article/opinion/columns/europe-data-privacy-9934892/ (accessed: 19 May 2025)

purposes and the standards as provided in the Schedule 2 of the Rules are followed.[58] Further, even the Draft Rules published does not mention the word "domestic" anywhere and leaves it open. It is also interesting to note: when the parent legislation uses the term personal, then the meaning of the same may not be constrained with the help of the subordinate legislation. Hence, the personal use cannot be simply restricted to research, archiving and statistical purpose. It is expected that the Rules will take into consideration this aspect and provide meaning and context of personal and domestic use.

3.3. Rights and Duties of Data Principals

The individual to whom the personal data[59] relates to, is called under the Act as Data Principal (DP) including child as well as any person with disability[60]. The Act recognizes various rights and duties of the DPs. One of the interesting things to be noted in the drafting of DPDP Act is that it uses the expression 'she' or 'her' to refer to all individuals as against the use of 'he', 'his' or 'him'. This is a welcome step to remove the linguistic bias that hitherto has dominated the legal language. The information providers under the DPDP Act have been called as Data Principal, which is departure from the GDPR nomenclature where they are called data 'subject'.[61] This may be a symbolic step but a better jurisprudential approach towards addressing the individuals as principals than the subjects of data concerning them.

---

[58] See the Draft Rules, Ministry of Electronics and Information Technology Notification, G.S.R. 02(E), Rule 15.

[59] The phrase 'personal data' has been defined to mean data about an individual who is identifiable by or in relation to such data. Thus, any data which contains the attribute(s) with the help of which an individual can be identified then such data becomes personal data. See Section 2 (t). The word 'data' has been used to mean information, facts, concepts, opinions or instructions if they are represented in a manner suitable for communication, interpretation, or processing by human beings or by automated means. See Section 2 (h). The word individual is used in the sense of natural person or human being. See Section 2 (s).

[60] The DPDP Act. Section 2 (j).

[61] See generally, Official Journal of the European Union, Regulation (Eu) 2016/679 of The European Parliament and of the Council, of April 27, 2016 "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) (GDPR), available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679 (accessed: 19 May 2025)

The first and the foremost right given to DPs is the right to obtain access to information about personal data available with the DFs. It is provided that when the DPDP Act comes into force, all the DFs are required to provide a summary to DPs about personal data which is already being processed. The DPDP Act also entitles the DPs to know about the identities of all other DFs and data processors who are processing such data. Also, DPs are entitled to information about the description of data which is being processed by such entities. In addition, the government may also prescribe by the Rules that what other information related to personal data of is required to be disclosed by the DFs.[62]

The next right—correction, completion and updating-- is dependent on the first right. If the DPs realize, after obtaining the information of the data available with the DFs, that there is error in data which is being processed by the DFs or on behalf of the DFs then DPs may get the same corrected, completed and updated.[63] This right of the DPs extend up to erasure of such data.[64] Exercise of such right of correction, completion, update and erasure has to be through a request made by DPs in the prescribed manner as provided by the DFs. However, in the legitimate State interest, despite the request for erasure being made, the data may be required to be retained for other specified purposes as may be prescribed under other legal obligation of the DFs under any other law.[65]

Another important right of the DPs relates to right to nominate any other individual in the event of death or incapacity of the DPs who can exercise the rights of the DPs in such eventualities.[66] The right to grievance redressal is also recognized as one of the important rights of the DPs.[67] It is provided that the DPs have right to grievance redressal by readily available means as provided by the DFs or data processors. This imposes a corollary obligation on the DFs and data processors to provide for access to such mechanisms which can provide opportunity of grievance redressal. The grievance redressal has to be within the specified timeline for which the rules is to be prescribed by the Central Government.[68]

---

[62] The DPDP Act. Section 11 (1) (c).

[63] Ibid. Section 12 (1).

[64] Ibid.

[65] Ibid. Section 12 (3).

[66] Ibid. Section 14.

[67] Ibid. Section 13.

[68] Ibid. Section 40(2)(o) .

The DPDP Act also provides for some of the duties that DPs are required to observe while exercising their rights. Though, the exercise of the rights is not dependent on performance of duties, however, it is a laudable provision where the DPs are expected to contribute in the better implementation of the Act. These duties include compliance with the provisions of the Act and all other relevant laws while exercising the rights under the DPDP Act. There is a duty not to impersonate another person while providing the details of another person for specified purposes, duty to ensure that there is no suppression of material information while providing personal data etc., there is a duty to not register false or frivolous grievance under the Act and duty to furnish only verifiable authentic information while exercising right to correction, update or erasure of data under the Act. The DPDP Act also empowers the Board to issue warning or impose cost in case of false or frivolous complaint being made by the DPs.[69]

3.4. Obligations of Data Fiduciaries and Data Processors

As stated above, the preamble of the Act recognizes lawful processing[70] of digital personal data[71] as one of the primary objectives of the legislation. The person[72] who determines the purpose and means of processing of data is called Data Fiduciary (hereinafter as DFs).[73] For the purpose of the Act, the DFs have been divided in two classes—Data Fiduciaries and Significant Data Fiduciaries (hereinafter as SDFs).[74] This SDFs is a special class of data fiduciary within the generic class. The Central Government is required to notify the persons who shall be considered as the SDFs on the basis of various factors such as volume

[69] Ibid. Section 28(12).

[70] Processing in relation to personal data, 'means a wholly or partly automated operation or set of operations performed on digital personal data, and includes operations such as collection, recording, organisation, structuring, storage, adaptation, retrieval, use, alignment or combination, indexing, sharing, disclosure by transmission, dissemination or otherwise making available, restriction, erasure or destruction' .See section 2 (x).

[71] The phrase 'digital personal data' is defined to mean personal data which is in digital form. See Section 2 (n) Even when personal data collected in non-digital form, but later on it was digitized, the Act becomes applicable to such data.

[72] Ibid. Section 2 (s) defines the word person in inclusive manner to include long list of juridical entities whether incorporated or not.

[73] Ibid. Section 2(i) .

[74] Ibid. Section 2(z) .

of data being processed by them, the risks to rights of Data Principals, impact on sovereignty and integrity, security of the State, public order and risk on electoral democracy.[75]

Apart from the DFs, another person who may be processing data is termed as Data Processor, they process data on behalf of DFs.[76] Other relevant concepts such as legitimate use,[77] specified purpose,[78] State[79] have been discussed later at appropriate stages. Obligations of the DFs can be understood as the core or the fulcrum of entire legislation. The first and the foremost obligation of the DFs relates to the compliance with the DPDP Act and other by-laws under the Act as a general-obligations.[80] It is provided that DFs shall process data only in accordance with the provisions of the Act and for lawful purpose only.[81] Lawful purpose refers to any processing which is not expressly forbidden by law. Processing may also occur for certain legitimate purposes as well.[82] The scope of legitimate purposes is defined in the Act to include various things discussed later in this part.[83] The next general obligation of the DFs relates to ensuring that data is complete, accurate and consistent when such data is to be utilized for the purposes of decision making related to DPs or when the same is being disclosed to any other DFs.[84] This obligation should be read along with the corollary right of the DPs to update, correct and complete data being processed by the DFs.

One of the most important obligations of the DFs relates to implementing the appropriate technical and organizational measures to ensure effective observance of provisions and rules prescribed under the Act.[85] The DFs are required to ensure that data in their possession remains protected and all measures reasonably necessary for such protection by them or the data processors should be in place as per the mandate of the law. In the event of breach of such data, there is an obligation

[75] Ibid. Section 10 (1).

[76] Ibid. Section 2(k).

[77] Ibid. Section 2(d).

[78] Ibid. Section 2(za).

[79] Ibid. Section 2(zb).

[80] Ibid. Section 8(1).

[81] Ibid. Section 4(1).

[82] Ibid.Section 4(2).

[83] Ibid. Section 7.

[84] Ibid. Section 8(3).

[85] Ibid. Section 8 (4).

on the DFs to intimate the same to the Board.[86] Also, data cannot be kept with the DFs for indefinite period and the same is required to be erased once the time period as specified in law is met or the consent has been withdrawn by the DPs unless retention of data is mandated by the law.[87] DFs are required to ensure that if the data is with the data processor on behalf of them, then such data is erased by the data processor. The DFs are also required to appoint the DPO (only in case of SDFs) or any other person who will answer the queries relating to data to the DPs.[88] Also, they have to ensure that business contact information of data protection officer (only in case of SDFs) or a person who is able to answer the queries raised by DPs relating to processing of personal data is made available to DPs. Also, the DFs are required to establish effective grievance redressal mechanism for DPs.[89]

It is the duty of DFs to provide notice to DPs for obtaining consent for data processing.[90] Such notice needs to contain the purpose of obtaining the consent in relation to data processing by the DFs. The consent by the DPs must be free, specific, informed, unconditional and unambiguous.[91] The consent should be obtained by a clear and affirmative action which should signify agreement to the processing of personal data for specific purpose and consent will be limited to such specific purpose as necessary for processing. Also, only that much data will be processed by the DFs as is necessary for the specified purposes for which the consent is obtained.[92] The contents of such notice have to be either in English or any other language as specified in the Eighth Schedule of the Constitution of India. Further, contents of the notice must be clear and in plain language.[93] Also, the notice itself should contain contact details of DPO or any person authorized by DFs to respond to communications from DPs for queries, concerns and exercising rights under the Act by DPs.[94] The DPDP Act also envisages similar duty of obtaining consent of the DFs in the transitory period as well. It is provided that when the

[86] Ibid. Section 8 (6).

[87] Ibid. Section 8 (7).

[88] Ibid. Section 8 (10).

[89] Ibid. Section 8 (9).

[90] Ibid. Section 5 (1).

[91] Ibid. Section 6 (1).

[92] Ibid.

[93] Ibid. Sections 5(3) and 6 (3).

[94] Ibid. Section 6(3).

consent of DPs was obtained prior to the enforcement of the Act, then at the time of the commencement of the Act, as soon as reasonably practicable, the DF must obtain consent as described above.[95] The consent in case of personal data of child or a disabled person refers to the consent of parent of such child or lawful guardian of such persons.[96]

Furthermore, it is the duty of DFs to inform DPs about the manner in which they can exercise various rights as recognized under the Act qua DFs such as right to correction, update or removal of data, right to withdrawal of consent, right to grievance redressal of the DPs etc. Also, DFs are required to ensure that process of withdrawal of consent has to be as easy as the process of obtaining the consent by the DFs.[97] In addition, DFs are required to inform DPs about the manner in which they can complain to the Board in case their grievances are not redressed by the DFs.[98]

DFs are required to cease processing of data once DPs have withdrawn their consent from such processing. Once the consent is withdrawn, then processing of data should not occur, except for the legitimate uses as prescribed by law. It is to be noted that the burden of proving that the processing of data is legitimate lies on the DFs. Also, the fact that DPs have not performed their duties as expected by the Act may not absolve the DFs from performing their duties or obligations.[99]

In the case of data concerning children, the law makes it obligatory that data processing must not take place in a manner that is detrimental to well-being of the child. Such processing must not lead to behavioral monitoring or targeted advertising and the processing must be in a manner which is verifiably safe manner.[100] There is additional obligation imposed on SDFs. They are required to mandatorily appoint a Data Protection Officer[101] and Data Auditor.[102] Also, SDFs are required to undertake periodic impact assessment of data protection, periodic audit and other actions as may be prescribed by the Rules in this regard.[103]

---

[95] Ibid. Section 5(2).

[96] Ibid. Section 9(1).

[97] Ibid. Section 6(4).

[98] Ibid. Section 5(1) (iii) read with Section 13(3).

[99] Ibid. Section 8(1).

[100] Ibid. Section 9.

[101] Ibid. Section 8(9).

[102] Ibid. Section 10 (2) (b).

[103] See generally Ibid. Section 10.

3.5. Legitimate Processing of Data and Exemptions

Legitimate use of personal data has been recognized under the Act in addition to processing of data for which consent has been obtained by DFs. Legitimate use may be by the DFs itself or by the State or any of the State instrumentalities.[104] The provision on legitimate use contains various grounds. The first legitimate use which is recognized relates to processing of data that has been shared by the DPs for specific purpose to the DF; and the processing of data by the DF for any other purpose for which she has not indicated that consent is not given to the use of personal data. Where personal data was shared at an earlier occasion by DPs for obtaining some benefit or any grant from the State, then in such situation, processing of data by the State instrumentality for granting any such or other benefits is considered legitimate.[105] Also, when data held by the State was in non-digital form, processing of the same may also occur for digitization purposes. However, in both cases standards and procedure for data processing must be in accordance with the Rules prescribed for the same.[106]

Further, personal data can be processed by the State for the purposes of performing any function as prescribed under any law or in the interest of sovereignty or integrity of India or for security of the State. Similarly, when data processing is necessary for fulfilling any obligation under any law which mandates disclosure of any information to the State, then also the data processing will be covered by legitimate use. Such processing is also required to be in adherence with the Rules in this regard. Additionally, processing of data in compliance of decree, judgment or order of the court, tribunal or any other regulatory institution which relates to contractual or civil nature may also be processed. Similarly, processing of data for employment purposes or for safeguarding the employer from loss or liability is also considered as a legitimate use. Another set of use which relates to safety of life and mitigation or prevention of disaster by various measure of provisioning of relief in such situation is also considered as a legitimate use.[107]

In addition to legitimate use, there are certain situations where specific purpose of processing is exempted from compliance of the man-

[104] See generally Ibid. Section 7.

[105] Ibid. Section 7(b).

[106] Ibid.

[107] See generally Ibid. Section 7, various clauses.

date of the law in relation to the obligations of DFs and rights of DPs.[108] These situations are: processing of data for enforcement of legal right; processing of data by the court, tribunal or an quasi-judicial or regulatory institutions; processing of data in the interest of prevention, detection, investigation or prosecution of any offence or contravention of any law; processing for the purposes of corporate restructuring as approved by the tribunal or any other authority established by law; processing of data of financial defaulters; contractual processing of data, where the DPs are not located in India and data is being processed in India by the contract where any of the parties to the contract is not located in India.

3.6. Powers and Functions of Various Functionaries under the Act

Apart from DPs and DFs, there are other functionaries which have been conferred various obligations, functions and powers under the DPDP Act. These functionaries are Consent Manager, Data Protection Officer, Data Auditor, Data Processors, Data Protection Board of India, Appellate Tribunal and the Central Government. The obligations, powers and functions of these functionaries are discussed below.

Consent Manager. It refers to a person who is registered with the Board for the primary function of acting as a single point contact for DPs on behalf of the DFs.[109] The Consent Manager is required to enable the DPs in managing, reviewing and withdrawing of consent in the accessible, transparent and interoperable manner.[110] Thus, Consent Manager acts like a bridge between the DFs and DPs. They have been made accountable to DPs.[111] The qualifications to register as consent manager and other technical requirements for the same are to be prescribed by the Rules to be notified by the Central Government.[112] If any grievance is made by the DPs, then Consent Manager is required to respond to such grievance within time specified in this regard.

Data Protection Officer (DPO). That Officer to be appointed by SDF[113] is required to represent the SDFs and acts as point of contact for griev-

[108] See generally Ibid. Section 17, various clauses.

[109] Ibid. Section 2 (g).

[110] Ibid. Section 6 (7).

[111] Ibid. Section 6 (8).

[112] Ibid. Section 6(9).

[113] Ibid. Section 10.

ance redressal mechanisms under the Act.[114] The individual based in India alone can be appointed as DPO and it will be responsible to the Board of Directors of the DFs. The functions of DPOs are similar to functions of person appointed by DFs for representing them under the Act as prescribed under Section 6(3).

Data auditor. Data Auditor refers to a person appointed by SDFs. The primary function of data auditor relates to carrying out data audit and other assessments and taking measures for data protection by SDFs.[115]

Data processors. It processes data on behalf of DFs. They are required to act as per the instruction of DFs.[116] The relationship between the DFs and Data Processors are supposed to be contractual and such contract has to be a valid contract.[117] Though, the Act does not expressly mention that the contract has to be written one, but it is expected that the Central Government may prescribe for the same through the Rules in this regard.

Data Protection Board of India. It is the prominent regulatory institution under the Act.[118] The Central Government is required to establish the same by a notification. Board is a body corporate. The Board shall comprise of a chairperson and other members.[119] Number of members are to be specified by the Central Government. The qualifications of chairperson and the members are same. It is provided by the Act that they should be persons of integrity and standing. The relevant experience may be related to the field of data governance, administration or implementation of laws related to social or consumer protection, dispute resolution, ICT, digital economy, law, etc. which in the opinion of the CG, may be useful to the Board. However, there must be at least one member from the discipline of law.[120] Primary functions of the Board relate to ensuring the Act is implemented properly.

The Act envisages that all consent managers will be registered with the Board[121] and such registration shall be based on essential conditions as prescribed by Rules relating to technical and other requirements ap-

[114] Ibid. Section 8(9).

[115] Ibid. Section 10 (2) (b).

[116] Ibid. Section 2(k).

[117] Ibid. Section 8 (2).

[118] Ibid. Section 2(c).

[119] Ibid. Section 19.

[120] Ibid. Section 19(3).

[121] Ibid. Section 6(9).

plicable to Consent Managers. The Board is expected to act as a first reporting authority in cases of data breach.[122] It is an obligation of DFs to inform the Board about breach in the manner prescribed. Once the Board receives the intimation about breach, the Board may give directions for mitigation and other purposes to contain the breach. It may conduct inquiry as well, into the cause of such breach.

Further, the Board is required to conduct inquiry and impose penalties in case of non-adherence of other mandates of law as prescribed by the Act or rules. The Board may receive complaint from the DPs with regard to data breach or non-adherence of the mandate in respect of rights of DPs about grievance not being addressed by DFs or consent manager. The Central and State government may also make a reference to Board, also, any court may also refer the matter to the Board for inquiry in relation to data protection or data processing.

In case of data breach or non-fulfillment of any obligation by DFs, the Board is required to conduct inquiry and it may impose penalty in case it is found that the breach is a significant one. Thus, a discretion has been conferred on Board that it may decide not to impose penalty in all cases. The discretion by the Board will be exercised keeping in mind nature of data breach, or other violation of mandate of law, along with factors such as gravity, duration of breach, type or nature of personal data affected by such breach, whether breach is recurrent one or repetitive, the nature of gain, if any, or loss to the person whose data has been breached, nature of mitigative steps taken by the person at default, the promptitude of the, the proportionality of the monetary fine imposition, and impact of fine if the same is imposed on person at fault.[123] Also, the Board is empowered to issue warnings or impose cost in those cases where it appears that nature of complaint is false or frivolous one.[124]

The Board is required to adhere to the principles of natural justice in proceedings before it and the law mandates that the Board will function as a digital office and physical appearance of the parties is to be avoided.[125] For the purpose of carrying out the functions under the Act, Board has been conferred with powers of a civil court.[126] The Board should

[122] See generally: Ibid. Section 27 deals with functions of the Board.

[123] Ibid. Section 33.

[124] Ibid. Section 28 (12).

[125] Ibid.Section 28.

[126] Ibid.Section 28 (7).

make an attempt to dispose the disputes or other grievances with the help of mediation amongst the parties or it may also decide to dispose the matter if the DFs make voluntary undertaking in matters where the primary grievance relates to non-compliance with the provision of the Act or the rules specifying the time for the compliance by DFs.[127]

Appellate Tribunal. To hear appeals from orders of the board, Appellate Tribunal has been provided for. Specific timelines have been provided under the Act with respect to disposal of the cases in appeal by the Tribunal.[128]

The Central Government. The Central Government is conferred with various powers under the Act in addition to notifying the commencement of the Act. For instance, the establishment of Board and appointment of chairperson and members of the Board are to be done by the Central Government. The primary responsibility of the Central Government relates to enactment of various types of Rules making implementation of the Act effective.[129] In addition, it is also the deciding authority with respect to exemptions of the mandate as provided for processing of data of children by such DFs for specific age bracket, who have adopted verifiably safe measures.[130] And the Central Government is also empowered to notify specific class of DFs who will be considered SDFs for the purposes of the Act.[131]

The Government is empowered to notify the countries where the data transfer will be prohibited[132]. The exemption from operation of law may be provided by the Central Government to any DFs are State instrumentality necessary for protection of the sovereignty or integrity of the nation, security of the state, friendly relations with any foreign state etc.[133]

In addition, the Central Government may also exempt the operation of law for research, archival or statistical purposes.[134] The Central Government is also empowered to provide exemption to startups.[135] Tempo-

[127] Ibid. Section 31.

[128] Ibid. See generally Sections 29-32.

[129] Ibid. Section 40.

[130] Ibid. Section 9 (5).

[131] Ibid. Section 10.

[132] Ibid. Section 14.

[133] Ibid. Section 17.

[134] Ibid. Section 17 (2) (b).

[135] Ibid. Section 17 (3).

rary exemption can also be notified by the Central Government when the notification for the same is notified in the initial five years from the date of the commencement of the Act.[136]

Apart from above powers, the Central Government's power with respect to blocking of the access of data by intermediaries is a powerful tool which is to be utilized cautiously and only when the conditions for the same are satisfied.[137] These conditions of blocking can be considered as triple test. Firstly, the Central Government should receive a reference from the Board intimating that a particular DFs has been imposed with a fine twice and secondly, the board advices that it is in the interest of general public that specific type of data should be blocked on the basis of which DF is able to offer the goods or services in India to DPs. Thirdly, the Central Government is also satisfied that such blocking is necessary for general interest of public. However, such blocking by the Central Government will be only after giving an opportunity of being heard to DFs.

4. Informational Privacy and Artificial Intelligence Algorithms

The DPDP Act is a remarkable piece of legislation protecting informational privacy. It is intended to provide a robust legal framework for processing of digital personal data while attempting to balance the rights of the Data Principals, and the need for data processing for the growth of business and other legitimate purposes. It is gratifying to note the Act, meets the international standards[138] and in certain cases is an improvement over those standards. The Act expressly contains data minimization principle and lawfulness principle with respect to processing of data.[139] Also, the Act provides that data processing is possible only with the consent, and for other reasons such as legitimate State interest, le-

[136] Ibid. Section 17 (5).

[137] Ibid. Section 37.

[138] See generally"India's Digital Personal Data Protection Act vs. the GDPR: A Comparison", available at: https://www.lw.com/admin/upload/SiteAttachments/Indias-Digital-Personal-Data-Protection-Act-2023-vs-the-GDPR-A-Comparison.pdf (accessed: 19 May 2025). This report provides a tabular analysis of each and every provision of the DPDP Act, 2023 with GDPR and points out the parameters where the DPDP Act, 2023 matches with the GDPR. Also, it points out the cases where it has gone beyond GDPR and what provisions are lacking in comparison to GDPR.

[139] DPDP Act. See Section 4(1).

gal obligations and contractual necessity. These principles are generally considered as essential components of law dealing with personal data.

However, the Act makes no express mention of data processing by AI algorithms, though it seems to be within the ambit of the Act, as the definition of 'processing' refers to 'automated' processing as well. Further, concerns such as data bias and biased decision making due to algorithms do not find place in the legislation. However, the Act is not expected to operate in vacuum or isolation. The Indian legal framework specifically provides rights relating to equality, non-discrimination, respecting liberties of individuals which can be curtailed on specific grounds as prescribed by the Constitution of India and that too within the reasonable and proportionate measures of restrictions. Thus, the Constitutional regime mandates that decisions about a person by the State or any of its instrumentality cannot be arbitrary as the same goes against the principles of equality as envisaged under the Constitution of India.[140] Further, any decision which adversely affects an individual must be taken only after giving a reasonable opportunity of being heard and by respecting other principles of natural justice. Thus, if any decision, adversely affecting a person, is being made solely on the basis of AI, the same can be challenged on the ground of arbitrariness which violates the principles of equality and natural justice.[141] However, the mandate of natural justice, reasonableness, non-arbitrary decisions are applicable to State or State instrumentalities only. These principles are not binding, per se, on private persons as the fundamental rights are enforceable against the State only. Thus, to uphold fairness and accountability, DPDP Act should require that data-driven decisions of material consequence involve substantive human evaluation beyond algorithmic inference.

Further, the line between personal and non-personal, anonymized and non-anonymized data is subtle and blurred the era of AI. Thus, the DPDP Act is required to ensure that anonymization of personal data must be robust. The law needs to ensure that data cannot lead to identification of individuals or classes of individual even by a combination of anonymized data when the same is not expected. However, such provisions do not appear in the DPDP Act in its current form. The power of the Central Government in relation to Rule making may be utilized for such Rules which can prescribe such robust framework of anonymization of data.

---

[140] See generally the case of *Maneka Gandhi* v. *Union of India*, 1978 AIR 597.

[141] See R. Pal and P. Samaraditya. *MP Jain Indian Constitutional Law*. Chapter XXI.

Further, threats relating to the use of personal data of individuals, especially in the area of medical, health and life insurance surely pose challenges. Such data may provide real time analysis to insurance companies about health and lifestyle condition of individuals and may be highly determinative factor in deciding to offer of insurance and premium of the same. Therefore, law should provide for regulation of such data being used by companies. Thus, the law should contain provisions that ensure that the adverse decision making on the basis of data is supplemented by human intervention and is not based merely on the processed data. Provision may also be made that minimal data processing through AI should occur for legitimate State interest, contractual necessity. The legal obligation principle should be made a condition precedent for processing of the personal data through AI algorithms.

Conclusion and Suggestions

Since 2017, right to informational privacy is available, in India, against the State as a fundamental right and against non-state actors as a legal and common law right. The biggest challenge with respect to informational privacy arises from usage of internet, mobile technology and Internet of Things which have led to accumulation of large amount of data. Data about an individual or group of individuals can be used for various purposes and the same may prove beneficial as well as harmful to the individual and the society. The legal framework under IT Act 2000, has allowed and promoted the digital growth and various types of businesses have flourished in India in the last two and half decades. However, concerns of digital and cyber frauds etc. have rapidly escalated in the last decade, due to data leakage and data breach.

In 2023, the Indian Parliament enacted the standalone and dedicated law relating to informational privacy known as Digital Private Data Protection Act (DPDP Act). The DPDP Act may be regarded as a legislative framework that aligns with international standards for data protection and, in several aspects, even surpasses those global standards. A more in-depth analysis of the DPDP Act, however, reveals certain areas necessitating consideration for its improvement and enhancement.

Firstly, the DPDP Act, while encompassing automated data processing within its scope, does not explicitly address data processing by AI algorithms. Provisions to tackle some of the crucial issues like algorithmic bias and the resulting discriminatory decisions are absent from the DPDP Act. Though, the Act functions within India's broader consti-

tutional framework, which upholds equality, non-discrimination, and individual liberties and an adverse decision based solely on AI, can be challenged for violating constitutional guarantees, these constitutional safeguards primarily apply only to State actions and not to private entities. Hence, there is a pressing need for incorporating specific provisions in the DPDP Act mandating that consequential decisions derived from data analytics be subject to human oversight, rather than relying exclusively on algorithmic outputs.

Secondly, the line between the personal and non-personal, anonymized and non-anonymized data is becoming thinner and blurred in the era of AI. Hence, DPDP Act is required to ensure that anonymization of personal data must be robust and the same does not lead to identification of the individual or class of individuals even by a combination of anonymized data.

Thirdly, the classification of personal data and sensitive personal data, which has been dropped in the present Act finds relevance in this context. The Rules may prescribe that some sort of very personal data should be kept out of the purview of the processing by AI.

Fourthly, the DPDP Act not only fails to provide for the compensation to the victim of data breach, it also repeals Section 43A of the IT Act, 2000 that prescribed compensation to the victim of data breach. Again, the Rules may contain suitable provisions for the compensation.

Fifthly, DPDP Act does not prescribe maximum time limit for data retention by the State and this requires reconsideration by the legislature.

Lastly, the DPDP Act should certainly provide for educating the masses on informational privacy and the same should be made one of the primary functions of the Data Protection Board. The functions of the Board may also include carrying out and funding research in the area of informational privacy.

Addressing the abovementioned areas, through Rule Making or amendments, will surely strengthen the evolving right to informational privacy in India.

References

1. Al-Khassawneh Y.A. (2023) A Review of Artificial Intelligence in Security and Privacy: Research Advances, Applications, Opportunities, and Challenges. *Indonesian Journal of Science and Technology,* vol. 8, no. 1, pp. 79–96.

2. Artzt M., Tran V.D. (2022) Artificial Intelligence and Data Protection: How to Reconcile both Areas from the European Law Perspective. *Vietnamese Journal of Legal Sciences,* vol. 7, no. 2, pp. 39–58.

3. Bakshi P. M. (2025) *The Constitution of India*. Delhi: Universal Law Publishing, 205 p.

4. Carey P. (2020) Data Protection: a Practical Guide to UK Law. Oxford: Oxford University Press, 689 p.

5. Dass R., Sharma A. et al. (2024) Artificial Intelligence in Media Marketing and Law. Delhi: Bloomsbury, 224 p.

6. Halder D., Jaishankar K. (2012) *Cyber Crime and the Victimization of Women: Laws, Rights and Regulations.* Hershey: IGI Global, 267 p.

7. Jain A. K., Jain S. (2025) Lead Smart in the AI Era. Delhi: Rupa Publications, 280 p.

8. Kamath N. (2012) Law Relating to Computers, Internet and E-Commerce. Gurgaon: LexisNexis, 847 p.

9. Kranenbarg W., Leukfeldt R. (2021) Cybercrime in Context: the Human Factor in Victimization, Offending, and Policing. Cham: Springer, 407 p.

10. Kumar S. (2021) Textbook on Information Technology Laws. Delhi: Whitesmann Publishing Co., 464 p.

12. Lumsden K., Harmer E. (2019) *Online othering. Exploring Digital Violence and Discrimination on the Web.* Cham: Palgrave Macmilan, 407 p.

11. Kuner C. et al. (2018) Expanding the Artificial Intelligence-Data Protection Debate. *International Data Privacy Law,* vol. 8, no. 4, pp. 289–292.

14. Pal R., Samaraditya P. (2025) *MP Jain Indian Constitutional Law*. 6th ed. Delhi: Lexis Nexis, 499 p.

13. Nanda S. K. (2021) Media Law. Prayagraj: Central Law Publications, 497 p.

15. Radu R. (2019) *Negotiating Internet Governance*. Oxford: Oxford University Press, 228 p.

16. Rajput B. (2020) *Cyber Economic Crime in India: an Integrated Model for Prevention and Investigation.* Cham: Springer, 262 p.

17. Ryder R.D., Naren N. (2020) Internet Law. Delhi: Bloomsbury, 539 p.

18. Shah N. (2024) *AI and Social Ethics: Gandhian Approach.* Jaipur: Rawat Publications, 220 p.

19. Sharma V., Sharma S. (2023) *Information Technology Law and Practice: Cyber Laws and Laws Relating to E-Commerce, Privacy, Social Media, Defamation.* Delhi: LexisNexis, 694 p.

20. Viano E.C. (2017) *Cybercrime, Organized Crime, and Societal Responses: International Approaches.* Cham: Springer, 378 p.

21. Westin A.F. (1968) Privacy and Freedom. *Washington and Lee Law Review,* vol. 25, no. 1, p. 166.

23. Yanamala A.K., Srikanth S. (2023) Advances in Data Protection and Artificial Intelligence: Trends and Challenges. *International Journal of Advanced Engineering Technologies and Innovations,* no. 1, pp. 294–319.

23. Završnik A., Simončič K. (eds.) (2023) Artificial Intelligence, Social Harms and human Rights. Cham: Palgrave Macmilan, 276 p.

Information about the authors:

U. Tandon — Senior Professor.

N.K. Gupta — Assistant Professor.

The article was submitted to editorial office 26.05.2025; approved after reviewing 12.06.2025; accepted for publication 12.06.2025.

Legal Issues of Digital Age. 2025. Vol. 6, no. 2.
Вопросы права в цифровую эпоху. 2025. Том 6. № 2.

Research article
JEL: K00
UDK: 340
DOI:10.17323/2713-2749.2025.2.118.133

Smart Digital Facial Recognition Systems in the Context of Individual Rights and Freedoms

Oleg A. Stepanov[1], Denis A. Basangov[2]

[1, 2] Institute of Legislation and Comparative Law under the Government of the Russian Federation; 34 Bolshaya Cherymushkinskaya Str., Moscow, Russia 117218

[1] soa-45@mail.ru. https://orcid.org/0000-0003-1103-580x

[2] d_basang@mail.ru, https://orcid.org/0000-0002-2776-4241

Abstract

The authors discuss the problem of digital facial recognition technologies in the context of implementation of individual rights and freedoms. The analysis is focused on whether their use is legitimate and on interpretation of the provisions behind the underlying procedures. The authors note a significant range of goals to be addressed through the use of smart digital systems already at the goal-setting stage: economy, business, robotics, geological research, biophysics, mathematics, biophysics, avionics, security systems, health, etc. Higher amounts of data and a broader range of technologically complex decision-making objectives require to systematize the traditional methods and to develop new decision-making methodologies and algorithms. Progress of machine learning and neural networks will transform today's digital technologies into self-sustained and self-learning systems intellectually superior to human mind. Video surveillance coupled with smart facial recognition technologies serves above all public security purposes and can considerably impact modern society. The article is devoted to the theme of legitimate use of digital facial recognition technologies and to the interpretation of provisions laying down the underlying procedures. The authors' research interests assume an analysis of legal approaches to uphold human rights as digital facial recognition systems are increasingly introduced into social practices in Russia, European Union, United Kingdom, United States, China. The purpose of article is to shed light on regulatory details around

the use of AI systems for remote biometric identification of persons in the process of statutory regulation. Methods: formal logic, comparison, analysis, synthesis, correlation, generalization. Conclusions**:** the analysis confirms that facial recognition technologies are progressing considerably faster than their legal regulation. Deployment of such technologies make possible ongoing surveillance, a form of collecting information on private life of persons. It is noted that accounting for these factors requires amending the national law in order to define the status and the rules of procedure for such data, as well as the ways to inform natural persons that information associated with them is being processed.

Keywords

smart digital systems; facial recognition; regulation; remote biometric identification of persons; balancing; private and public interests.

For citation: Stepanov O.A., Basangov D.A. (2025) Smart Digital Facial Recognition Systems in the Context of Individual Rights and Freedoms. *Legal Issues in the Digital Age,* vol. 6 , no. 2, pp. 118–133. DOI:10.17323/ 2713-2749.2025.2.118.133

Background

Digital technologies play a key role in transforming modern societies and in reinventing public governance practices. Meanwhile, their introduction into social relations raises serious concerns over security of individuals and the state.

Awareness of the potential to record someone's actions on a storage device is a major factor containing personal behavior [Gordon B., 2021: 1–29].

A study fulfilled in 2004 by B. Welsh from the University of Massachusetts and D. Farrington from the University of Cambridge has showed: where CCTV cameras were installed, street crime declined by 21 percent, with the highest decline observed in parking lots and in locations that, apart from being provided with cameras, were well-lit.

The progress of such technologies relies today on capabilities of AI systems, with society to adapt to the challenges and opportunities enabled by these systems in the process of automatic remote identification of individuals based on unique physical, biological or behavioral features.

Mordor Intelligence, a market research firm, estimated the facial recognition market at USD 6.61 billion in 2024, with prospects to reach

USD 14 billion by 2029 (at the average growth rate of 16.20 percent over the forecast period of 2024–2029)[1].

1. Facial Recognition Technologies

Technological corporations such as Amazon Web Services, Microsoft Azure and Google Cloud are currently validating different tools that use facial recognition to unlock smartphones, as well as services like Google's *Find My Face* that law-enforcement bodies use to counter terrorist threats and mass riots [Grigoriev V.N., 2021: 334–355].

A major outcome of their development and dissemination is that biometric data (such as the face geometry), once in the hands of unauthorized persons, cannot be altered since they are directly associated with a particular person[2].

Moreover, biometric identification methods used simultaneously (in parallel) both online and offline will make the boundary between man and his digital twin very much arbitrary. Thus, the existence and operation of the digital twin (avatar) endowed with a number of capabilities in the virtual space will directly impact the rights, duties, freedoms and legitimate interests of the real human person. A leakage of these biometric data will compromise them, only to considerably restrict not only their possible use but also recovery of the violated rights [Kitchin R., Dodge M., 2021: 112, 114, 125].

For instance, Apple's *Face ID*, a facial recognition technology, was hacked by the Vietnamese company B kav immediately after it went on sale. Source data to produce a human face mask by 3D printing at the cost of approximately USD 150[3] may be easily borrowed from a person's profile on social media. This circumstance aggravates the threat from unauthorized use of private data including as part of the critical infra-

[1] Analysis of the facial recognition market size and shares—growth trends and forecasts (2024–2029). Available at: URL: https://www.mordorintelligence.com/ru/industry-reports/facial-recognition-market (accessed: 26.11.2024)

[2] Daly M.P. et al. Biometrics Litigation: An Evolving Landscape (Practical Law Litigation, April–May 2016). Available at: URL: https://uk.practicallaw.thomsonreuters.com/w-001- 8264?lrTS=20170720182117024&transitionType=Default & context Data=(sc.Default)&first Page=true (accessed: 26.11.2024)

[3] Facial recognition from A to Z for video analytics, video surveillance and access control. Available at: URL: https://securityrussia.com/blog/face-recognition.html (accessed: 26.11.2024)

structure operation since technological capabilities to store information are not confined to national borders, with cloud storage available anywhere across the planet [Huang J., 2020: 1283–1308].

The rapid introduction of facial recognition technologies calls for enhancing the regulatory role of law in this process as their use can result in serious problems of non-selective coverage and inordinate number of individuals subject to arbitrary identification whereas only specific persons need to be identified (for example, at airports and railway stations). No-touch nature of such identification already raises problems associated with a lack of proper legal basis to process personal data, only to result in negative implications for the persons concerned.

Thus, in October 2020, a certain Mr. A. Leushin was held in custody by the guards at Moscow's Auchan supermarket before arrival of the police when a facial recognition system identified him as someone who had stolen78 thousand of rubles worth of fine spirits from this supermarket, with the error not admitted until hours later. In February 2023, the hydrologist A. Tsvetkov was detained when boarding a plane when a neural network decided his face was 50 to 60 percent similar to that of a video fit. In custody for a year on charges of murders dating back to 2002, the scientist had a heart attack and was not released until February 2024 following a vigorous public campaign[4].

If, in view of the above cases, we define the facial recognition technologies as digital algorithms which by comparing two or more facial images can identify or verify them using data in databases for biometric authentication to determine the data's owner[5], the use of such technologies can considerably impact the exercise of individual rights and freedoms.

It is worth noting that video surveillance systems have become widely used in Russia since 2016, with the first 1.5 thousand cameras installed outdoors and in doorways in Moscow for testing[6]. In 2018, this system

[4] 5 cases when facial recognition systems nearly destroyed human lives. Available at: URL: https://skillbox.ru/media/business/5-sluchaev-kogda-sistema-raspoznavaniya-lits-edva-ne-razrushila-zhizn-cheloveka-po-oshibke/ (accessed: 27.11.2024)

[5] Sarabdeen J. Protection of the rights of the individual when using facial recognition technology. Available at: URL: https://pubmed.ncbi.nlm.nih.gov/35309394/ (accessed: 28.10.2024)

[6] Facial recognition system allowed to identify almost 1.500 criminals in Moscow over 3.5 years. Available at: URL: https://www.tadviser.ru/index.php/

was expanded following trial at the World Football Cup where surveillance cameras helped to detain almost 180 persons wanted by the federal police. Moscow's surveillance system was also used in 2020 during the pandemic to identify and penalize more than 200 breaches of lockdown and self-imposed isolation[7]. By that time, the *Safe City* public system was deployed in 40 constituent territories of Russia, with smart digital systems for remote facial identification in use across 13 constituent territories (including cities of Saint-Petersburg, Ryazan and Saratov, also Kamchatka and the Crimea)[8]. In 2021, the Moscow City Office arranged for shopping centers to be connected to the video surveillance system under the administrative procedure, to be followed by Moscow schools [Bobrinskiy N.A., 2020: 91].

2. Regulatory Challenges of Facial Recognition in Russia and Elsewhere

The Russian Federation presently does not have statutory regulation of remote facial recognition systems that would strike a balance between individual interests to safeguard privacy and those of the state related to security and optimization of specific procedures (such as personal identity verification at transport, sport shows, etc.) [Zharova A.K., 2019: 73].

Since it is not determined to what extent digital video surveillance systems with personal identification capabilities are allowed to invade privacy, the result is prosecutorial bias of judicial and other law enforcement practices.

In Federal Law No. 152-FZ "On Personal Data" of 27 July 2006[9], biometric personal data are defined as describing physiological and biological features that identify an individual. Moreover, this Law regulates how these data are processed in absence of the individual's consent, for example, to uphold the national security and defense or combat terrorism (Article 11).

Проект:Как_устроена_система_распознавания_лиц_в_Москве?ysclid=m2uaa7tv2v145186906 (accessed: 26.11.2024)

[7] Smart Moscow City. Video surveillance system in Moscow. Available at: https://www.tadviser.ru/index.php (accessed: 26.11.2024)

[8] Gaynutdinov D., Koroteev K. Facial recognition: a foretaste of dystopia: a report. Available at: URL: https:// runet.report/static/core/doc/Facial_recognition.pdf (accessed: 28.11.2024)

[9] Collected Laws of Russia, 31.07.2006, № 31 (part 1), Art. 3451.

Meanwhile, the procedures to collect and analyze big data including digital footprints is still not properly regulated by law [Pashentsev D.A., Zaloilo M.V., Ivanyuk O.A., 20]; [Maslovskaya T.S., 2019: 59–69]. Provisions of Federal Law No. 149-FZ "On Information, Information Technologies and Data Protection" of 27 July 2006[10] (Federal Information Law) do not capture the progress of relationships in big data as to legal mechanisms behind public enforcement decisions and follow-up control by public institutions. From the perspective of the constitutional right to privacy, automatic collection and processing of biometric data is not regulated by any statutory provision [Kartashov A.S., 2022: 107].

In this regard, it is only logical to raise the question whether facial images from CCTV cameras installed in public places (shopping centers, airports, railway stations etc.) could be considered biometric data.

Under the Civil Code (para 2, part 1, Article 152.1), if someone's image was taken in a public place, provided that this image is not the main thing being used, no consent will be sought. However, this legal stance for the use of biometric data was formulated long before smart digital systems for remote facial identification started to be deployed in Russia. The content of privacy did not capture someone's presence in public places — moreover, being in a public place at a certain time was opposed to private life[11]. This principle objectively allowed to draw a line between the private and the public sphere before the latter was inundated with smart digital systems for facial identification[12].

Since 29 December 2020, the list of grounds for non-consented biometric data processing was legally extended to notarial needs[13] following

[10] Collected Laws of Russia, 31.07.2006, No. 31 (part 1), Art. 3448.

[11] As was explained by the Roskomnadzor (2013), facial images are not deemed biometric data before they are sent to competent authorities for identification. It is enough for the administration of a public place to make textual or graphic announcements to visitors that they might be under surveillance by photo and/or video cameras. According to the agency, once these conditions are met, no consent to surveillance is required // Explanations on treating photo and video footage, fingerprints and other information as biometric personal data and their processing procedure. Available at: URL: https://pd.rkn.gov.ru/press-service/subject1/news2729/ (accessed: 26.11.2024)

[12] Pozdnyakov V. The legitimacy of facial identification by cameras in public places. Available at: URL: http://www.it-lex.ru/faq/zakonnost-raspoznavaniya-lic/ (accessed: 28.10.2024)

[13] See: Federal Law No. 480-FZ "On Amending the Fundamental Law on the Notarial System and Specific Regulations of Russia" of 27.12.2019 // Collected Laws of Russia, 30.12.2019, No. 52 (part I), Art. 7798; Federal Law No. 537-

connectivity of the notarial system to the universal biometric system of the Russian Federation for biometric identification of those referring to notarial services.

In 2017, the Federal Information Law came to include Article 14.1[14], regulating the process of identification through the use of personal biometric data. This was the first step towards the regulatory framework of the universal biometric system[15] for remote identification of individuals by using control templates with appropriate biometric details. Further legislative changes[16] considerably expanded the opportunities for identification to cover any natural persons, not only Russian nationals. Moreover, the Federal Information Law does not constraint possible use of private information systems to process any biometric personal data including collection and storage.

Under the Law on the Universal Biometric System, in force since December 2022[17], primary biometric samples are to be stored in the public system in a coded form, with mathematical data codes (vectors) available to businesses. Also, while the storage of biometric personal data is prohibited by law, UBS vector processing is allowed. It is assumed that personal data of users will be impossible to decode in the event of leakage from the business storage. Since 1 June 2023 the law allows individuals to withdraw from collecting and storing biometric personal data for the purpose of identification and authentication.

Following the passing of Federal Law No. 127-FZ "On Amending Specific Regulation of Russia" of 14 April 2023[18] (Military E-Summons

FZ "On Amending the Federal Law on Non-State Pension Funds of 30.12.2020 Regarding Protection of Rights and Legitimate Interests of Insured Persons in Choosing the Insurer for Mandatory Pension Insurance, and Article 42 of Russia's Fundamental Law on the Notarial System" // Collected Laws of Russia, 04.01.2021, No. 1 (part I), Art. 76.

[14] Federal Law No. 482-FZ "On Amending Specific Regulations of Russia" of 31.12.2017 // Collected Laws of Russia, 01.01.2018, No. 1 (part I), Art. 66 (voided).

[15] While the Ministry of Telecommunications announced the creation of the UBS back in 2016, consistent efforts to draft the regulatory framework for this initiative were made in the following years.

[16] Federal Law No. 479-FZ "On Amending Specific Regulations of Russia" of 29.12.2020 // Collected Laws of Russia, 04.01.2021, No. 1 (part I), Art. 18.

[17] Federal Law No. 572-FZ "On Identifying and/or Authenticating Natural Persons through the Use of Biometric Personal Data, Amending Specific Regulations of Russia and Voiding Specific Provisions" от 29.12.2022 // Collected Laws of Russia, 02.01.2023, No. 1 (part I), Art. 19.

[18] Federal Law No. 127-FZ "On Amending Specific Regulations of Russia" of 14.04.2023 // Collected Laws of Russia, 17.04.2023, No. 16, Art. 2764.

Law), digital facial recognition will be used to identify those evading conscription.

In 2024 the Ministry of Digitization, Ministry of Transport and the RZhD (Russian Railways) have announced an experiment to verify passengers by their biometric data. The use of biometric data for identification when boarding the train will be voluntary, with the service not to be denied to those who refuse[19].

In accordance with Federal Law No. 197-FZ "On Amending the Federal Law on Motor ways and Road Management in Russia and on Amending Specific Regulations" of 29 May 2023, the information on location of stationary and mobile speed cameras and/or transport routes with installed speed cameras should be made public since 1 September 2024 at the official website of the Ministry of Interior[20].

However, no agency in Russia assumes overall responsibility for processes related to facial recognition, and no mechanism allows to check compliance with the procedure for deletion of incorrect biometric data from smart CCTV systems as the key is to define to what extent the biometric identification by facial geometry and other anthropometric data is allowed. Moreover, it is crucial is to account for the difference between footage from video cameras scattered across public places and those integrated into a single smart system for remote facial recognition.

In this connection, it is of interest to discuss A. Popova's appeal against the IT Department of Moscow and Moscow's Head Office of the Ministry of Interior in 2019 regarding the municipal CCTV system. In support of her claims at the trial, the appellant has indicated that the said system violated a number of individual rights guaranteed by the Constitution (Articles 23, 24). In her view, any biometric data processing by the operator should be consented by the affected individual. If this requirement is violated, the constitutional right to privacy is not guaranteed[21].

---

[19] Mass media reported the RZhD planned experiment to identify passengers by their faces. Available at: URL: https://www.forbes.ru/tekhnologii/493537-kommersant-uznal-o-planah-poeksperimentirovat-s-licami-passazirov-poezdov?ysclid=m2vk5w5wjd40876130 (accessed: 26.11.2024)

[20] Starting from 1 September 2024, speed cameras are subject to specific requirements. Available at: URL: https://www.consultant.ru/law/hotdocs/80387.html (accessed: 16.01.2025)

[21] Information on case No. 02A-0577/2019- Available at: https://www.mos-gorsud.ru/rs/savyolovskij/services/cases/kas/details/988f386e-be51-47b0-b48f-e871043ef1fc (accessed: 26.11.2024)

In dismissing the claims[22], the Savyolovsky District Court of Moscow has noted that the use of this technology did not constitute prohibited methods of information processing. Where no personal identification procedure is invoked, video images of an individual cannot amount to biometric personal data. For this reason, public agencies do not need to seek a person's consent for processing biometric personal data.

The court also has emphasized that since the surveillance system directly served the public security purposes, it was not the source of personal data in the sense defined by the personal data law.

This decision of the Savyolovsky District Court later constituted the crucial enforcement instrument behind the legitimacy of video surveillance systems both in Moscow and elsewhere in Russia.

Remote biometric identification by smart digital technologies with restricted access to data under the law of criminal procedure and other regulations does not prevent courts from recognizing it automated personal data processing [Andreeva I.O., 2019: 12]. However, biometric data processing should envisage specific guarantees to avoid misuse of this digital technology, a provision needed to avoid violation of constitutional rights of individuals in absence of uniform enforcement practices [Zorkin V.D.].

According to the lawyer E. Abashina, no provision indicates to what extent two images should be similar for corrective action to apply to an individual, be it additional law enforcement intelligence or court action on administrative offense[23]. This prompts a more profound scrutiny of the question on behavior of someone not involved in a misdeed, in particular, whether it is legitimate to process behavioral data in the continuous mode where the person did not consent to be identified by the smart system[24].

Detractors of facial recognition technologies believe they arbitrarily expand the scope of authority of the police and other special services by offering a tool too attractive and uncontrolled to avoid misuse.

---

[22] The Moscow City Court upheld this decision as the appellate instance.

[23] How the authorities use cameras and facial recognition against protestors. Available at: URL: https://ai-news.ru/2022/01/kak_vlasti_ispolzuut_kamery_i_raspoznavanie_lic_protiv_protestuushih.html (accessed: 23.11.2024)

[24] Dual use cameras: dangers of the facial recognition system. Available at: https://www.rbc.ru/spb_sz/10/10/2019/5d9efecb9a794718418b1e64?ysclid=m2orb29ofm715521384 (accessed: 26.11.2024)

These concerns are caused by the technological possibility to set up cameras to detect only faces of a certain race or ethnicity where it may be reasonable from the statistical point of view.

These concerns are shared by the European Data Protection Board (EDPB), digital sector regulator, whose representatives in 2021 invited the governments to give up expansion of the surveillance camera network and dismantle those already installed. In their view, the current facial recognition practices violated the European rights to privacy and freedom of movement.

Also in 2021, the Advisory Committee of the Council of Europe[25] proposed to prohibit using facial recognition technologies to identify sex, race, color of the skin, ethnicity, social status, health condition, religion and other parameters.

The European AI Act of March 2024[26] allows for restricted use of biometric recognition technologies under very limited scenarios related to crime prosecution and investigation where decisions are to be made promptly such as searching for missing children, preventing terrorist attacks and armed assaults etc. While the Act will not take force before 2026, a number of EU member-states either support the toughest regime of its application or complete prohibition of such technologies in the national territory, especially in the public space.

Thus, the European Union believes law enforcement and judicial uses of AI should not be regarded just as a technological capability but as a policy decision serving the operational purposes of law enforcement agencies and criminal justice systems.

For example, the European Court of Human Rights has handled in 2017 a claim by two professors of mathematics from the University of Montenegro against the installation of surveillance cameras in auditoriums that they believed to restrict the right to privacy[27]. They have

[25] Established under the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data of 1981 // Council of Europe. Available at: URL: https://rm.coe.int/1680078c46 (accessed: 26.11.2024)

[26] Artificial Intelligence Act. Available at: URL: https://www.europarl.europa.eu/RegData/ etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf (accessed: 26.11.2024)

[27] European Court of Human Rights judgment of 28.11.2017 on case of Antovic and Mirkovic v. Montenegro. ECHR 1068. Available at: URL: https://hudoc.echr.coe.int/eng#{%22sort%22:[%22kpdate%20Descending%22],%22itemid%22:[%22001-178904%22]} (accessed: 26.11.2024)

argued that surveillance was unlawful while the university administration did not exercise any necessary control of the relevant procedures. In dismissing the claim, the national courts have noted that since video surveillance was in public places (public space), the university administration did not restrict the right to private life. However, the European Court was critical of the arguments brought by national courts.

The European Court has noted that the notion of "private life" in the meaning of Article 8 of the Convention for the Protection of Human Rights and Fundamental Freedoms included "private social life" related to the possibility to develop one's own social identity and relationships with other people. Operation of surveillance cameras in public places without a legitimate basis (purpose) as an exclusive way to achieve a purpose restricted, in the Court's view, the guaranteed right to privacy and violated the relevant provisions of the national law.

Another notable case handled in 2020 by the Court of Appeal (England and Wales) concerned Ed Bridges from Cardiff who challenged the legitimacy of the police use of facial recognition. The appellant was subject to unauthorized remote biometric identification during Christmas shopping in City of Cardiff (2017) and during a lawful protest (2018). In the appellant's opinion, this technology for biometric data analysis which had been arbitrarily tracking hundreds of thousands people without their consent clearly violated their right to freedom of movement in the absence of strict control by public authorities. During hearing of the case, the solicitors noted that the procedure for biometric data retrieval through facial scanning in violation of the British law was analogous to non-consented taking DNA or fingerprints.

The Court held there was no legal basis for using facial recognition cameras including a watch list, qualifying criteria for locations to install such surveillance systems, or secure storage and use of biometric personal data. In the Court's view, the police was to make sure the algorithms of digital facial recognition technology were free of a gender or racial bias. To be fair, it should be noted that the Court observed a balanced restriction of human rights by this technology as its benefits to the appellant outweighed the likely constraints on privacy[28].

[28] The UK recognized the facial recognition technology as unlawful. The system was used by the South Wales police. Available at: URL: https:// metronews-ru.turbopages.org/metronews.ru/s/novosti/world/ reviews/v-velikobritanii-priznali-nezakonnym-ispolzovanie-tehnologii-raspoznavaniya-lic-1700348/ (accessed: 23.10.2024)

Meanwhile, there are over 420 thousand cameras in London alone, of which some are capable of identifying suspicious items and recognize faces of individuals wanted by the police[29].

Thus, the UK regulation allows law enforcement agencies to use smart video surveillance systems installed in public places for remote facial identification while the law and enforcement practices provide an exhaustive list of terms and grounds for legitimate and admissible use of such surveillance[30]. The use of hi-tech systems by public authorities accounts for the position of civil society institutions including private interests of the population.

The San Francisco city council prohibited the facial recognition technology since 14 May 2019 as the public believed that it posed a threat to the fundamental right of local inhabitants to privacy and other inalienable civil freedoms. No CCTV system can be used by the municipal authorities and the police. The system is not for use by the police as the underlying facial recognition algorithms are obviously unreliable and non-transparent. "System errors can result in innocent black people being involved in police investigations where their lives may be at risk", said Matt Cagle, lawyer of the American Civil Liberties Union of North California[31].

As a compromise between systems' deployment and their full prohibition, a moratorium could be introduced during the period of perfecting the technology because it can be of considerable benefit to society in criminal investigations such as searching for missing persons, victims of human trafficking, potential terrorists. Meanwhile, facial identification technologies are widely and unrestrictedly used by private firms, and by the administration of the San Francisco international airport and seaport as facilities subject to the federal jurisdiction[32].

---

[29] Sharafiev I. London boasts un unprecedented number of CCTV cameras. Available at: URL: https://hightech.fm/2019/08/01/cctv. (accessed: 26.11.2024)

[30] Surveillance Camera Code of Practice. Available at: URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1010815/Surveillance_Camera_Code_of_Practice__update_pdf (accessed: 26.11.2024)

[31] San Francisco to become the first American city to ban the facial recognition technology. Available at: URL: https://forbes-ru.turbopages.org/forbes.ru/s/tehnologii/376099-vlasti-san-francisko-zapretili-ispolzovanie-tehnologiy-raspoznavaniya-lic (accessed: 26.11.2024)

[32] Ibid.

A number of large U.S. metropolitan centres have imposed a similar ban on this technology for fear of its unauthorized use, with three states — California, New Hampshire and Oregon passing laws to prohibit the use of facial recognition in body cameras of police officers. In 2020, following the *Black Lives Matter* riots in the United States, IBM, Amazon and Microsoft restricted or suspended sales of facial recognition products.

Under the law of the State of Illinois[33], processing of biometric data should be consented by the individual concerned for the sale, exchange or other profiting from data to be legitimate. The requirements to processing biometric personal data are aimed at ensuring privacy, impartiality and non-discrimination [Kharitonova Yu.S., 2021: 490].

It is noteworthy that 8 out of the top 10 most "watched" cities in the world are in China[34], with CCTV cameras ensuring security of the territory to identify in some cases a misdeed or a person behind it. Once a face is recognized as belonging to the individual on the watch list, the system will signal an outstanding fine, traffic offense, overdue debt or alimony.

The Eyecool smart CCTV system deployed in the majority of airports and railway stations will daily report to the Sky Net mass surveillance system over two million images of suspects.

China's Sky Net national project is a technologically controllable system for comprehensive surveillance of the population using more than 800 million cameras with the facial recognition capability, one per each citizen[35]. Deployed since 2005, the system is not confined to public security: the project is crucial for the anti-corruption system, as well as the Social Credit System that brings together the information from each citizen's trustworthiness digital profile[36].

---

[33] Biometric Information Privacy Act. 740 ILCS 14 // Illinois General Asse mbly. Available at: URL: https://www.ilga.gov/legislation/ilcs/ilcs3.asp?ActID=3004&ChapterID=57 (accessed: 26.11.2024)

[34] For comparison, the national surveillance system comprises approximately 50 million cameras in the United States, 5–6 million in the United Kingdom and about 300 thousand in Russia.

[35] How information security is implemented in China. Available at: URL: https://nvo.ng.ru/nvo/2023-01-26/13_1222_security.html?ysclid=m2uiwgandp510405287 (accessed: 26.11.2024)

[36] CCTV with facial recognition to be deployed in Moscow's metro befor e 1 September. Available at: URL: https://www.m24.ru/news/mehr-Moskvy/23012020/104711?utm_source=CopyBuf (accessed: 26.11.2024)

Unlike people in Europe, Chinese nationals perceive the wide deployment of CCTV systems in the national territory quite favorably, with 67 percent approving and nearly 9 percent disapproving the installation of such smart digital systems in China [Kostka G., Steinacker L., Meckel M., 2021: 671–690].

As a result, almost all population of China (over 1.4 billion of human beings) is covered by the facial recognition database.

Digital facial recognition technologies based on access to databases of social media and mobile network operators help the police to identify and penalize traffic violators while also allowing to reduce traffic load, reinforce security and improve the system's performance.

In 2017, the State Council of China developed the New Generation Artificial Intelligence Development Plan[37] that envisages the collection of data and evidence for criminal investigations, and analysis of legal instruments for a smart judicial system.

Under Article 26 of the Personal Information Protection Law of the People's Republic of China (PIPL)[38], that is in force since 1 November 2021, the equipment for image recording or facial recognition will be installed in public places as may be necessary for national and public security in accordance with qualifying criteria to be established. Personal images and identification features may be collected only to serve national security and no other purpose, unless consented by data subjects to serve other needs.

Thus, the whole of biometric data collected through the use of digital video surveillance systems is governed by legal provisions that regulate the requirements to personal data security whereby personal data may be collected only if consented by the data subject exclusively for "legitimate, necessary and specific purposes" [39].

Since 1 August 2021 the Supreme Court of Peoples' Chinese Republic has prohibited private firms from using the outcomes of biomet-

[37] China has missed out on the industrial revolution but will not miss out on the digital one. Available at: URL: https://russiancouncil.ru/analytics-and-comments/analytics/kitay-upustil-promyshlennuyu-revolyutsiyu-ne-propustit-tsifrovuyu/?ysclid=m2uj0etc3c884897317 (accessed: 26.11.2024)

[38] Available at: https://digichina.stanford.edu/news/translation-personal-information-protection-law-peoples-republic-china-effective-nov-1-2021 (accessed: 12.11.2024)

[39] Personal Information Security Specifications, in force since 1 May 2018.

ric video identification, unless consented by the individuals concerned, with the principles of legitimacy, fairness, objectivity and security, protection of state and business secrets, privacy and personal information to be strictly observed for any use of facial recognition technologies[40].

Conclusion

Thus, continuous operation of smart facial recognition systems in the public space serves to record and collect data related, in particular, to private life. This circumstance requires to amend the national law accordingly to define how these data will be processed and to regulate how natural persons will be advised in this respect.

Meanwhile, regulation of social relations associated with the use of smart facial recognition systems should be aimed at striking a balance between private and public interests in retrieving, processing and updating biometric personal data through the use of such digital systems. This calls for a compromise between the observance of human rights and public security requirements based on possibilities to safeguard privacy and on technological conditions behind the use of smart facial recognition systems.

References

1. Andreeva I.O. (2019) Face Recognition Technologies in Criminal Proceedings: Issue of Legal Basis behind the use of Artificial Intelligence. *Vestnik Tomskogo gosudarstvennogo universiteta*=Bulletin of the Tomsk State University, no. 11, p. 12 (in Russ.)

2. Artemova S.T., Zhiltsov N.A. et al. (2020) Digital Divide and Constitutional Guarantees of Digital Equality. *Konstitutsionnoe i munitsipalnoye pravo*=Constitutional and Municipal Law, no. 10, pp. 41–45 (in Russ.)

3. Bobrinskiy N.A. (2021) Moscow's Punitive Innovation: Tentative Results. *Zakon*=Law, no. 6, pp. 89–95 (in Russ.)

4. Gordon B. (2021) Automated Facial Recognition in Law Enforcement: The Queen (On Application of Edward Bridges) v. The Chief Constable of South Wales Police. *Potchefstroom Electronic Law Journal*, no. 24, pp. 1–29.

5. Grigoriev V.N. (2021) Information Technologies in Riot Investigation. *Vestnik Sankt-Peterburgskogo gosudarstvennogo universiteta*. *Pravo*=Bulletin of Sankt Petersburg University. Law, no. 2, pp. 334–355 (in Russ.)

[40] China' Supreme Court has prohibited private firms from using facial recognition without consent. Available at: URL: https://3dnews.ru/1045567/verhovniy-sud-kitaya-zapretil-chastnim-kompaniyam-ispolzovat-raspoznavanie-lits-bez-soglasiya-lyudey/ (accessed: 26.11.2024)

6. Huang J. (2020) Applicable Law to Transnational Personal Data: Trends and Dynamics. *German Law Journal*, vol. 21, no. 6. pp. 1283–1308. https: URL: // doi:10.1017/glj.2020.73.

7. Kartashov A.S. (2022) Realization of Constitutional Rights and Freedoms in 'Smart Cities': Main Risks and Ways to Minimize it. In: Constitutional Reform and Legal Development of Ethnic Groups in Russia. Kazan: Buk, pp. 104–111 (in Russ.)

8. Khabrieva T. Ya. (2018) The Law in the Digitalization Conditions. *Zhurnal rossiyskogo prava*=Journal of the Russian Law, no. 1, pp. 85–102 (in Russ.)

9. Kharitonova Y.S., Savina V.S., Pagnini F. (2021) AI Algorithmic Bias: Ethics and Law. *Vestnik Permskogo gosudarstvennogo Universiteta. Jurisprudencia*=Bulletin of the Perm State University. Jurisprudence, no. 3, pp. 488–515 (in Russ.)

10. Kitchin R., Dodge M. (2021) (Un) security of Smart Cities: Risks, Mitigation and Prevention of Negative Consequences. In: City Networks: People. Technologies. Authorities. E. Lapina-Kratasyuk et al. (eds.). Moscow: Novoe Literaturnoe obozrenie, pp. 105–130 (in Russ.)

11. Kostka G., Steinacker L., Meckel M. (2021) Between Security and Convenience: Facial Recognition Technology in the Eyes of Citizens in China, Germany, the United Kingdom, and the United States. *Public Understanding* of *Science*. no. 30, pp. 671–690.

12. Maslovskaya T.S. (2019) Digital Sphere and Constitutional Law: Facets of Interaction. *Konstitutsionnoye i munitsipalnoye pravo*=Constitutional and Municipal Law, no. 9, pp. 18–22 (in Russ.)

13. Pashentsev D.A. et al. (2019) Digitizing Law-Making: Search for New Solutions. Moscow: Infotropic, p. 20 (in Russ.)

14. Rassolov I.M., Chubukova S.G. et al. (2020) Biometrics in the Context of Personal Data and Genetic Information: a Legal Dimension. *Russkiy zakon*=Lex Russica, no. 1, pp. 108–118 (in Russ.)

15. Talapina E.V. (2021) Surveillance (Spying) and Human Rights: New Risks in the Digital Age. *Sravnitelnoye konstitutsionnoe obozrenie*=Comparative Constitutional Review, no. 6, pp. 123–136 (in Russ.)

16. Zharova A.K. (2019) Regulating Information Security in the 'Smart Cities'. *Yurist*=Lawyer, no. 12, pp. 69–76 (in Russ.)

17. Zorkin V.D. (2021) Under the Sign of the Fundamental Law. Constitutional Court at the Turn of the Fourth Decade. *Rossiyskaya Gazeta*=Gazette of Russia. 27 October, no. 247 (in Russ.)

Information about the authors:

O.A. Stepanov — Doctor of Sciences (Law), Professor.

D.A. Basangov — Candidate of Sciences (Law), Senior Researcher.

The article was submitted to editorial office 30.05.2025; approved after reviewing 06.06.2025; accepted for publication 06.06.2025.

Legal Issues in the Digital Age. 2025. Vol. 6, no. 2.
Вопросы права в цифровую эпоху. 2025. Том 6. № 2.

Research article
JEL: K 00
UDK: 34.096
DOI:10.17323/2713-2749.2025.2.134.160

Brain-Computer Interface 5.0: Potential Threats, Computational Law and Protection of Digital Rights

Said S. Gulyamov

Tashkent State University of Law, 7 Iftihor, 8 Yunusabad,Tashkent, Uzbekistan 100057,

said.gulyamov1976@g.mail.com, ORCID: https://orcid.org/0000-0002-2299-2122,

Abstract

The development of neurotechnologies is now at a critical point where direct readout and modulation of brain activity has passed from test studies to business applications, only to urgently require adequate legal and technological guarantees. The relevance of this study is prompted by the rapid development of the fifth generation brain-computer interface (BCI 5.0), a technology that provides unprecedented potential of direct access to neural processes while at the same time creating principally new threats to digital rights of individuals. The existing legal mechanisms have turned out to be inadequate for regulating altogether new risks of manipulating consciousness, unauthorized access to neural data and compromised cognitive autonomy. The study is focused on legal and technological mechanisms for protection of digital rights in the context of introducing the fifth generation neural interface technologies including analysis of regulatory gaps, technical vulnerabilities and possible security guarantees. Methodologically, the study is based on the multidisciplinary approach bringing together neuroscience, law and information technology, and on the comparative analysis of regulatory framework and inductive inference of specific regulatory mechanisms. The main hypothesis is: legacy regulatory mechanisms for data protection in biometric and telecommunication technologies are structurally inadequate for BCI 5.0 while digital rights could be protected only by a hybrid system combining special provisions with technological guarantees via mechanisms of computational law. The author puts forward a minimum set of viable security and confidentiality standards, comprehensive cryptography and blockchain-based ap-

plications, as well as detailed legislative advice for ethical and safe neurotechnological development with secure guarantees of fundamental human rights in the digital age. Findings of the study are of considerable practical value for legislators, those involved in the development of neurotechnologies, regulatory bodies and advocacy organizations by proposing specific evidence-based tools and mechanisms to strike an effective balance between the innovative development and the imperatives of protecting human dignity, mental autonomy and cognitive freedom.

Keywords

digital rights; computational law; neural privacy; cognitive freedom; neuron technologies; legal regulation; data protection; ethical governance.

For citation: Gulyamov S.S. (2025) Brain-Computer Interface: Potential Threats, Computational Law and Protection of Digital Rights. *Legal Issues in the Digital Age,* vol. 6, no. 2, pp. 134–160. DOI:10.17323/ 2713-2749.2025.2.134.160

Background

The emergence of brain-computer interface technologies (BCI) opens up an enormous potential not only for improved communication between individuals and computers but also for new opportunities in the event of disability.

However, these rapidly advancing technologies are fraught with altogether new regulatory challenges for digital rights of individuals. This article provides an overview of BCI 5.0 innovations, identifies the main threats to rights, discusses the current regulatory principles worldwide, shows the implications of inefficient legal guarantees and proposes viable technical and policy standards for confidential, safe and responsible introduction of BCI 5.0.

BCI technologies will directly link the brain with external devices bypassing traditional neuromuscular outputs. While BCI 1.0 and 2.0 were only for auxiliary applications for locomotor and communication disabilities, BCI 3.0 offers a basic device control potential by analyzing EEG, and BCI 4.0 is capable of hands-free texting, web browsing and gaming at up to 60 characters per minute. BCI 5.0 will elevate these capabilities to a new height through a high-density wireless EEG for seamless conversation, unrestricted environmental management and access to rich virtual worlds.

For example, Facebook's sensory headband prototype allows people to type by simply thinking while Kernel brain prosthetic aims to repro-

duce hippocampus memory function, and Neuralink strives to help paralyzed persons to control digital devices using a wireless BCI implant. This is nothing short of a neurotechnological revolution since such an invasive, ubiquitous EEG access will profoundly threaten privacy, security, identity and behavior. Notably, consumer EEG headsets are quite vulnerable to spoofing, signal injection and neural data theft.

It is equally worth noting BCI may be manipulated, only to malignantly alter the user's perception, behavior and memories [Burwell C., 2017: 1–12]. Those patenting such capabilities including Elon Musk's Neuralink are not subject to any mechanism for accountability, compensation of damage or civil supervision of likely harm [Sample M., 2021: 159]. A lack of proper legal protection from these emerging risks creates an instant policy gap to be filled. Thus, the article looks into what has been achieved in terms of protection based on the rights needed to access and contain BCI 5.0 capabilities. It analyzes the threats to individual rights from unauthorized access to neural data, assesses the adequacy of regulatory approaches adopted worldwide for meaningful control of technologies and highlights the need for governance mechanisms to encourage ethical and responsible BCI innovations, broader rights and opportunities available to users in respect of their neural data, and for protection of rights.

BCI 5.0 is emerging in a complex technological landscape shaped by huge neurotechnological changes, fragmented political ecosystems and strong private interests.

The potential disruptive power of BCI 5.0 comes from a number of trends, with the rapid progress of EEG software providing for high-definition wireless sensing [Musk N., 2019].

Portable devices such as headbands have a promise of ongoing ex-vivo brain monitoring [Das S. et al., 2021: 5746]. Advanced machine learning architectures can now decode cognitive states using their EEG signatures whereas new standards of communication such as 5G and WiFi 6 enable real-time data transfer between the brain and a cloud, only to open the door to widely available consumer BCI with unprecedented capabilities. However, with much utility promised, such ubiquitous access creates risk. EEG data carry sensitive markers of identity, psychology and intentionality valuable to advertisers, insurers and public agencies and potentially usable to secretly manipulate emotions, filter information and enable behavioral micro-targeting in an unsolicited way observed in the latest research of Facebook's emotional contamination.

Another issue on the agenda could be neurological discrimination leading, like genetic discrimination, to refusal of opportunities. Brain penetration could also effectively threaten user intentions and memories. Thus, uncontrolled BCI 5.0 systems, apart from their benefits, will critically threaten rights and liberties. These likely implications have been magnified by prevailing policy failures. For the most part, BCI applications are still unregulated and fraught with major legal gaps with regard to data access, confidentiality and security.

For example, direct access to personal thoughts, unlike communication, is not protected while only a few meaningful mechanisms ensure the transparency of BCI audit logs or user control over the joint use of neural data. Options to claim a compensation of damage from neurotechnologies are poorly defined, with a lack of specific guarantees to remove new BCI risks extending the scope of violation even more. In addition, global technology companies fast track BCI commercialization in absence of adequate accountability setups. In this regard, Facebook's aggressive acquisitions assume a combination of persuasive power of social media with direct access to cognitive vulnerabilities.

Technological monopolies would repeatedly get hold of user data for profit and manipulation, only to demonstrate the threats inherent in such access to neural data. Their unparalleled resources and lobbying power can dishabilitate any policy response to protect individual rights. Governance gaps and incentives for anti-social business models make regulation an urgent focal point for assuring public interests.

This study assumes that legacy regulatory mechanisms for data protection in biometric, telecom and computer technologies are structurally inadequate in the face of new capabilities of BCI 5.0 that involve direct access to neural processes. It is assumed that only a hybrid regulatory system combining special provisions with technological guarantees embedded into BCI architecture via computational law mechanisms can effectively protect individual digital rights at the time of the fifth generation neural interface. To test this hypothesis, a profound review of emerging opportunities, constraints and risks of BCI 5.0 is performed to inform the plausible design of comprehensive political and technical guarantees for ethical innovations respectful of user rights promoting socially valuable applications.

It has a sense to discuss the study's purposes and objectives. First, the likely benefits and risks of BCI 5.0 are made clear in the light of modern understanding of neural science and documented technological paths

for empirically grounded assessment of problems to be addressed. Second, the article offers an overview of the current legal framework from the perspective of adequacy while also identifying gaps in the meaningful regulation of BCI 5.0 capabilities. It also specifies a key objective: it is necessary to have a minimum set of viable standards and mechanisms for BCI 5.0 adapted to its new technological properties to encourage secure, privacy safe, user-controlled systems. This is followed by a description of extra legislative policies and tools of computational law that will allow individuals to better protect their rights. Finally, one of the purposes is to propose guiding principles and recommendations to various stakeholders on the basis of summarized conclusions.

These purposes entail the following objectives of research: a) an in-depth technical overview of the emerging methods including neural network sensors, focused ultrasound neuromodulation, Neuro Mesh implants and AI architecture to support BCI 5.0 applications; b) classification of likely threats to the above rights including unauthorized data access, user behavior manipulation and compromised security based on documented vulnerabilities and predictive scenarios.

Third, the study includes an analysis of the existing laws and assessment of their outreach to effectively address the issues of BCI potential. Fourth, it provides a description of technological guarantees (such as blockchain, differential privacy, federated learning) which can be harnessed to reduce BCI-related risks and embed policy standards. Fifth, there is a description of specific changes applicable to the effective law and a sample code of conduct or ethics charter for stakeholders in BCI. Lastly, the study purports to identify ways to ensure accountability, dispute resolution and liability assessment within the proposed structure.

The multi-level analysis is intended to design effective policies and technical guidance for ensuring security and ethical focus at the next stage of man-machine integration for developers, regulators and users. Recommendations should strike a balance between encouraging useful applications and designing preventive risk reduction policies by providing a roadmap to responsibly navigate the emerging neurotechnological frontiers.

To achieve the above objectives and test the proposed hypothesis, the study relies on a comprehensive methodology embracing three interrelated approaches.

Multidisciplinary data collection and synthesis. The study brings together a variety of fields of knowledge: technical sources from industry

journals (Nature Neuroscience, Neuron, Current Opinion in Neurobiology, Brain-Computer Interfaces) provide the details of new methods of neuron visualization/stimulation. Legal journals (Law Journal of the Higher School of Economics, Journal of Law and Biosciences, Journal of Law) make up the basis for analysis of regulatory implications. Multidisciplinary publications (Science and Society, Philosophy and Technology, Ethics and Information Technology, Innovation and Technology) allow to discuss technical issues in the context of rights, value and governance. As an extra source, patent databases, corporate reports and civil society contributions are used for comprehensive understanding of the BCI landscape.

Comparative analysis and inductive reasoning. The study compares BCI 5.0 extended capabilities with existing mechanisms for protection of data, privacy and security while analyzing gaps between technological capabilities and regulatory framework in various jurisdictions. Based on the identified inconsistencies, special guarantees and supervision mechanisms adapted to the unique properties of BCI technologies are inductively proposed. This approach allows to design political and technical responses to new social and technological challenges.

This methodology provides for empirically grounded, balanced approach to come up with advice that would account for both innovative potential and the need to protect individual rights at the time of the fifth generation neural interfaces.

1. BCI 5.0 Technological Capabilities and Threats

1.1. Detailed Overview of Technological Capabilities and Innovations of BCI 5.0

A number of achievements have enabled a transition from laboratory-based and largely stripped-down iterations to ubiquitous, almost seamless integration of man and computer. SDK, such as Facebook's Brain-2Bot, will use consumer EEG headsets for automatic smart instrument control and environmental navigation free of portable devices. Startups such as Paradromics and Cortical Labs (Table 1) are working to make less heavy EEG sensors for high-density recording through the skin at resolutions comparable to FMRT [Sun Y., 2020: 310–324]. Thanks to the progress of machine learning methods, imagined speech and intended movements are now identified from neural activity together with semantic representations.

Taken together, these trends translate into "hands-free real-time interaction" with digital systems given the sole intent. Whereas an early BCI texting interface would identify EEG correlates of letters to type 90 characters per minute [Chen X., 2015: E6058-E6067], a recently decoded speech attempt has resulted in onscreen rate of 150 words per minute. This example gives an idea of how quickly we can have a seamless direct brain-computer link. However, compared to understanding, texting or dictation is a fragmented capability. The efforts to reconstruct perceptive experience, memories, emotions and conceptual thinking from decoded neural patterns foreshadow radically higher BCI throughput.

Table 1. Cortical Labs key features

Feature	Description
Biologically plausible neural networks	Research and simulation of neural network structure and behavior in animal/human brain for realistic AI design
Neuromorphic chips	Designing specific neuromorphic processors (Anthropic Neural Computers, ANC) optimized to launch such biologically plausible networks
General artificial intelligence	Models for general intelligence rather than specific tasks able to solve a wide range of cognitive problems just like man

Neuron visualization achievements set the stage for developing a capability to *read out* thoughts. Kernel's brain-chip interface attempts to capture hippocampus activities and to externalize memories [Hassabis D., 2021: 493–498]. Facebook's sensory headband aims to decode coded speech for augmented reality devices. Neuralink's 3000+ channel readouts have enabled real-time forecasting of limb movements in primates [Musk N., 2019]. Simultaneous innovations in stimulation technologies allow to *record* sensory and cognitive data. Examples of bidirectional communications are the experience induced in patients by temporal lobe stimulation and optogenetic induction in rodents.

Current developments also hold a promise of remote, wireless and possibly covert capabilities. Ultrasound neuromodulation can transcranially influence brain areas without a need for implants [Menz M., 2021: 2919-2933] while EEG biometry is capable of discreet user authentication [Sun Y., 2020: 31005]. The emerging reconfigurable neural sensors

can detect chronic states of the brain [Seo D., 2020: 1-17]. Portability also allows to track users in different environments, for example, as envisaged in Facebook's VR BCI. Miniaturization allows to embed applications as in Smart Stent's neurovascular interface (Table 2). The said trajectories are fraught with far-reaching implications affecting cognition, identity, privacy, behavior, justice and social cohesion, all of which require further discussion.

Table 2. Smart Stent key functions

Function	Description
Minimally invasive implantation	The device is to be implanted into brain blood vessels transvascularily without a need for open brain surgery
Brain activity recording	The device is to record signals from brain areas responsible for motion control
Auxiliary device control	Decoded neural data are to be used for control of external robotic systems, exoskeletons, other rehabilitation devices

1.2. Classification of the Key Threats to Rights and Liberties

In absence of proper supervision, the above adaptive capabilities will create major threats classified by this study in light of governance priorities. These threats include unauthorized data access, manipulative and discriminative applications, non-transparency and non-accountable commercialization.

Once ubiquitous, personal data collection creates a new risk of identity theft, emotional manipulations and discriminative refusal of opportunities. In fact, EEG biometry has been shown to distinctly identify people, with psychological profiling becoming a new application in its own right, only to result in unauthorized access or tracking. Selective data filtration based on decoded neural states will amount to manipulative censorship. In absence of proper checks, the identification of neural markers of risk, disease or demographic profile is a signal for predatory exclusion from service, a cognitive equivalent of genetic discrimination.

Direct neuromodulation is fraught with a number of extra risks of behavior compromise. Animal studies have shown that induced stimuli will cause specific behavior — for instance, one study [Adamantidis A., 2015: 420-424] points out to the potential for unauthorized influence.

Sensory manipulations can create neural evidence in favor of invalid assertions or sow discord by distorting perception and memories in event witnesses. These capabilities red-flag a forced and deceitful use to call for an extra level of control. They highlight the importance of the boundary between therapeutic applications for improved well-being and those that do not respect the autonomy of individuals.

1.3. Weak Security Provisions

Experiments with simple methodologies have demonstrated a potential for embedding malware into the brain via consumer headsets, neural signal spoofing and EEG data theft [Sun Y., 2020: 310]. With direct access to executive functions, BCI 5.0 will multiply the potential power of ransomware. Compromises between encrypting neural data and allowing crucial application are still an open question. Moreover, non-clinical BCI applications bypass supervisory standards for health devices despite health risks caused by direct brain stimulation. Such vulnerabilities highlight the need for special guarantees.

Non-transparency of business applications is itself a cause of concern since the incentives of dominant companies will often conflict with user well-being. In fact, the past study of Facebook's emotional contamination is an illustration of the willingness to discreetly manipulate users. With an opportunity to access or impact individual thoughts and feelings, behavior could become subject to unprecedented threats of persuasive power facilitated by the absence of guaranteed transparency and democratic supervision. With such applications deployed on a massive scale, proactive governance to prevent harm is a matter of priority.

2. Legal Regulation of BCI and its Constraints

2.1. Current Regulatory Principles Adopted Globally for Innovative Technologies

A policy framework for protecting individual rights from the emerging BCI applications should rely on the key governance principles designed for the existing technologies capable of affecting the brain.

The discussed international standards show a commitment to human rights, with a majority guided by democratic considerations in adopting the technologies that have an impact on human life. The declared prin-

ciples of international law include the universal declaration on bioethics and human rights that asserts human dignity, autonomy and consent in medical interventions on the brain. However, such declarations are devoid of mechanisms for enforcement which is left to the national law. Thus, the outcomes of protection will vary between jurisdictions.

Legal scholars believe direct access to thoughts to be sensitive enough to call for stricter supervision than is imposed on biometric and communication data. While some argue for applying restrictions only to ways of retrieving information outside human control, others are not as sure in respect of voluntary applications. Still others argue for control to maximize user autonomy over neural data flows. BCI also require informed consent with new designs that allow to use dynamically revocable and granular permissions.

In absence of specific rules applicable to BCI, some instructive precedents come from adjacent areas. For example, the law governing medical devices, human subject research and consumer goods offers comparative points of reference as for requirements to BCI system quality, safety and accountability. Protection of health data signals a need for cyber-security and access control policies to include, apart from openness and follow-up supervision of experiments to simulate high-risk BCI applications, advice on institutional bioethics. These mechanisms can be adapted to account for unique problems arising in BCI studies. Overall, the current framework is respectful of human dignity but also highlights awareness of the need for careful scrutiny in rapidly developing areas. Meanwhile, the current controls can only deal with new issues such as the constancy of access to thoughts, with a balance to be struck using the available principles as a backbone to provide special guarantees for BCI 5.0 capabilities.

2.2. Constraints of the Rules for BCI 5.0 Adoption

Making BCI 5.0 integration socially useful and respectful of rights requires governance adapted to new technological features. Conversely, the analysis shows constraints of direct application of the existing legal framework designed largely for biometric, communication and legacy computer technologies.

While intuitive logic assumes that access to thoughts requires higher levels of protection than those afforded to behavioral data, few of the existing regulatory differences will recognize this boundary.

As such, global data rules remain purely information-driven as they restrict the use of collected data. Real-time access to neural processes is out of scope of a vast majority of global data protection rules such as the General Data Protection Regulation (GDPR). Some persons believe it to be essentially a real-time access to another form of mental privacy that should be protected. There is a need for balance which would enable legitimate use of technologies without violating the boundaries of mental privacy.

The same ambiguity surrounds the concepts that define BCI admissibility. Freedom of cognitive improvement is a principle upheld by international policies, with coercive practices being forbidden. Meanwhile, even the legal definitions of coercion will normally revolve around obvious force or threat, only to exclude the opportunity for more delicate manipulations enabled by BCI. Therefore, a more nuanced governance should be established to distinguish applications for positive cognitive reinforcement from those that undermine mental integrity.

For example, the present-day rules will focus on illicit use or dissemination of information. BCI have exclusive access to thoughts even where ongoing recording is not assumed. Control of the actual real-time data collection is thus desirable due to implied sensitivity. An adequate way of protecting designed for BCI's constant neural access routes involves multi-level access control and revocable permissions, a deviation from a normal data protection framework.

Finally, sectoral regulation will often exclude consumer technologies even of high social impact, with higher standards applicable to medical equipment for restricted access to therapeutic devices. Once adopted, multi-level supervision for a balance between innovations and proportional control of high-risk applications can overcome this constraint. On the other hand, responsive governance is achieved by adapting some of the aspects such as tentative market overview and post-launch monitoring of BCI elaborations across all sectors.

2.3. Implications of Inadequate Legal Protection of Digital Rights

In absence of meaningful guarantees to match special capabilities enabled by BCI 5.0, the above categorized risks to individual rights become highly probable. Beyond breaches of neural privacy and actions as such, they threaten to routinely erode civil liberties as a whole.

What is the most disturbing, unregulated BCI commercialization can make normal the breaches of neural privacy that will be intuitively perceived as negative. Legitimizing such access, even to a minor extent, creates alarming preconditions for thought control and ideological coercion by authoritarian governments in the future. Moreover, if protection of civil liberties from violation is not there for too long, human rights will be threatened [Jobin A., 2019: 389–399]. Careful approach is urgently needed because the slope from benevolent to repressive use is slippery [Hildt E., 2021: 1–12].

Finally, it is decentralized technical design that offers a unique potential for upholding rights in the emerging sectors. With data access architectures broadly available across societies like social media previously, post-factum regulation is unlikely to rectify violations. Embedding the elements of security, protection and consent control into the technology itself means that protection will be there in the event of adaptation. Political and technical guarantees are joined to reliably secure the interests of individuals from the emerging threats.

A lack of security provisions in BCI 5.0 is threatening to make low cyber-security standards normal despite being dangerous for public well-being in a broad sense. Designs supporting transparency, access protection and audit control would create incentives for a cultural shift towards data management. The risk of neural data theft and compromised sovereignty is socially unacceptable: security and democratic supervision should become priority number one. Unprecedented risks arising from uncontrolled BCI commercialization coupled with constraints of the legacy political framework highlight an urgent need for innovative governance to protect individual rights to cognitive freedom.

3. Technological and Legal Safeguards for Protecting Digital Rights in the Context of BCI 5.0

3.1. Technical Guarantees for Better Security and Privacy in BCI 5.0

Apart from policies, technical principles of design and architecture can also provide ways for secure and ethical evolution of BCI 5.0 ecosystems respectful of human rights.

Federated learning enables collaborative model training on user data without exchanging the data themselves to preserve privacy. Thus, us-

ers can benefit from crowdsourcing applications while preventing exploitation of their neural patterns. Mixing proxy data via algorithm development methods such as differential privacy leaves less room for re-identification while providing insights. Access to insights for achieving improvement without damaging user privacy is one of technological pillars of an ethical BCI.

Another core safeguard involves encryption and control of protection in line with the standards of the Health Insurance Portability and Accountability Act for electronic health records, something that ensures security in preserving utility. For instance, selective disclosures such as cryptographic registration details to confirm identity attributes without revealing raw biometric data will protect user interests via approaches proposed by [Soares J., 2012: 149–155]. Encryption of high-density neural data flows is currently constrained by the level of computational overheads.

Blockchain architectures (decentralized ledger) will also support secure audit for access and permission management service under user control. Action support mechanisms in connected technologies point to smart contracts that ensure limited purpose and revocable data exchange by cryptographic consent tokens rather than unconditional access. It is a combination of computational law tools with adaptive policies that can provide robust protection of the rights.

Human-centric privacy, accountability, democratic supervision technologies are indispensable supplements to top-down regulation where individual interests and liberties are to be protected from the emerging threats. Multi-level governance will blend the strengths of these approaches in a way that securely expands the potential of socially useful innovations while containing risks generated by the emerging neurotechnologies.

3.2. Normative Minimum Viable Standards for BCI 5.0

Security and privacy will be critical for establishing basic regulatory norms and expectations that are important for meaningful provision of individual rights and interests of BCI system developers and users.

A vital prerequisite for ethical integration of BCI could be the assurance of improved protection of mental privacy beyond what applies to communication and even biometric data due to special sensitivity of this issue. These legal definitions are supposed to prevent real-time access to

neural processes, not only permanent recording. While a need to ensure lawful applications requires nuanced approaches without total prohibitions, it also requires to avoid uncontrolled distribution.

This translates into a higher denial threshold before neural data could be collected or used in consumer applications, something really in line with medical ethics and proportional protection. In particular, consent-giving via multi-factor authentication for daily use or passive monitoring assumes a higher threshold than one-time approvals now predominant in digital systems. Dynamic revocation and granular permissions will additionally secure user actions. This will put the burden on the developer who should substantiate the need for access.

Technological protection is another pillar. The requirements modeled upon HIPAA security standards — those for tracking, logging and attempting to prevent healthcare data breaches — provide for accountability via data encryption, access control, audit and a lot more in fighting abuse and cyber-threats (Fig. 1).

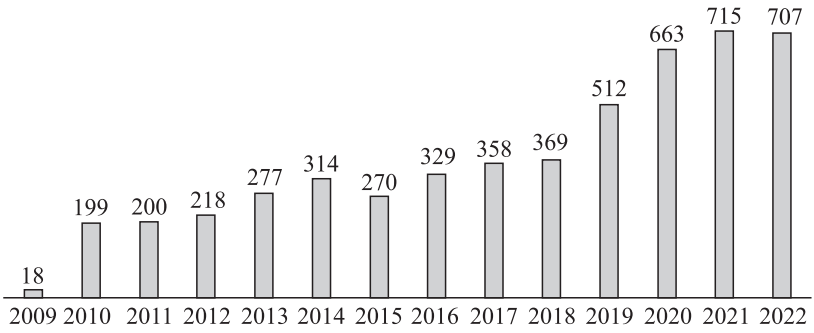

Fig. 1. Healthcare data breaches (reported by HIPAA, 2009–2022)

Federated analytics, differential privacy are some of the ways to maximize utility without harm to user interests. Security provisions designed for BCI threats ensure continuous protection along with applications.

Finally, the development of responsibility and compensation arrangements reflects recognition of the fact that some of the emerging technologies will cause harm even despite due care. Well-informed ways for compensation of damage, flexibility to adapt to ever evolving stock of impact evidence and participatory mechanisms in supervision regimes could ensure accountability. In combination, such basic reasonable guarantees strike a balance between unfettered innovations and provision of necessary safeguards against BCI pitfalls.

3.3. Material Amendments to the Data Protection act in Light of BCI-related Challenges

Full drafting of policies and rules for disruptive adoption of BCI requires to revisit the existing policy framework designed largely in response to technologies of the past. Given below are specific amendments to support priority reforms for more comprehensive rights protection in light of the analyzed risks.

While the effective law has a strong focus on regulating how the collected information is used, real-time access to thoughts requires better protection at the very initial level because sensitivity passes by the authorization and use restriction requirements [Ienca M., Haselager P., 2016: 117–119]. Provisions that restrict unwarranted collection of neural data combined with the existing rules of use will provide consistent protection.

Narrow definitions of coercion and inappropriate influence in regulating the persuasive technologies should be expanded to account for intricacies in BCI. The evidence that neural processes can be manipulated to induce relationship, behavior and memories without any obvious force or deceit means that governance should counter such influences on psychic integrity.

Moreover, customized supervision mechanisms will address the problem of combining medical and consumer uses of BCI and will balance innovations with supervision in proportion to the identified risks. The rules may require high-risk interventions to be subject to security checks in the same way as pharmaceuticals or medical devices while transparency provisions target the applications designed for consumers. Thus, nuanced models can enable the consideration of specific risk profiles.

Responsive governance is possible via the expansion of rights and methods of compensation for damage combined with flexible liability funds compensating documented harm. With such arrangements, mandatory insurance of developers from verifiable abuse will contain risks while allowing unfettered innovations by avoiding preventive restrictions, and will make these emerging laws compatible with BCI realities.

3.4. Codifying Rights and Restrictions for BCI 5.0 via Computational Law

Apart from policies, there is a good chance to embed the rules of legitimate use and protection of rights into technological architectures

via computational law. Smart contracts will codify ethics embedded into technologies to enable granular, dynamic and transparent consent management. Users can preset access restrictions to be automatically enforced to prevent any future abuse. A certain revocation of consent can trigger guaranteed cascade deletion of data. It is these computational iterations of law that that contribute to fail-free protection.

Such applications also allow real algorithmic output-related events to control codified supervision and regulation. Third-party audits certified to approve sound data processing practices could automatically extend operating licenses. Problem reports by representative civil juries can trigger inquiry and rectification cascades. The final goal is to embed social checks and balances via computational law and to uphold accountability.

Overall, careful integration of legal principles directly into technological architectures in an inventive way allows to preempt risks while ensuring unfettered innovation. Rather than responding by restrictions, computational law options offer proactive protection of individual rights and interests holding an enormous promise for ethical integration with neurotechnologies.

3.5. Automating Protection of Digital Rights Using Smart Contracts and Oracle AI Agents

A very promising area for drafting and enforcing policies via the emerging technologies of computational law — as in smart contracts, decentralized apps and tokenized consent systems — is automated enforcement and monitoring of policies proposed for protection of individual rights in the context of BCI.

Dynamic permission tokens can codify the above proposed type of granular consent policies directly into access control infrastructure via smart contracts. Users could manage such permissions on their own, for instance, by deleting EEG data exchange for business while maintaining all access data for health purposes. Pre-programmed rules can trigger the necessary deletion of data when the purpose expires. General consent management will also avoid any dependence on external coercion.

Meanwhile, AI agents trained as LegalTech applications can algorithmically identify breaches of codified rights to protection by system logs and user reports. As an illustration, applications could note unauthorized passive neural monitoring or flag business applications failing a risk assessment. To provide quick protection, applications could auto-

matically generate warnings, escalate to human review and technically disable systems when breaches exceed probative thresholds.

Wider/decentralized autonomous organizations where users manage policies on their own to balance innovation risks on a peer-to-peer basis rather than under a formal corporate order also offer promising ways to uphold rights. Such codified iterations of legal principles allow to go beyond upgrading restrictions on the use of technologies to a technology designed for mutually conceived supervision, with potential benefits explored in parallel with policy development.

Overall, the proposed guarantees, once implemented directly in the code, can sustainably ensure that the layer of rights is resistant to regulatory destruction. The technological architectures that embed supervision and balance the incentives for innovation with social well-being can complement some key reforms on the way to the ethical neurotechnological future. This will require intensive multidisciplinary collaboration all along the way — from conceptualization to implementation and testing.

4. Liability and BCI-Related Dispute Resolution

4.1. Methods to Demonstrate Claims for Compensation and to Settle Disputes in the Event of Unauthorized Use

In the pursuit of risk preemption, good governance will also assume setting up specific mechanisms for rectification in case of verifiable damage that can arise even with strict guarantees in place.

Encrypted logging and watermarking methods allow to identify a single path of traceable evidence of unauthorized use thus ensuring restitution. For example, users can register personal EEG signatures to allow for attribution as soon as such activity is accessible or synthesized by unauthorized parties. Embedded digital watermarks allow to check neural data for commercial appropriation and licensing breaches. Any abuse should be proved with inalterable records for possible further action.

Proportional liability funds supported by mandatory security deposits rather than penalties or criminalization will facilitate settlement. Claims can be processed and compensation distributed by democratically governed independent supervisory boards with civil membership.

In other words, the availability of probative evidence resulting from novel judicial methods along with channels for compensation will make

it possible to use non-punitive but rectifying mechanisms to uphold justice. The regulatory design's focus on direct mitigation of damage rather than on preventive restriction can bring benefits while securing reliable guarantees. Technological and political synergies can provide a robust protection of individual rights.

4.2. Current Regime for Distribution of Liability for Damage

hile striving to minimize unnecessary damage, a realistic assessment will recognize that unintended effects from rapid progress of the emerging technologies such as BCI 5.0 are inevitable. Applying the existing legal principles with regard to distribution of liability for so-called "unintended but inevitable" harm assumes a point of departure where there is no provable malice of any kind.

The effective regimes for products admit different distributions of liability between producers and consumers based on the analysis of due care on both sides in light of the reasonable care standard [Miller J., Goldberg R., 2004: 149-155]. Producers adhering to the acknowledged best practices would face limited liability for unforeseen errors. However, consumers in violation of the due care obligation (such as failing to turn on security functions) would face the distribution of liability within this extent.

The application of similar principles to balance accountability, innovation and precaution in BCI use would uphold equity. Scenarios of unintended harm via compromised devices or careless use of functions would result in a mixed model. On the contrary, where security is weak due to negligence or deployment of risky unauthorized applications, it would be fully justifiable to impose stricter liability on producers. Overall, the existing nuanced framework offer some initial guidance on the arising problem.

However, new technological spaces also require to consider extended social liability models for aggregate public effects. Isolated disputes clearly inadequately capture the general implications of harm as BCI 5.0 is promising to be profoundly transformative both on individual and collective scale. Going deep into integrated compensation, rehabilitation and recovery systems to achieve real social outcomes provides the best opportunity to maximize the protection of rights. This is worthy of more careful scrutiny.

4.3. The Importance of Establishing Guilt in Criminal Activities with Compromised BCI Systems

While the previous sections deal with mitigating unintended harm, it would be realistic to discuss pragmatic adversarial settings in the face of quasi-dualistic nature of integrated neurotechnologies.

Seamless BCI 5.0 integration obfuscates agency attribution and, therefore, guilt for criminal action in systems made vulnerable by malignant actors. Where the executive function control was seized, it will be hard to identify with sufficient certainty whether the criminal intrusion was committed by the user or hackers. If the guilt cannot be ascribed, fair responses are difficult.

But arbitrary attribution of fault will punish the victims of manipulation. Too much zealous prosecution will also suppress incentives to report and disclose the information required for better protection. However, due to ambiguity, universal immunity escapes accounting, only to allow exploitation. On should proceed with care in these dilemmas.

Technological options such as blockchain-based data recorders, access and threat logs are potential sources of evidence to identify liability [Kshetri N., 2024: 117–119]. Behavioral forensics would establish a deviation from personal baseline as indicative of compromise. Still less than perfect reconstruction is a reality as to the existence of a barrier to satisfy probative thresholds. There is a special need to develop verifiable diagnostics [Froomkin A., 2020: 513].

This broad problem underlines tension between justice, freedom and security due to the risks and ambiguities arising from BCI. But a repressive bend would be as much dangerous as reckless indulgence. Governance projecting the importance of sincere strife to the truth and reconciliation leads to socially approved outcomes. A multidisciplinary analysis that necessarily follows will discuss ways to uphold ethics.

4.4. Call for an Ethics Charter to Prevent Unauthorized Use

Interrelated risks in all these analyses point to the development of a culture of responsibility to secure socially useful future outcomes.

One such setup would include the principles of necessity and proportionality just as those of consent and privacy, transparency and accounting, harmlessness via inclusive discussion. It would define the duties of

producers that consider social aspects of effects, characterize risks, embed protection systems into technological design, provide remedies in the event of harm, and discourage harmful business models as much as currently possible. The relevant duties of users would include bona fide consent-giving, problem and unauthorized use reporting, and provision of feedback for system improvement.

Charters endorsed by producers and representative consumer groups define voluntary but mutually binding obligations in accord with the proposed regulatory guarantees. They carry non-punitive signals beneficial for public confidence and create provisions. While the framework will need an upgrade in view of the lessons learned and expectations, the original pacts will lay the brickwork for subsequent collaboration.

Despite inherent risks due to rapid dissemination, charters embodying ethics through corporate responsibilities achieve responsiveness and self-regulation. They extend powers to stakeholders rather than create isolated authoritarian restrictions. While value-based partnerships cannot do without formal policies, they have been found to meaningfully uphold secure and equitable innovation paths [Yuste R. et al., 2017: 159–163]. Science, politics, business and society — all should join forces to secure this obligation.

5. Import of Findings and Their Implications

The study has endeavored an in-depth analysis of the emerging capabilities and constraints of BCI 5.0 to design policies and technical interventions aimed at protecting user rights. The above sections contain a synthesis of robust technical assessment, comparative policy analysis, predictive risk modeling and the relevant proposals for governance.

Predictive analyses based on the experience of related sectors confirm the emerging threats created by commercialization unbridled by incentives that agree with user well-being. They highlight how governance should discourage antisocial uses before their dissemination is deeply rooted. Nevertheless, excessive care is fraught with the risk of containing useful development. Navigating through these competing tensions requires nuanced, adaptive and multimodal interventions that were proposed here.

Drafting minimum viable guarantees and amendments to upgrade protective framework and tools for application of computational law provides some of the ways to maximize opportunities for securing rights.

Embedding ethical practices directly into technological architectures and organizational models reliably secures protective capabilities — often beyond what is achievable by the external regulation [Frolova E., Lesiv B., 2024: 15]. Synthetic recommendations on technological, political and cultural interventions provide comprehensive guidance to implement positive future outcomes.

Overall, this integrated technical and social analysis presents key ideas and tools to inform the efforts to prepare stakeholders for forthcoming dissemination of neurotechnologies. As such, this is an enormous step towards human improvement that needs to be re-formatted for preemptive governance in going forward towards equitable innovation to secure beneficial outcomes across society.

6. Current Analysis Constraints

While this study achieves an extensive range through a synthesis of social science, engineering, legal and ethical perspectives, its findings should be treated with care recognizing inherent constraints that stipulate their use and identify steps to follow.

As the most general point, all analyses are underpinned by efforts to predict what is likely to occur in the near future but is still in the making. Though they are based on demonstrated prototypes, the exact functionality that leaves room for risk is an open empirical question. Real practices may deviate from forecasts in ways that cannot be foreseen.

Moreover, complex social and technical phenomena have evolved due to unpredictable shared constitution between technological and social entities [Volos A., 2024: 90]. Statistical analyses are in peril of neglecting the emerging future outcomes with new opportunities inducing unforeseen uses, adaptations and harm in need of permanent reassessment; therefore, one should monitor the ongoing co-evolution.

In this way, the discussed study provides a necessary basis for multiple paths via prospects and risks of a rapid launch of integrated neurotechnologies. Wise use, however, is necessary where foresight has reached its limits. Ongoing reassessment and understanding of the arising divergence is needed to stay on track.

The current findings indicate a number of key pathways for further research and studies of the questions that are left open.

Technical studies of analytics for preserving privacy of high-density neural data flows would finally enable progress in the proposed guaran-

tees, with better consistency between non-invasive BCI and implanted systems to improve the diagnostic and therapeutic utility. Moreover, studies of user interfaces for effective consent and understanding of risks are crucial for creating human-centric designs.

Finally, studies of transition management approaches that link the upcoming business realities with long-term aspirations will help maintain pragmatic focus. For example, studies of voluntary sectoral ethics charters can provide insights into the best practices for early efforts. Practical testing of the proposed computational law tools assumes checking their efficiency in the real world. Such empirical steps to translate principles into reality remain an important complement of conceptual policy development.

These pathways to perfection demonstrate how responsible BCI innovations could be maximized. Taken collectively, technical, sociological and philosophical understanding of success to be achieved can provide the basis for interaction with already occurring fundamental shifts and for joint projecting of equitable outcomes in the future. The discussed study provides a tentative structured outline for such urgent joint endeavors.

Aggregating the identified opportunities, gaps and risks results in a balanced set of political, technical and cultural recommendations that allow to responsibly steer the implementation of BCI 5.0 capabilities while securing social values and rights. In particular, it is recommended to:

design in light of the above discussion of opportunities, gaps and risks a multi-level adaptive policy recognizing the unique risk-benefit compromises in BCI applications while avoiding universal governance;

require higher consent modeled on medical ethics for access to neural data due to sensitivity of the issue;

draft technical standards and design incentives for better security and privacy, and for user supervision opportunities;

provide proportional mechanisms of accountability and compensation for verifiable but unintended harm;

encourage civil participatory supervision and multidisciplinary expert contribution to governance;

embed ethical principles and protection directly into technologies via computational law wherever possible;

encourage collaborative sectoral self-governance via associations and voluntarily adopted ethics charters;

invest into multidisciplinary predictive studies to inform the emerging policies;

provide for more civil engagement and participative innovation design in agreement with social values;

achieve international consensus on fundamental principles with room left for regulatory diversity.

With such holistic, adaptive, human-centric governance, the transformative potential of BCI 5.0 could be equitably and safely geared to serve the purpose of prosperity for all. Sustainable, inclusive public discussion combined with bona fide policy design can thus channel these historical opportunities towards moral goals.

Conclusion

The study is an attempt of multidisciplinary research of the emerging BCI 5.0 systems to propose special governance arrangements for balancing capabilities brought by innovation with preventive rights protection.

It provides an overview of the current achievements in BCI that are rapidly approaching ubiquitous and seamless man-computer integration for unconstrained communication, control of environment, extended memory and a number of other improvements potentially within reach. They are also fraught with risks of ongoing neural data monitoring, hostile manipulations, compromised security and other breaches in absence of appropriate guarantees.

A comparative overview was conducted to understand constraints for direct application of the existing legal framework for privacy, security and protection of users in adequately managing BCI capabilities. Regulatory gaps pending removal were identified in respect of real-time data access and use, updated definitions of mental privacy and especially adaptive approaches to monitoring. An uncontrolled progress of these technologies can mean normalization of such invasive practices.

Predictive modeling based on what has been learned from the related sectors of persuasive computing, biometry and personalized medicine highlights the likely risks of poorly coordinated economic incentives and inadequate guarantees. Thus, innovative governance will be required to avoid potential threat for individual and collective rights.

A synthesized structure of specialized guiding principles currently allows to design a road map for bidirectional tracing of political and tech-

nical paths towards preservation of rights and achievement of socially useful outcomes. Stakeholders could systematically implement the recommendations via the following practical steps:

Policymakers could continue the dynamic upgrade of mental privacy laws to regulate real-time access, address supervisory gaps and tighten the requirements to user safety using the proposed multi-level risk model. A phased introduction can enable iterative improvements.

This could include the implementation of privacy preservation architecture, ethical risks assessment and transparency obligations, and even the development of the best practices and supervisory bodies for the sector as a whole — everything that can promote accountability. Users would thus have a voice in respect of security provisions, informed consent, duty of care in the process of use, expression of concern and request for damage compensation mechanisms for independent agency in BCI device integration.

A combined implementation of such recommendations can ensure equitable achievement of many critical advantages.

Acting on the proposed guiding principles could catalyze ethical innovation ecosystems in BCI in the interest of scientific community, businesses, regulatory bodies and the public at large.

A focus on privacy and security can expand the range of BCI research by enhancing user confidence in secure neural data exchange. Introducing the ethical review and supervision mechanisms will improve the conduct of research.

Responsible innovative channels will create long-term social and regulatory confidence crucial for sustainable success. Voluntary self-regulation will preempt restrictive policies that hold back progress.

For the public at large, innovative trajectories focused on social values and rights will generate more options for useful access. Channels for monitoring and damage compensation create ways for participation. Overall, general responsibility can create a driver for significant progress in the living standards of population.

A paradigm shift for ethical innovative BCI ecosystems should be realized by all stakeholders in a coordinated way.

Policymakers should promote multi-level regulatory models to balance unfettered innovations with supervision of high-risk applications based on the established principles; propose incentives to design privacy

preserving architectures; systematically engage expert community and civil society for contribution to the development of specialized governance; and invest into foresight to manage adaptation.

Developers should embed transparency, auditability, secure design and user-led supervision into block chain design and development options; adopt the practices of ethical risk assessment and monitoring; take part in promotion of the best practices and professional association ethical standards.

Users should be allowed to provide informed consent for BCI use including privacy provisions. They should take precautions for use and monitoring; provide feedback and report problems to help improve systems; demand efficient claim processing mechanisms.

Researchers should study social, ethical and legal implications of BCI use in a wide range of sectors. They should explore practical ways to implement the proposed guarantees and guiding principles; provide inclusive discussions and participatory supervision mechanisms.

Civil interest groups can monitor BCI achievements and commercialization, raise issues and advocate policies and business models that serve the interests of society.

This could include civil society participation in responsible innovations. BCI 5.0 prospects are a cause of surprise and concern. Nevertheless, equitable and safe development respectful of rights can bring promising future outcomes to reinforce human potential in all communities. Let out collective action rise to the challenge with urgency and wisdom required by the current moment.

References

1. Adamantidis A. et al. (2007) Neural substrates of awakening probed with optogenetic control of hypocretin neurons. *Nature*, 450 (7168), pp. 420–424.

2. Allen C. et al. (2020) Artificial morality: Top-down, bottom-up, and hybrid approaches. *Ethics and Information Technology*, vol. 22, no.3, pp.149–155.

3. Anumanchipalli G. et al. (2019) Speech synthesis from neural decoding of spoken sentences. *Nature*, 568 (7753), pp. 493–498.

4. Bublitz J. C. (2013) My mind is mine?! Cognitive liberty as a legal concept. In: Freedom of the Mind. Baden-Baden: Nomos, pp. 233–264.

5. Burwell S. et al. (2017) Ethical aspects of brain computer interfaces: a scoping review. *BMC Medical Ethics*, vol. 18, no.1, pp. 1–12.

6. Chen X. et al. (2015) High-speed spelling with a noninvasive brain–computer interface. *Proceedings of National Academy of Sciences,* vol.112, pp. E6058–E6067.

7. Das D.M. et al. (2021) Brain-computer interface: Advancement and challenges. *Sensors*, no. 17, p. 5746. https://doi.org/10.3390/s21175746

8. Frolova E., Lesiv B. (2024) Sources and Forms of Law: a Modern View on Basic Theoretical Provisions. *Law*. *Journal* of *the Higher School of Economics*, vol. 17, no. 1, pp. 4–39. https://doi.org/10.17323/2072-8166.2024.1.4.39 (in Russ.)

9. Froomkin A.M. (2020). Regulating mass surveillance as privacy pollution: Learning from environmental impact statements. *University of Illinois Law Review,* no. 2, pp. 513–572.

10. Gulyamov S.S., Rodionov A.A., Rustambekov I.R. and Yakubov A.N. (2023) The Growing Significance of Cyber Law Professionals in Higher Education: Effective Learning Strategies and Innovative Approaches. International Conference on Technology Enhanced Learning in Higher Education, pp. 117–119, doi: 10.1109/TELE58910.2023.10184186. https://ieeexplore.ieee.org/document/10184186

11. Hassabis D. et al. (2021) Neuroscience-inspired artificial intelligence. *Neuron*, no. 3, pp. 493–498.

12. Hildt E. (2021) Multi-person brain-to-brain interfaces: Ethical considerations. *Frontiers in Neuroscience*, no. 15, pp. 1–12.

13. Hiremath S. et al. (2015) Wearable Internet of Things: Concept, architectural components and promises for person-centered healthcare. In: 2015 International Conference on Wireless Mobile Communication and Healthcare-Transforming Healthcare through Innovations in Mobile and Wireless Technologies, pp. 304–307.

14. Ienca M., Haselager P. (2016) Hacking the brain: Brain-computer interfacing technology and the ethics of neurosecurity. *Ethics and Information Technology*, vol.18, no. 2, pp.117–119.

15. Jasanoff S. (ed.) (2015) Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power. Chicago: University Press, 360 p.

16. Jobin A., Ienca M. (2019) Global landscape of AI ethics guidelines. *Nature Machine Intelligence*, vol. 1, no.9, pp. 389–399.

17. Kramer A. et al. (2014) Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of National Academy of Sciences,* vol. 111, pp. 8788–8790.

18. Kshetri N. et al. (2024) Blockchain technology for cyber defense, cybersecurity, and countermeasures: Techniques, solutions, and applications. https://doi.org/10.1201/9781003449515

19. Lotte F. et al. (2018) A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. *Journal of Neural Engineering*, vol. 15, no. 3, p. 031005.

20. Menz M., Oralkan O. et al. (2019) Precise neural stimulation in the retina using focused ultrasound. *Journal of Neuroscience*, vol. 39, no.15, pp. 2919–2933.

21. Miller J., Goldberg R. (2004) *Product liability*. Oxford: University Press, 386 p.

22. Musk N. (2019) An integrated brain-machine interface platform with thousands of channels. *Journal of Medical Internet Research, vol.* 23, no. 10, p. e30903.

23. Naseer N. (2015) NIRS-based brain-computer interfaces: a review. *Frontiers in human neuroscience,* no. 9, p. 3.

24. Reijers W., O'Brolcháin F. (2018) Governance in blockchain technologies & social contract theories. *Ledger*, no. 3, pp. 1–17.

25. Sample M., Bauer Z. (2021) Brain-computer interfaces and personhood: Interdisciplinary deliberations. *Cambridge Quarterly of Healthcare Ethics*, vol. 30, no.1, pp. 157–169.

26. Soares J. et al. (2012) Privacy-preserving attribute-based encryption for brain-computer interfaces. *Journal of Medical Systems*, vol. 36, no.1, pp. 149–155.

27. Sun Y., Zhang H. et al. (2020) EEG-based biometric authentication: A comprehensive survey. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, vol. 2, no. 4, pp. 310–324.

28. Volos A. (2024) Concept of Weak Party in Civil Matter in Context of Digitalization. *Law. Journal of the Higher School of Economics,* vol. 17, no. 3, pp. 84–105. https://doi.org/10.17323/2072-8166.2024.3.84.105 (in Russ.)

29. Yeung K. (2021) Algorithmic regulation: a critical interrogation. *Regulation & Governance*, vol. 12, no. 4, pp. 505–523.

30. Yuste R. et al. (2017) Four ethical priorities for neurotechnologies and AI. *Nature*, 551(7679), pp.159–163.

Information about the author:

S.S. Gulyamov — Doctor of Sciences (Law), Professor.

The article was submitted to editorial office 28.10.2024; approved after reviewing 14.01.2025; accepted for publication 18.02.2025.

Legal Issues in the Digital Age. 2025. Vol. 6, no. 2.
Вопросы права в цифровую эпоху. 2025. Том 6. № 2.

E-Government

Research article
JEL: K23, K 38
UDK: 342.5, 342.7
DOI:10.17323/2713-2749.2025.2.161.182

Transparency in Public Administration in the Digital Age: Legal, Institutional, and Technical Mechanisms

Pavel P. Kabytov[1], Nikita A. Nazarov[2]

[1, 2] Institute of Legislation and Comparative Law under the Government of the Russian Federation, 34 Bolshaya Cheryomushkinskaya Str., Moscow 117218, Russia,

[1] kapavel.v@yandex.ru, https://orcid.org/0000-0001-8656-5317

[2] naznikitaal@gmail.com, https://orcid.org/0000-0002-3997-0886

Abstract

The article contains a comprehensive analysis of the very relevant topic of ensuring transparency and explainability of public administration bodies in the context of an ever-increasing introduction of automated decision-making systems and artificial intelligence systems in their operations. Authors focus on legal, organisational and technical mechanisms designed to implement the principles of transparency and explainability, as well as on challenges to their operation. The purpose is to describe the existing and proposed approaches in a comprehensive and systematic manner, identify the key risks caused by the non-transparency of automated decision-making systems, and to evaluate critically the potential that various tools can have to minimise such risks. The methodological basis of the study is general scientific methods (analysis, synthesis, system approach), and private-scientific methods of legal science, including legalistic and comparative legal analysis. The work explores the conceptual foundations of the principle of transparency of public administration in the conditions of technology transformation. In particular, the issue of the "black box" that undermines trust in state institutions and creates obstacles to juridical protection, is explored. It analyses preventive (*ex ante*) legal mechanisms, such as mandatory disclosure of the use of automated decision-making systems, the order and

logic of their operation, information on the data used, and the introduction of pre-audit, certification and human rights impact assessment procedures. Legal mechanisms for *ex post* follow-up are reviewed, including the evolving concept of the "right to explanation" of a particular decision, the use of counterfactual explanations, and ensuring that users have access to the data that gave rise to a particular automated decision. The authors pay particular attention to the inextricable link between legal requirements, and institutional and technical solutions. The main conclusions are that none of the mechanisms under review are universally applicable. The necessary effect may only be reached through their comprehensive application, adaptation to the specific context and level of risk, and close integration of legal norms with technical standards and practical tools. The study highlights the need to further improve laws aimed at detailing the responsibilities of developers and operators of the automated decision-making system, and to foster a culture of transparency and responsibility to maintain public administration accountability in the interests of society and every citizen.

Keywords

transparency; explainability; automated decision-making; artificial intelligence; legal regulation; *ex ante* mechanisms; *ex post* mechanisms; right to explanation; black box; protection of citizens' rights.

Acknowledgements: The research was carried out with the Russian Science Foundation grant No. 23-78-01254, https://rscf.ru/project/23-78-01254/.

For citation: Kabytov P.P., Nazarov N.A. (2025) Transparency in Public Administration in the Digital Age: Legal, Institutional and Mechanisms. *Legal Issues in the Digital Age,* vol. 6, no. 2, pp. 161–182. DOI:10.17323/ 2713-2749.2025.2.161.182

Introduction

Introduction of automated decision-making systems and artificial intelligence (AI) systems into the operations of public administration bodies marks a new era in the development of public administration, which can be loosely described as the "automation of public administration." Its main purpose is to increase efficiency, optimise resources and enhance the quality of government services that may be provided automatically, i.e. without direct human involvement. In this case, citizens interact directly with the technology envelope of public administration. Hence, this creates a range of challenges, and maintaining transparency and explainability of the decisions taken holds a special place among them.

Historically, the principle of transparency (openness) of public authority activities evolved as a fundamental guarantee that the authorities would be accountable to society, citizens' rights would be protected, and

the basis for trust between the state and its citizens would be laid. As decisions affecting the rights and legitimate interests of individuals are increasingly made or drafted without the direct participation of a human person (public servant), the so-called "black box" problem arises that consists in the opacity of the internal decision-making logic and the prerequisites for making a certain final decision.

Thus, the lack of understanding how and on what grounds the automated decision-making system has come to a particular conclusion undermines trust in state institutions, creates obstacles to juridical protection and is capable to lead to systemic violations of legal guarantees, and human and civil rights. Therefore, our article aims to provide a comprehensive analysis of and offer a system for existing and proposed legal mechanisms aimed at ensuring transparency and explainability of automated decision-making systems and AI systems in public administration. It explores the conceptual foundations of the transparency principle in the context of new technology realities, identifies the key risks associated with the opacity of algorithm systems, and critically assesses the potential and limitations of various legal instruments (both preventive ones, *ex ante*, and subsequent control ones, *ex post*) in addressing the issue under review. A special emphasis is placed on the need to integrate legal, organisational and technical approaches in order to establish an effective system of safeguards.

1. Conceptual Foundations of and Challenges to Opacity in the Context of Automation of Public Administration

1.1. Automated Decision-Making and Artificial Intelligence in the Public Sphere: Essence and Key Parameters

In the past years, public administration has been actively exploring the potential of automated decision-making, i.e. the procedure of making decisions where information technologies are used either to facilitate the formation of judgements by decision-makers, or to replace them, partially or completely. In this situation, it should not be of critical importance which particular technology (whether a simple rules-based system or a neural network) has influenced the outcome. Undoubtedly, the specificity of technology must be taken into account in creating a regulatory requirements framework for the development, implementation and operation of such systems. At the same time, the very fact that

the process of making a decision that affects the rights and freedoms of a person is automated plays the decisive role in determining the item subject to regulation. The existing law-enforcement practice confirms this: even automated decision-making systems that use software code and that, according to some classifications, do not belong to AI systems (e.g., self-learning systems) in a strict sense can influence the lives of citizens and the activities of organisations in very serious and sometimes critical ways[1]. In view of the above, one should positively assess the approaches of such systems of justice where the "automated decision-making process" as such is the special subject of regulatory influence, regardless of the complexity of the underlying system. It enables a broader and more technology-neutral legal regulation and thus covers the risks associated with automation[2].

Automated decision-making systems can be classified on various grounds:

by their application sphere: law enforcement, legislative, judicial activities;

by the level of their automation: partially automated (a human operator supports the decision-making process), delegated (the system initiates and makes the decision but hands over to a human operator in case of a problem), and fully automated decision-making;

by their legal significance: decisions that have direct legal consequences; intra-organisational decisions; decisions that have other significant effects.

by the technologies used: systems based on rigidly defined rules, systems based on statistical methods, AI-based systems (machine learning, deep learning, etc.).

---

[1] See: Automating Society Report 2020. Available at: URL: https://automatingsociety.algorithmwatch.org (accessed: 07.05.2025); Automating Society 2019. Available at: URL: https://algorithmwatch.org/en/automating-society-2019/ (accessed: 07.05.2025)

[2] See: Directive on Automated Decision-Making. 2019. URL: https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592 (accessed: 11.12.2023); Gesetz über die Möglichkeit des Einsatzes von datengetriebenen Informationstechnologien bei öffentlich-rechtlicher Verwaltungstätigkeit (IT-Einsatz-Gesetz ITEG) Vom 16. März 2022. Available at:мURL: https://www.gesetze-rechtsprechung.sh.juris.de/bssh/document/jlr-ITEGSHpP1 (accessed: 10.12.2023); Article 28(1) Förvaltningslag (2017:900); Articles 41 и 42 Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público; Article 35a Verwaltungsverfahrensgesetz (VwVfG); Articles L311-3-P311-3-1-2 Code des relations entre le public et l'administration

At the same time, by introducing automated decision-making, the state seeks to improve the efficiency of public administration, minimise "human factor" errors, cut costs, and reduce corruption risks. However, these advantages come with major challenges including threats to human rights, difficulty to ensure human control, the problems of diffusing responsibility and, of particular importance for our study, the fundamental issue of making such systems transparent and explainable.

1.2. The Principle of Transparency in Public Administration: Theoretical and Legal Dimension

The principle of transparency (openness) of the activities of public administration is the cornerstone of a modern state governed by the rule of law. Historically, the idea of the openness of power has come a long way from the first legislative acts (for example, the Swedish Law 'On Freedom of the Press' of 1776) to its global recognition and enshrinement in international documents and national legal systems, including the Russian Federation Constitution (Part 2, Article 24).

However, to characterise the phenomenon in question, modern Russian legal doctrine and legislation use terms that are different, although close in meaning: 'transparency', 'openness', 'transparency', 'glasnost', 'publicity', 'publicity' [Silkin V.V., 2021: 20–31]. Such diversity, as noted in the literature, "results in a certain conventionality in the use of this or that term, the blurring of the concepts in question" [Pogodina I.V., 2023: 29–31]. This may make it difficult to develop a unified approach to their enshrinement in law and to their enforcement in the specific context of automated decision-making systems.

Despite the nuances in terminology, the essence of the principle lies in a mode of functioning of public authorities, which ensures that information on their activities is accessible to the society, creates conditions for public control and participation of citizens in the management of state affairs, and promotes the development of mutual trust between the state and society.

The transparency principle in Russian law includes the following key elements:

information openness: the obligation of state and local self-government bodies to actively publish information about their activities (e.g., on official websites, and in the media) and provide this information upon requests from citizens and organisations. Federal Law No. 8-FZ

"On Access to Information on the Activities of State Bodies and Local Self-Government Bodies" of 09.02.2009 describes in detail the possible ways of ensuring access to information. These include its publication in the mass media (Art. 12), placement on the Internet (Art. 13, 14), placement in the premises occupied by the authorities (Art. 16), provision of information upon request (Art. 18), and others. Federal Law No. 149-FZ of 27.07.2006 "On Information, Information Technologies and Information Protection" also enshrines the openness of information on the activities of government bodies and free access to such information as one of the principles of legal regulation (Art. 3);

comprehensibility and accessibility of information: information should be provided in a form that ensures that it can be perceived and understood by a wide range of people, and not specialists only. As the Concept of Openness of Federal Executive Bodies (approved by the order of the Government of the Russian Federation of 30.01.2014 No. 93-r) notes, the "comprehensibility" of information is important;

civil society involvement and public control: transparency creates prerequisites for a constructive dialogue between the authorities and society, for citizen participation in the process of developing and making decisions. Federal Law No. 212-FZ of 21.07. 2014 " Fundamentals of Public Control in the Russian Federation" explicitly states that one of the tasks of public control is "to increase the level of trust of citizens in the activities of the state, as well as to ensure close cooperation between the state and civil society institutions" (part 2, Article 2). For example, the Rules for Disclosure by Federal Executive Authorities of Information on the Preparation of Draft Regulatory Legal Acts and the Results of their Discussion (approved by Resolution of the Russian Federation Government No. 851 of 25.08.2012) are aimed at implementing the principle of transparency. These Rules provide for compulsory posting of draft regulatory legal acts on the portal <regulation.gov.ru>.

accountability and responsibility of the authorities: the openness of the activities of the authorities allows the public to assess their effectiveness, identify violations, and hold officials accountable for their actions. As academician O.E. Kutafin rightly emphasised, in the modern period "state power responsible to the people and the law" is one of the main criteria for the establishment of constitutionalism [Kutafin O.E., 2008: 18].

developing and maintaining trust between the state and society: as enshrined in Article 75.1 of the Russian Constitution, "conditions shall be created in the Russian Federation for sustainable economic growth of

the country and improvement of the welfare of citizens, for mutual trust between the state and society." Trust, in turn, serves as the basis of social institutions, "uniting people, guaranteeing them security, the success of collective endeavours and allowing them to direct their combined energies for the common good" [Narutto S.V., Nikitina A.V., 2022: 13–18].

Thus, transparency is not just a desirable attribute, but a fundamental legal principle of public authorities' activity in a modern state governed by the rule of law. It has deep roots and has been enshrined in international acts and national laws including the Russian Constitution. The contents of this principle is quite diverse: it includes information openness, clarity and accessibility; society involvement; accountability and responsibility of authorities; and society's confidence in the government.

Implementation of the transparency principle helps enhance mutual trust between the state and society, improve public administration efficiency, prevent corruption, and protect citizens' rights. Still, to achieve real transparency it would be necessary not only to pass laws and regulations, but also to develop the corresponding culture in government bodies, and for civil society to take a pro-active stance. It is important to balance openness with the need to protect legal interests.

In this context the article offers a comprehensive analysis of the very relevant topic of ensuring public administration bodies' transparency and explainability in the context of an ever-increasing implementation of automated decision-making systems and artificial intelligence systems in their operations.

1.3. From Legislative towards Scientific Understanding of the Prerequisites for Maintaining the Transparency of Automated and Artificial Intelligence for Public Administration

Scholars in the sphere of legal science emphasize that in addition to enshrining transparency as a basic principle of public administration there are other prerequisites to enshrine the requirement of transparency and explainability of automated decision-making systems:

Trust is a significant aspect of automated decision-making, and explainability and transparency are necessary to increase and fortify this trust [Fine Licht de K., Fine Licht de J., 2020: 917]. Algorithm explainability is more important than algorithm transparency both for the ordi-

nary citizen and for the person making decisions [Grimmelikhuijsen S., 2023: 242] because explainability allows to reveal the cause-and-effect relationship between the input data, the logic of the system operation, and the automated decision made, thus contributing to understanding its validity.

In addition, sociological surveys compare citizens' trust in the case of decision-making with or without human involvement. E.g., one of them noted that when an AI system solved a "technical" job scheduling task, there was no difference in ranking, but for tasks requiring "human judgement," namely making a hiring decision, algorithms were perceived as less trustworthy [Lee M.K., 2018: 1-16]. Another study shows that citizens have less trust in automated decisions that "lack transparency." However, there is no transparency in the decision-making process even for the decision makers themselves [Schiff D.S., Schiff K.J., Pierson P., 2022: 653–573].

Explanation and transparency contribute to the creation of a safer and more reliable product, and enable collecting evidence for accountability [Sokol K., Flach P.A., 2019: 1–4]. This is especially important in the field of diagnosis and treatment, because in the absence of such requirements, the fundamental principles of medical ethics are jeopardised, which may negatively affect the safety of the individual and society.

Transparency encourages the human user to participate in the decision-making process, and explanations allow to correct and find technical errors in the automated decision-making system [Srinivasu P.N. et al., 2022: 1–20].

Explainability and transparency are necessary conditions of accountability for both the decision-maker and the operator of the automated decision-making system. Transparency is an informational aspect of accountability and as such is a prerequisite for accountability. And the individual right to information or clarification is only one of the elements of a broader structure of regulation and supervision [Wischmeyer T., Rademacher T., 2020: 75–101].

Lack of algorithm transparency can hide discrimination, create room for manipulation, or make people blindly trust algorithm-based decision-making [Drunen M. Z., Helberger N., Bastian M., 2019: 220–235]. Price discrimination can be identified in addition to gender discrimination, which creates inequality among different segments of the population [Veale M., Edwards L., 2018: 401–402].

Transparency allows to remove information asymmetry between all actors. As a result of the use of automated systems, an information asymmetry may develop, first of all between a state agency (the system operator) and a citizen (the subject of the decision), where the advantage of one person arises precisely owing to information about the other person (including information against the other person). Information asymmetry can be used both to the advantage and to the disadvantage.

Explainability and transparency ensure the decision-making procedure is legitimate [Fine Licht de K., Fine Licht de J., 2020: 918–926]. Moreover, an automated decision, made in a way that is explainable and procedurally fair, helps to ensure that the decision is legitimate and that the decision-making body has credibility among citizens.

Explanation and transparency may be helpful to the applicant by helping to understand which inputs had the strongest influence on the decision made [Verma S., Boonsanong V. et al., 2022: 2]. In addition, these requirements allow an applicant to challenge a decision, for example, if their race was critical in determining the outcome. This may also be useful for organisations when testing their algorithms for systematic biases.

In some cases, explanation and transparency provide the applicant with feedback on the basis of which they can take action to get the desired outcome in the future.

Explanation helps to adhere to laws related to machine decisions, including Regulation No 2016/679 of the European Parliament and of the Council of the European Union "On the protection of natural persons with regard to the processing of personal data and on the free circulation of such data and repealing Directive 95/46/EC (General Data Protection Regulation)" (hereinafter GDPR).

At the same time, there are also opposing views arguing the requirements of explainability and transparency are unnecessary, especially in the context of public administration. The arguments proposed are that the pace of technology development, multiple transparency concepts, uncertainty about where transparency is required, how best to approach communication with different stakeholders, and how to build transparency measures into meaningful and organisationally realistic accountability measures all pose challenges to implementing these requirements, despite seemingly general agreement this is important [Felzmann H., Fosch-Villaronga E., Lutz C. et al., 2020: 3355]. These chal-

lenges may also include the risks of disclosure of algorithm developers' trade secrets, the possibility of system manipulation by knowledgeable actors ("gaming" the system), and the significant costs of developing and implementing truly effective explainability mechanisms for complex AI systems. Furthermore, there is a concern that excessive transparency requirements may slow down the adoption of innovative technologies in public administration.

2. Legal Mechanisms to Maintain Transparency and Explainability of Automated Decision-making and AI Systems

2.1. Mechanism Categories: *ex ante* Approach and *ex post* Approach

Contemporary and proposed mechanisms aimed at maintaining transparency and explainability of automated decision-making systems and AI systems in public administration may be categorised on various grounds. Legal doctrine and related fields of knowledge offer various grounds for categorising such mechanisms[3].

Categorisation by the goal of transparency and explainability. Under this approach, items that fall under the requirement of transparency and explainability, can be grouped by the two main aspects:

Transparency and explainability of the decision-making process (algorithm) implies disclosure of information about the system itself, its architecture, logic of functioning and data used (e.g., what factors the system takes into account and how when making decisions);

Transparency and explainability of the outcome (a particular decision): focuses on providing information that justifies a specific decision made by the system in relation to a particular actor or situation (e.g., why a particular decision was made in this case and what data of the actor influenced it).

Categorisation by the timing of the explanation and the nature of the transparency. This approach differentiates mechanisms depending on

[3] See: Explaining decisions made with AI. 2022. Available at: URL: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/explaining-decisions-made-with-artificial-intelligence/ (accessed: 07 April 2025)

the stage in the life cycle of automated decision making and AI systems in which they are implemented, and subdivides transparency into:

ex ante mechanisms. These mechanisms are implemented before automated decisions are made, and independently of a particular decision. Their purpose is to prevent risks, ensure the predictability of the system's operation, and inform the public and stakeholders about the principles of its operation and potential consequences;

ex post mechanisms. These mechanisms are applied after the automated decision has been made, especially if it affects the rights and legitimate interests of the subjects. Their purpose is to ensure accountability, enable effective appeal, correct errors and analyse the performance of the system for future improvement.

Categorisation by the levels and types of transparency. The following interrelated levels and types of transparency can be identified depending on the item of information disclosure:

data transparency: disclosure of information about the data used. This aspect is critical because the quality and characteristics of the data directly affect the functioning and performance of the system and the AI.

algorithm transparency: disclosure of information about the algorithm itself. The purpose is to maintain understanding of how the system processes information and arrives at conclusions. In some cases, this may involve disclosure of the source code or model of the AI, although this carries risks to intellectual property, various secrets, and information security (e.g., identifying system vulnerabilities);

results transparency: the ability of a system or its associated mechanisms to explain in ways understandable to a human why a particular decision was made and how certain inputs led to a particular conclusion.

These approaches to classification emphasise the multidimensionality of the concepts of "explainability" and "transparency" in relation to the automated decision-making system. At the same time, different types of transparency and explainability mechanisms are not mutually exclusive. On the contrary, they should complement each other, forming a comprehensive system at all stages of the system's life cycle in public administration.

Our analysis of foreign academic literature, laws and law enforcement practices allows us to identify a number of basic legal, institutional and technical mechanisms aimed at ensuring the transparency and explain-

ability of the system. We believe in the beginning it would be expedient to group them according to one of the key classifications, namely the timing of the explanation (*ex ante* and *ex post*):

2.2. Mechanisms of Preventive Control (*ex ante*)

Ex ante mechanisms create conditions for inherent predictability, controllability and legitimacy of automated decision making.

Disclosure of the use of an automated decision-making system. Obligation to inform actors that a decision has been made using the above system. This is a basic requirement related to the right to information and necessary for the realisation of other rights (request for information; call for human intervention: right to appeal a decision made using an automated decision-making system);

Disclosure of the order or logic of decision-making (under personal data laws). Personal data law (e.g., the general requirements for informing the person contained in Federal Law No. 152-FZ "On Personal Data" of 27.07.2006), contains rules requiring operators to explain how automated decisions are made or to provide "meaningful information about the logic involved", although the level of detail isn't as significant as in Articles 13-15 of Regulation (EU) 2016/679 of the European Parliament and of the Council of 27.04.2016 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) in the EU. However, this mechanism is limited in scope only to decisions based solely on automated processing of personal data with legal consequences. Other limitations include aspects of protection of IP and trade secrets, and the difficulty of explaining the logic of complicated models to non-specialists. Moreover, even if such a right does exist, its implementation may be hampered by the lack of clear criteria for the "meaningfulness" of information about the logic and about the limits on the disclosure of such information so as not to infringe the rights of developers. Another open question is the efficiency of such disclosure for complex self-learning AI systems because their logic is not always deterministic and is able to evolve over time;

Disclosure of information about the data used for the development and operation of the automated decision-making system. Provision of information about the sources, types, and characteristics of data on which the system has been trained and operates. This allows the poten-

tial impact of the system to be assessed, the impact of the system to be investigated, and biases to be identified. This also includes disclosure of data in the form of open data sets (with due observation of confidentiality), which facilitates public scrutiny and encourages innovation;

Disclosure of the programme code and (or) AI model. Providing access to the source code or detailed description of the model. This mechanism allows for the most in-depth public scrutiny. On the other hand, it faces serious constraints related to the protection of intellectual property and trade secrets. International practice offers various examples in this respect.

Pre-audit, certification, and impact assessment. Independent checks of automated decision-making systems prior to implementation for their compliance with the law, ethical standards, and to identify risks. These may range from government oversight mechanisms to voluntary certification or internal audit systems. Such internal audit may assess the suitability of the system for its stated purposes, the quality and representativeness of the data used for training, the existence of discrimination prevention mechanisms, the reliability and security of the system, and the adequacy of measures to ensure transparency and explainability. Another promising field is developing standardised methodologies for conducting such assessments, including criteria for assessing data and algorithm biases, and accrediting independent auditors with relevant competencies in both legal and technical areas.

2.3. Legal Mechanisms of Subsequent Control (*ex post*)

The aim of this mechanisms is to maintain basis of a decision already taken is understood and may be challenged.

"Right to an explanation" of an individual decision. An evolving concept involving the legislated ability of a person affected by an automated decision to receive comprehensible explanations of the system's role in a particular decision and its underlying determinants. An example of enshrining such a right is Article 86 of Regulation 2024/1689 of the European Parliament and of the Council of 13.06.2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). This right implies not only stating that an automated decision-making system was

used, but providing the user with personalise information about the factors that influenced a particular decision and, if possible, the logic that guided the system. However, implementation of this right directly depends on the technical ability to provide such an explanation in a form understandable to the human user, in particular in the case of complex AI systems;

Counterfactual explanations. Providing information about what changes in the inputs or conditions could have led to a different (e.g., desired) outcome. This approach helps in understanding the logic of the system and its sensitivity to various factors, and offers practical pieces of advice to the user. Counterfacts answer the question "what if" and can reveal hidden biases. On the other hand, in implementing this approach, one is faced with the multiple possible explanations problem ("the Rashomon effect") and the difficulty to take into account all the relevant factors. In addition, generating counterfactual explanations can be computationally expensive and resource intensive. It is also important to bear in mind that while such explanations can be useful for understanding the system's sensitivity to changes in the inputs, they sometimes fail to show the real cause for the decision made; instead of it they show how a different outcome could have been achieved. That said, they have a significant potential in enhancing the understanding and extending the field of the user's opportunities to interact with the system;

Disclosure of the data that served as the basis for a particular automated decision. Ensuring that user has access to specific data that the automated decision-making systems used to make a decision about him or her. This allows to check the data for correctness and completeness, identify irrelevant or discriminatory factors, exercise the right to correct the data; furthermore, it is the basis for a reasoned challenge to the decision.

3. The Role of Organizational and Technical Solutions in the Legal Support of Transparency

On the other hand, the purpose of legal mechanisms largely depends on the existence of adequate organisational and technical tools for implementing them. The rules of law that enshrine principles and duties require adequate technical tools for putting them into practice. Without proper technology solutions, many legal requirements, such as the right to explanation or the obligation to disclose the system's logic, may

remain declarations. That enhances the role of technological methods that can either make the systems inherently more understandable or provide tools for *ex post factum* analysis of how they work.

3.1. *Ex ante* Organisational and Technical Approaches: Interpretable Models and "Transparency by Design"

The key strategy is to create and use systems designed with interpretability or explainability features. That includes:

artificial intelligence models that are interpretable and explainable "by default." Initially interpretable or explainable models thus directly promote the implementation of *ex ante* legal mechanisms. For instance, the use of such models facilitates due diligence audit and certification, because their logic is more open to analysis. Besides, it facilitates disclosure of information about their decision-making procedure or logic and about the data used to develop the system. Legislative codification of requirements or recommendations to use such models, especially for automated high-risk decision-making in public administration, could be an important step towards building transparent automated decision-making systems. That may be implemented via standards, guidelines for developers and state clients, and also via assessment criteria used in the procurement of artificial intelligence systems for public needs;

forming publicly accessible registers of the automatic decision-making systems used in public administration, indicating their purpose, applications (specific state functions or services), type of the data used (including the availability and sources of personal data), degree of automation (decision-making support system or fully or partially automated decision), developer and operator information, information on conformity assessment or audit passed (where applicable), and contact information for requesting explanations or appealing against decisions. Such registers should be easily accessible to citizens and be updated regularly. Keeping them could be entrusted to a special authority or integrated into existing State service and open data portals;

delegation of specific powers to an existing or newly established public authority to supervise automated decision-making systems in the public sector. Such powers might include: keeping the above-mentioned register, development of transparency and explainability guidelines and standards, holding scheduled and extraordinary checks (audits), issuing orders to correct any irregularities, and initiating studies to assess the

risks and the automated decision-making systems' impact on individuals' rights. Given the specifics of operating within the public administration system, it is important to make sure that such an authority is independent, impartial, and possesses the required expertise and technical resources. A potential mechanism that could strengthen confidence in the findings is the adoption of procedures that keep the audit findings unchanged and truthful using e.g. distributed registry technology or other cryptographic methods to record the findings, and in some cases expressly defined by law, for records on formal aspects of the audit, notarisation;

Adherence to Privacy/Transparency by Design approaches. As noted by L. Edwards and M. Veale, the newly passed GDPR introduces a number of new provisions that attempt to create an environment in which less "toxic" automated systems will be built in future. These ideas come out of the long evolution of Privacy by Design engineering as a way to build privacy-aware or privacy-friendly systems, generally in a voluntary rather than mandated way. [Edwards L., Veale M., 2018: 46–54]. While, historically, Privacy by Design focused on privacy, its principles (proactivity, integration in design, and focus on the user) are also applicable to the pursuit of transparency and explainability in a broader sense as they lay a basis for Transparency by Design.

Besides, the above concept should extend into a principle of heightened requirements to models for high-stakes decisions. Thus, one study states that the legislator should call for greater efforts to ensure the safety of, and confidence in, machine learning models that support high-stake and highly significant decisions [Rudin C., 2019: 206–215]. This principle leads developers and customers to choose or create more reliable and, potentially, more transparent models at the *ex ante* stage for critical automated decision-making systems in public governance.

3.2. *Ex post* Organisational and Technical Approaches: Explainable Artificial Intelligence Tools for Decision Analysis

The analysis of decisions already taken by systems (especially, by "black boxes") employs methods of an explainable AI system (Explainable AI, XAI):

Explainable artificial intelligence (XAI) methods. A set of techniques that help generate explanations for individual decisions that suit the specific case and the user's level of understanding (e.g., LIME — Local

Interpretable Model-agnostic Explanations; SHAP — SHapley Additive exPlanations);

Interactive visualisation and What If analysis tools. These enable both users and experts to examine the model's behaviour and understand how different input data will affect the result, which is closely related to counterfactual explanations;

Intelligent decision assistance. Automated decision-making was shown to have many benefits for both business and society, but that comes at a cost. It has long been known highly automated decision-making may have various drawbacks such as biased decisions and loss of professional skills by employees. Authors have analysed those two disadvantages to develop a new decision support system, namely Intelligent Decision Assistance [Schemmer M., Kühl N. et al. 2021: 1–10]. That system complements the human decision-making process with explainable AI, while offering no concrete recommendations. Such an approach may be used *ex post*, so that the human reviewer can understand AI contribution to the decision taken and assess it for relevance, which is important for human supervision and challenge mechanisms;

Establishing a procedure for challenging automated decisions. Development of an administrative and judicial procedure for appealing against decisions that were taken using automated decision-making systems, including the definition of the standard of proof and burden of proof distribution. Human control must remain in place and permit revision of an automated decision. Thus, whatever the automation level may be, there should remain an opportunity to appeal to a human and have the decision revised. The procedure should also take into account the specifics of automated decision-making systems, e.g. permit requesting the system's technical logs (subject to any limitations on access to legally protected secrets) and engaging artificial intelligence experts to analyse whether the system is functioning correctly.

Those tools form an institutional and technical basis for exercising the right to explanation and can be used for system audit by individuals and supervisory authorities.

3.3. Integration of Organisational and Technical Solutions into Legal Regulation: the Need and Prospects

Automated decision-making and artificial intelligence systems cannot efficiently be made transparent and explainable unless the legal rules

are closely integrated with the development, implementation and use of the relevant technical standards, tools, and methods. That is because legal regulation should not just proclaim duties and principles, but also create efficient mechanisms for putting them into practice by stimulating technological development and channelling it towards the observance of human rights and good governance.

Firstly, assessment of how the above legal mechanisms are codified in the light of the state policy is one of the key modalities. The authors of the current paper believe that such mechanisms should accompany every stage of an automatic decision-making system's lifecycle. As an additional reference point, we can consider developing criteria and clear recommendations for the developers of those systems that could help create reliable systems with an emphasis on the protection of the state's core values and the rights of individuals.

That may be achieved particularly by establishing:

minimum requirements on the interpretability of artificial intelligence models depending on the degree of risk and the significance of the decisions taken (e.g. mandatory use of verifiable and explainable models for high-risk systems);

formats and protocols for giving explanations that make them understandable to various categories of users (laypersons, officials and/or experts);

standards and requirements on data quality that guarantee the reliability of that fundamental element of artificial intelligence by providing accurate, up-to-date, representative and complete data that will underlie an automated decision;

requirements on logging the automated decision-making system's actions, which is critical for conducting audit, investigating incidents and providing evidence in a decision challenge process. The said logs must contain information about the input data, key information processing stages, and the resultant decision indicating time.

Secondly, law should encourage and regulate the use of specific technical tools and techniques that enhance transparency. This includes:

supporting the development and implementation of explainable artificial intelligence (XAI) tools such as LIME, SHAP or analogues, adapted for use in state information systems. The state could either commission such developments or facilitate their advent into the market;

creation and support of platforms for testing and verifying an automated decision-making system for compliance with transparency and non-discrimination requirements and with other ethical and legal rules. Such sandboxes could be used by developers as well as supervisory authorities;

development of methods for assessing the automatic decision-making systems' effect on human rights (Human Rights Impact Assessment), to include technical aspects of system analysis and the assessment of potential social consequences of their adoption.

Thirdly, legal conditions should be created to support efficient use of technically generated explanations and data in legal procedures. Law should establish requirements on the quality, completeness and understandability of technically generated explanations so that individuals can use them to protect their rights, and courts and administrative authorities can use them to assess decisions for lawfulness. The legal status and evidential force of information obtained from an automated decision-making system (such as logs and explanations) should be defined. This should include development and codification of procedures for requesting, receiving and challenging such explanations that will guarantee prompt provision of understandable information and easy access to the procedure itself.

Fourthly, it is important to develop interdisciplinary co-operation. The transparency of an automated decision-making system can only be successfully and efficiently achieved through deep integration of legal, organisational and technical solutions. Close co-operation and interaction among lawyers, AI developers, researchers, ethicists, and members of civil society is thus required. Here we should assume the very implementation of automated public administration is impossible without a better "digital literacy" and understanding of the work of AI by two groups: on the one hand, by public officials, judges and other law enforcers. On the other hand, by citizens who are both recipients of such decisions and the principal actors, and are thus expected to know and understand their own rights and duties, including procedure for challenging an automated decision.

Consequently, amid rapid evolution of the governance paradigm, any well-developed legal order aiming to protect human rights and interests as the supreme value should include, as justified and necessary actions, active studies of the world's best practices (including the approaches embedded in the EU AI Act) and encouraging domestic research and

practical developments in the field of explainable and trusted artificial intelligence. Such an interdisciplinary and international approach will support the adoption of advanced technology subject to the basic principles of a state governed by the rule of law, where transparency is central to the government's accountability to society.

Conclusion

One of the key objectives of current law and order is to maintain the transparency and explainability of automated decision-making and use of artificial intelligence systems in public administration.

Analysis has shown that, despite active development of law and doctrine in the field, the existing legal mechanisms — both preventive (*ex ante*) ones and those providing for posterior (*ex post*) control — are fraught with certain limitations and cannot always and fully protect citizens' rights and keep the authorities accountable amid algorithm-based governance.

None of the mechanisms discussed is a universal solution; efficiency can only be achieved through their comprehensive application and adaptation to the specific context around the use of automatic decision-making and artificial intelligence systems, with due regard to the risk level associated with the decisions in question, their social significance and the technical complexity of the systems being used. Most importantly, legal requirements must be closely integrated with the development and implementation of relevant organisational and technical solutions that can ensure real, not declarative, transparency and explainability.

Development of legislation and jurisprudence should aim to specify the obligations of developers and operators of automated decision-making systems, establish clear-cut criteria for assessing the adequacy of the explanations returned, and to strike the optimal balance between the needs for openness, protection of intellectual property and trade secrets, and information security.

A deep-rooted and conscious culture of transparency and responsibility should become an important feature, both in public authorities and among developers and operators of artificial intelligence systems. Only in this way can we ensure that the adoption of advanced information technology really fosters safer and more equitable, efficient and accountable governance that meets both society's and every citizen's interests.

References

1. De Fine Licht K., De Fine Licht J. (2020) Artificial Intelligence, Transparency, and Public Decision-Making: Why Explanations are Key When Trying to Produce Perceived Legitimacy. *AI & Society*, no. 35, pp. 917–926. doi: https://doi. org/10.1007/s00146-020-00960-w

2. Drunen M.Z., Helberger N., Bastian M. (2019) Know Your Algorithm: What Media Organizations Need to Explain to their Users about News Personalization. *International Data Privacy Law, v*ol. 9, no. 4, pp. 220–235. doi: https://doi.org/10.1093/idpl/ipz011.

3. Edwards L., Veale M. (2018) Enslaving the Algorithm: From a Right to an Explanation to a Right to Better Decisions? *IEEE Security & Privacy*. no 3, pp. 46–54. doi: https://doi.org/10.1109/MSP.2018.2701152.

4. Felzmann H., Fosch-Villaronga E., Lutz C. et al. (2020) Towards Transparency by Design for Artificial Intelligence. *Science and Engineering Ethics,* no. 6, pp. 3333–3361. doi: https://doi.org/10.1007/s11948-020-00276-4.

5. Grimmelikhuijsen S. (2023) Explaining Why the Computer Says No: Algorithmic Transparency Affects the Perceived Trustworthiness of Automated Decision-Making. *Public Administration Review,* no. 2, pp. 241–262. doi: https://doi.org/10.1111/puar.13483.

6. Kutafin O.E. (2008) The Russian Constitutionalism. Textbook. Moscow*:* Norma, 544 p. (in Russ.)

7. Lee M.K. (2018) Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic Management. *Big Data & Society,* vol. 5, no. 1, pp. 1–16. doi: https://doi.org/10.1177/2053951718756684.

8. Narutto S.V., Nikitina A.V. (2022) Constitutional Principle of Trust in Modern Russian Society. *Konstitucionnoe i municipalnoe pravo*= Constitutional and Municipal Law, no. 7, pp. 13–18 (in Russ.)

9. Pogodina I.V. (2023) Forming Culture of Transparency with Help of ICTs. *Gosudarstvennaya vlast i mestnoe samoupravlenie*=State Power and Local Self-Government, no. 11, pp. 29–31(in Russ.)

10. Rudin C. (2019) Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. *Nature Machine Intelligence,* no. 5, pp. 206–215.

11. Schemmer M., Kühl N. et al. (2021) Intelligent Decision Assistance versus Automated Decision-Making: Enhancing Knowledge Work through Explainable Artificial Intelligence, pp. 1–10. doi: https://doi.org/10.48550/ARXIV.2109.13827.

12. Schiff D.S., Schiff K.J., Pierson P. (2022) Assessing Public Value Failure in Government Adoption of Artificial Intelligence. *Public Administration,* vol. 100, no. 3, pp. 653–673. doi: https://doi.org/10.1111/padm.12742.

13. Silkin V.V. (2021) Transparency of the Executive Power in the Digital Age. *Rossijskij yuridicheskij zhurnal*=Russian Law Journal, no. 4, pp. 20–31 (in Russ.)

14. Sokol K., Flach P.A. (2019) Counterfactual Explanations of Machine Learning Predictions: Opportunities and Challenges for AI Safety. *Safe AI AAAI*, pp. 1–4.

15. Srinivasu P.N., Sandhya N. et al. (2020) From Black Box to Explainable AI in Healthcare: Existing Tools and Case Studies. doi: https://doi.org/10.1155/2022/8167821.

16. Veale M., Edwards L. (2018) Clarity, Surprises, and Further Questions in the Article 29 of Working Party Draft G=Guidance on Automated Decision-Making and Profiling. *Computer Law & Security Review*, no. 2, pp. 398–404. doi: https://doi.org/10.1016/j.clsr.2017.12.002.

17. Verma S., Boonsanong V. et al. (2022) Counterfactual Explanations and Algorithmic Recourses for Machine Learning: A Review. Counterfactual Explanations and Algorithmic Recourses for Machine Learning. arXiv:2010.10596 [cs, stat]. arXiv, pp. 1–23.

18. Wachter S., Mittelstadt B., Floridi L. (2017) Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. *International Data Privacy Law, v*ol. 7, no. 2, pp. 76–99. doi: https://doi.org/10.1093/idpl/ipx005.

19. Wischmeyer T., Rademacher T. (2020) Artificial Intelligence and Transparency: Opening the Black Box. Regulating Artificial Intelligence. Cham: Springer International Publishing, pp. 75–101.

Information about the authors:

P. P. Kabytov – Candidate of Sciences (Law), Leading Researcher.

N. A. Nazarov – Junior Researcher.

Contribution of the authors:

P.P. Kabytov–Introduction, Conclusion; N.A. Nazarov–Chapters 1,2,3; Introduction, Conclusion.

The article was submitted to editorial office 06.03.2025; approved after reviewing 21.04.2025; accepted for publication 12.05.2025.

Legal Issues in the Digital Age. 2025. Vol. 6, no. 2.
Вопросы права в цифровую эпоху. 2025. Том 6. №2.

Research article
JEL: K23, K 38
UDK: 342.5, 342.7
DOI:10.17323/2713-2749.2025.2.183.212

The Artificial Intelligence Influence on Structure of Power: Long-Term Transformation

Vladimir A. Nizov

Sber, 19 Vavilova St., Moscow, Russia 117312,

vnizov12@gmail.com, ORCID: 0000-0002-0933-8775, Science Index: 5176-7786

Abstract

Integration of artificial intelligence (AI) into public administration marks a pivotal shift in the structure of political power, transcending mere automation to catalyze a long-term transformation of governance itself. The author argues AI's deployment disrupts the classical foundations of liberal democratic constitutionalism — particularly the separation of powers, parliamentary sovereignty, and representative democracy — by enabling the emergence of algorithmic authority (algocracy), where decision-making is centralized in opaque, technocratic systems. Drawing on political theory, comparative case studies, and interdisciplinary analysis, the researcher traces how AI reconfigures power dynamics through three interconnected processes: the erosion of transparency and accountability due to algorithmic opacity; the marginalization of legislative bodies as expertise and data-driven rationality dominate policymaking; and the ideological divergence in AI governance, reflecting competing visions of legitimacy and social order. The article highlights AI's influence extends beyond technical efficiency, fundamentally altering the balance of interests among social groups and institutions. While algorithmic governance promises procedural fairness and optimized resource allocation, it risks entrenching epistocratic rule — where authority is concentrated in knowledge elites or autonomous systems — thereby undermining democratic participation. Empirical examples like AI-driven predictive policing and legislative drafting tools, illustrate how power consolidates in executive agencies and technocratic networks, bypassing traditional checks and balances. The study examines paradox of trust in AI systems: while citizens in authoritarian regimes exhibit high acceptance of algorithmic governance, democra-

cies grapple with legitimacy crises as public oversight diminishes. The author contends "new structure of power" will hinge on reconciling AI's transformative potential with safeguards for human dignity, pluralism, and constitutionalism. It proposes a reimagined framework for governance — one that decentralizes authority along thematic expertise rather than institutional branches, while embedding ethical accountability into algorithmic design. The long-term implications demand interdisciplinary collaboration, adaptive legal frameworks, and a redefinition of democratic legitimacy in an era where power is increasingly exercised by code rather than by humans.

Keywords

artificial intelligence; separation of power; structure of power; algocracy; epistocracy; liberal democracy.

For citation: Nizov V.A. (2025) The Artificial Intelligence Influence on Structure of Power: Long-Term Transformation. *Legal Issues in the Digital Age,* vol. 6, no. 2, pp. 183–212. DOI:10.17323/ 2713-2749.2025.2.183.212

Introduction

It is difficult to find a developed country that does not recognize the vital importance of implementing artificial intelligence (AI) in public administration. Of course, the question of how to define AI remains subject to debate; however, the overall trend toward its integration into governance is robust and sustainable. In electronic government, the role of AI has become more significant than it was previously. The reason for this shift is straightforward: AI can perform certain tasks in ways that surpass human capabilities. As a result, public administration can become faster, less expensive, and more efficient one through the implementation of AI technologies. The countries who will avoid the implementation of the AI in the public administration may become degenerative exceptions due to the fact of the international rivals.

In modern history, governments have continuously sought tools to automate basic human functions. Initially, the primary goal was the development of military technologies. Beyond defense, computers have been employed for decades by government agencies to support administrative and data management tasks, including tax collection and the operation of large national benefit programs [Relyea H., Hogue H., 2004: 16].

Today, the implementation of new governance systems based on AI can be either fully automated or semi-automated [Danaher J., 2016: 247]. Removing the human element introduces both structural advan-

tages and disadvantages. This new era of decision-making without human intervention requires thorough and foundational analysis.

The potential for rapid advancements in AI technology has prompted widespread concern, including calls for government regulation of AI development and restrictions on its deployment. Such concerns are not unprecedented — fear of technological change and demands for governmental oversight have accompanied nearly every major technological innovation.

Therefore, it is crucial to understand the legal, political, and ethical obstacles societies face in the full implementation of AI in governance and public administration. Public decision-making typically requires moral and political legitimacy [Peter F., 2017]. Scholars have identified different approaches to understanding AI: the technical approach, which studies algorithms as computational tools; the sociological approach, which examines algorithms as products of interactions among programmers and designers; the legal approach, which considers algorithms as entities within legal frameworks; and the philosophical approach, which explores the ethics of algorithmic decision-making [Barocas S., Hood S., Ziewitz M., 2013: 3].

The hypothesis of the research is implementation of AI in public administration leads to a transformation of the classical structure of state power. Implementation of AI usually necessitates reconfiguring existing processes, and the current power structures are no exception. The present model of political decision-making is increasingly misaligned with the development of AI. Society must either slow the pace of AI development or reform the existing governance system to better accommodate these changes. The author argues the most pressing challenges lie not primarily in legal or technical domains, but in philosophical and ethical considerations. These emerging issues may ultimately challenge classical political philosophy and contemporary legal systems.

The author focuses on the heart of liberal democratic constitutionalism such as separation of powers and representative democracy. These principles have historically ensured checks and balances within state institutions and safeguarded citizens from arbitrary governance. However, the deployment of AI challenges these foundational elements by introducing new forms of authority — often opaque, technocratic, and centralized ones — that do not easily align with democratic frameworks.

The author also explores how the integration of AI into public administration disrupts the classical structure of state power and poses significant risks to liberal democratic constitutionalism. It investigates

whether algorithmic governance can maintain democratic legitimacy, especially when decision-making becomes less transparent and more reliant on epistemic elites or autonomous systems. Furthermore, it examines how AI may erode parliamentary sovereignty.

Ultimately, the article seeks to answer the central research question: How does the integration of AI into public administration challenge the foundational principle of liberal democratic constitutionalism — separation of powers? In doing so, it calls for a rethinking of governance structures that can accommodate technological advancements without compromising democratic ideals.

This article consists of five sections, including the introduction. The first chapter outlines the core functions of public administration and proposes a classification relevant to the current research. The second chapter examines the legal, political, and ethical challenges associated with replacing human decision-makers with AI. The third one presents conceptual proposals for the long-term integration of AI in governance. Finally, the conclusion summarizes key findings and discusses implications for future research and policy.

1. Use AI in Public Management

To understand how AI transforms public administration researchers must examine two interrelated dimensions: the nature of AI technologies and their impact on current and future social processes; and the evolving concept and structure of power. This chapter focuses on the AI's role in public management — with the latter being explored in detail in the subsequent chapter.

The author does not attempt to offer a definitive definition of AI applicable across all domains of public management. Indeed, no universally accepted definition of AI is available, even among experts in the field. Citing Alan Turing's foundational work, highlight an approach that emphasizes AI's capacity to «act humanly» — a perspective rooted in early conceptions of machine behavior [Turing A., 1950: 442]. However, what distinguishes AI from earlier technologies is its ability to operate autonomously. Already, AI systems can perform complex tasks such as driving vehicles or managing investment portfolios without direct human supervision.

For the purposes of the study, the definition proposed by the High-Level Expert Group on Artificial Intelligence serves as a comprehensive

framework: "software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal"[1]. Despite its complexity, this definition captures the breadth of AI applications and provides conceptual coherence for the analysis.

However, it is quite important to understand that the implementation of AI in public management represents a form of algorithmization — a process wherein decision-making and administrative functions are increasingly governed by algorithms. As Kushner notes, algorithms do not merely perform tasks but also construct and implement regimes of power and knowledge [Kushner S., 2013: 1243–1244]. Their deployment carries normative implications [Anderson C., 2011: 530], shaping how authority is distributed, exercised, and perceived. The system where algorithms make decisions and (or) implement decisions has a different name in the literature: algorithmic authority [Shirky C., 2009] or algorithmic governance [Musiani F., 2013: 3]. More pragmatic term we find in Dodge and Kitchin "automated management". They describe this term as decision-making processes that are automated, automatic and autonomous; outside of human oversight [Dodge M., Kitchin R., 2007: 270].

While algorithmization is not a novel phenomenon: examples exist even in ancient administrative systems [Miyazaki S., 2012: 1–3], but the pace and depth of change driven by AI are unprecedented. Unlike traditional automation, which follows predefined rules, modern AI systems can learn and adapt, potentially expanding the scope of tasks they can perform.

From a technical standpoint there is no inherent distinction between algorithmizing private sector operations and public administration. However, the political significance lies in identifying which state functions are deemed essential and how they should be classified. First, the present chapter will give the general understanding of the public administration from the AI implementation perspective. Second, the examples of the AI projects in public administration will be given. Third, the chapter present the brief classification of the public administration process.

[1] The European Commission. Ethics Guidelines for a Trustworthy AI. Brussels, 2020, p. 36.

From the cybernetic perspective the algorithmizing of the processes might be possible without the informational technology. The informational technologies' functions were the prerogative of the humans. The humans did the simple tasks, such as delivering letter, collecting the papers, etc. The effectiveness of the public administration was and still depended on these simple tasks. AI goes further and tries to implement even more complicated tasks. However, AI is limited by the possible options, which were programmed for it. Self-educated systems may enlarge the possible options for the activity, but the origin of the code establishes the red line for such activity.

The implementation of AI may influence on political system and foster tremendous social changes. It's obviously not the first time that a techno-scientific field's promise to bring about utopia (or dystopia) has been exploited. Given the behaviorist core of today's celebrated AI systems, it's worth revisiting the 20th century debates on behaviorism-based visions of a future society. In a critique of B.F. Skinner's promises that human behavior can be reshaped to produce a desirable society using the scholar methods of reinforcement, Noam Chomsky wrote: "One waits in vain for psychologists to make clear to the general public the actual limits of what is known. Given the prestige of science and technology, this is a most unfortunate situation" [Chomsky N., 2010].

Slavery, feudalism, capitalism and socialism are the systems that gave the answer for the main question: how society must be organized. The main feature and precondition in these systems is the status of the different people within society. The situation in legal, economic, and political spheres predetermined the answer for the general question. There is no doubt that AI influences on all three spheres. That is why the society needs to find appropriate model for the future governing. The phenomenon of algorithmic governance is a part of a long historical process and since the time of Max Weber, the approach to the legal-bureaucratic organization of the state is subject to the same modernizing trends as the design of industrial factories. The continuation of this trend we may find, for example, in New Public management. The speed, scale and ubiquity of the modern technologies that make algorithmic governance possible are grander [Danaher J., Michael J., Hogan M., 2017: 2, 7] and may change the classic structure and the essence of the public administration (see below).

For the understanding of the AI implementation in public administration, it has a sense to use Kitchin methods. He argues a major goal

of algorithm studies is to find answer for the question: how algorithmic governance systems are designed and implemented [Kitchin R., 2017: 16–17].

In spite of the fact that the research is inclined to give picture for the state system, it is impossible to avoid the steps of the transformation of the public administration with AI implementation. Author will use Coglianese and Ben Dor classification of the "spectrum of digital technologies". They provide three main point of the spectrum: digitization, algorithmic tools and machine learning. The closest step begins with simple *digitization*. This step is a building bridge to the possibility of the AI implementation because it can facilitate the availability of the "Big Data" on which machine learning is based. Next point is *algorithmic tools* that is, traditional, human-created statistical models, indices, or scoring systems that are then used as decision tools. Only the final step called a *machine learning* constitutes what we will consider AI, because learning algorithms essentially work "on their own" to process data and discover optimal mathematical relationships between them [Coglianese C., Ben Dor L., 2021: 795–796].

Thus, the AI is possible only in some situations of the public administration where the machine learning is possible for modern technologies and provide effective results compare with human activities. However, the new technology tools sit closer to the decision-making point, and thus entail greater displacement of human discretion, than past rounds of innovation [Coglianese C., Lehr D., 2019: 23]. The observed trend leads us to the conclusion that fully automated decision-making, leaving progressively less to human discretion and analysis system is possible in future [Ho D., Engstrom D., 2021: 59]. Some researches even dream of creating 'master algorithms' that will be able to learn and adapt to any decision-making situation without the need for human input or control [Domingos P., 2015: 23–56].

Despite variations in political regimes, AI technologies are largely standardized across the world. Differences arise primarily in how governments choose to apply them. The main difference is the aims and focuses in utilizing AI. Smart cities all around the world use surveillance technologies, such as facial recognition and cloud computing for ordinary policing. However, smart cities in China have bigger focus on these technologies [Roberts H. et al., 2021: 67]. In contrast, the European Union has taken a more cautious stance, prioritizing privacy and human rights — evident in its regulatory frameworks such as the General Data

Protection Regulation and the AI Act. The difference of the Chinese and European approach is not only in the focuses, but in deepness of the implementation of the AI technology. Europeans try to avoid direct implementation of AI in the public administration and governance, and China try to change social construction with a Social Credit System, where AI will play a central role [Ding J., 2018: 34].

The analysis of the governance system is a complicated task indeed, and it is necessary to employ two methodologies: analysis of the concrete functions of government and the analysis of the management process. In the first method may help to distinguish vital functions of the government and functions which are not necessary to exercise by government, the second method may foster the understanding of where it is possible to implement AI and where it is not. To sum up, the analysis needs to provide broader picture of the governance system: even the most essential function can be separated on many simple tasks. The answer for the analysis will based on the understanding where the modern social system of governance has "sensitive points" for the AI implementation.

The difference in automation of the concrete functions can be shown on robotic weapon systems. Citron and Pasquale proposed the next classification of robotic weapon systems [Citron D., Pasquale F., 2014: 6–7]:

Human-in-the-loop weapons: Robots can only select targets and deliver force with a human command.

Human-on-the-loop weapons: Robots can select targets and deliver force on their own, but there is human oversight and the possibility of human override.

Human-out-of-the-loop weapons: Robots act autonomously, selecting targets and delivering force without human oversight or override.

The classification of three elements ("human-in-the-loop", "human-on-the-loop" and "human-out-of-the-loop") can be universal for any public function. For example, Danaher use this classification for tax law enforcement systems [Danaher J., 2016: 248].

In the theory of the public administration, it is possible to find two main parts of the administration process: decision-making process and process of action. Additionally, four-step decision model that incorporates intelligence, design, choice and review can be appropriate further classification. It is a simplifying classification, but it is needed for structural analysis of all process. The scholars who investigate algorithm governance use the next classification: collection, processing, uti-

lization and feedback and learning [Zarsky T., 2013: 1504]; [Citron D., Pasquale F., 2014: 27–29].

To summarize the analysis of the decision-making process and the process of the implementation it is possible to state that automation and algorithmizing are possible on all stages. However, it much more important to understand concrete function: lawmaking and automatic boarder control may have the same stages, but the possibility of replacing human is different. Agencies have limited number of auditors, inspectors, and other enforcement personnel who must oversee a vast number of individuals and businesses to ensure their compliance with myriad pages of laws and regulations [Ho D., Engstrom D., 2021: 70]. Machine-learning algorithms can provide forecasts of the likelihood of violations, thus helping agencies allocate resources and decide which regulated entities to target [Kalhan A., 2013: 1119]. However, AI can implement even more creative and sensitive function as lawmaking and representation in the future.

That is why to understand the possible transformation of structure of power, it is crucial to understand real sense and function of each main element of the modern structure, examine them and propose which function AI may do better and in which circumstance.

2. The Sense of the Authority

The implementation of the AI in the public administration and governance opens the discussion of the sense of the authority. As it was mentioned in previous part, some researchers name the system where algorithms make decisions and (or) implement decisions — algorithm authority or algorithm governance. The establishment of the new type of authority links with the problem of the legitimacy. A key question arises: Can AI possess authority, and if so, under what conditions can that authority be considered legitimate? Drawing from classical theories of political legitimacy, particularly those of Max Weber and David Easton, this chapter examines foundations of belief in political systems and evaluates whether similar mechanisms can apply to AI-driven governance.

Max Weber's tripartite classification of authority — traditional, charismatic, and legal-rational — provides a foundational framework for analyzing legitimacy [Weber M., 1947: 328]. However, as this chapter argues, algorithmic authority does not neatly fit into any of these

categories. Instead, it introduces a new form of epistemic authority, grounded in expertise and data-driven rationality. Meanwhile, Easton's distinction between specific and diffuse support helps explain how citizens might come to accept AI governance—not necessarily because of satisfaction with specific outcomes, but through generalized trust in the system's perceived fairness, transparency, and purpose [Easton D., 1975: 436–437]; [Easton D., 1979: 278–319].

The algorithm authority cannot be the object of tradition. However, there is a room for assumption about charisma and legality. To generalize issue of the legitimacy the chapter proceeds in two parts: Exploration of belief and trust in AI systems; Discussion of ideology and ethics in algorithmic governance.

2.1. Belief and Trust

Trust constitutes a foundational element of any functioning political system. In democratic societies, belief and trust typically derives from shared values, transparent procedures, and institutional accountability mechanisms. However, the delegation of decision-making authority to opaque or autonomous AI systems disrupts traditional sources of trust. AI has often been characterized as a "black box", due to its complexity and lack of interpretability, which poses significant challenges for policymakers seeking to legitimize its use within public administration.

Jacopo Scipione identifies three essential preconditions for establishing trust in AI-based decision-making [Scipione J., 2020]. Alignment with human values; Responsiveness to human control; Direct oversight by humans. While these conditions may be effective in the short term, they may not fully address long-term shifts in public attitudes toward increasingly autonomous systems. For instance, historical analogies such as religious institutions and their role in legitimizing supernatural authority — demonstrate that trust does not always depend on transparency or human control. Priests, for example, gained authority not necessarily through democratic legitimacy, but through perceived divine endorsement. Similarly, if AI systems acquire symbolic or normative authority, they may not require continuous alignment with human values or direct oversight to gain acceptance.

Nevertheless, this paper focuses on modern liberal democratic frameworks where trust is grounded in rational-legal legitimacy. Within such contexts, one key factor influencing trust is transparency in the de-

cision-making process. As D. Estlund argues, opacity in administrative decisions can lead to non-compliance or diminished public confidence [Estlund D., 2003: 53–69]. G. Gaus further contends that decision-making procedures must be rationally acceptable to those affected by them in order to maintain legitimacy [Gaus G., 2010: 36–38].

However, full transparency is not always feasible. Commercial secrecy, national security concerns, and technical complexity often limit access to critical information. While absolute openness may not be attainable, it is crucial that core algorithms impacting public policy remain subject to scrutiny through mechanisms such as public audits or independent oversight bodies[2]. Ultimately, trust in AI governance is mediated through intermediary institutions, particularly legislative representatives who act as gatekeepers of sensitive information. When these actors lack sufficient access or influence over algorithmic processes, public trust erodes significantly—even in countries with strong parliamentary traditions like the United Kingdom or the United States, where suspicions of a "deep state" have grown.

A second challenge lies in the comprehensibility of AI systems. Even when information is publicly available, its complexity often exceeds the understanding of the general population. Unlike traditional expert knowledge, algorithmic logic operates at a level of abstraction that is inaccessible to most individuals [Andrejevic M., 2014: 1673–1689]. It creates what some researchers call "invisible barbed wire" — a subtle form of constraint where individuals outsource comprehension and decision-making to other AI systems, effectively reducing personal autonomy. The resulting "big data divide" exacerbates social inequalities between those who design and control AI systems and those who are governed by them.

The disbalance in society leads us to the concrete bargain: people gift their trust and their right to have access to the information, and they need protection of their interests in return. The implementation of the AI would not change the sense of that negotiations. Even if the agency will be artificial, it needs some mechanisms which may make people sure about the benefits of their contract. Today such disbalance is visible

[2] For example, the American state of Idaho has passed a law requiring all pretrial risk assessment tools be transparent, compelling the builders of these tools to make their algorithms' inputs open to public inspection and allow criminal defendants to request access to the calculations and data that determine their risk assessment scores. Idaho Code. § 19-1910. 2019.

problem for the modern democracies where the private companies intend to replace classic democratic institutions, because these companies know more about us than we know about them [Zuboff S., 2019: 38]. Additionally, the level of trust to the apps are higher than to the social institutions. Although trust in consumer applications often surpasses that in formal institutions, this dynamic should not be uncritically extended to governance. Public trust in AI requires robust safeguards against the concentration of unchecked authority.

The described desires to have access to the information are explained by the human fear: fear to lost control over AI and lost human dignity. The lack of responsibility provokes the decrease of the trust to the system. In a discretionary system, someone must be held responsible for those decisions and be able to give reasons for them. There is a legitimate fear that in a "black box" system used to produce a decision, even when used in coordination with a human counterpart or oversight, creates a system that lacks responsibility [Olsen H. et al, 2019]. Even through these analyses we distinguish the problem of the AI responsibility as a cornerstone of the topic.

Loss in human dignity is connected, but different side of the upcoming fear. If legal processes are replaced with algorithms, there is a fear that humans will be reduced to mere "cogs in the machine". The interaction with the same creature is more comfortable for human. However, "the from office" of the administration can be more "human". This issue extends beyond the scope of algorithmic accountability and reflects deeper shifts in societal values. The inclusivity in the society was the consequence of the mobilization of all masses. People was the important resource for the many projects: from the Egyptian pyramids to the battles in the Second World War. In the future people will be not so important because the majority of their functions would be made by AI. The people will lose their social utility which leads to the loss of the human dignity. The issue of the people's utility is another fundamental challenge, which is not the subject of the research.

Of course, the use of AI may have the opposite side. By limiting the role of human discretion and intuition and relying upon computer-driven decisions this process protects minorities and other weaker groups [Zarsky T., 2012: 33–35]. Fairness and discrimination in algorithmic systems are globally recognized as topics of critical importance [O'Neil C., 2017]. Danaher proposes to balance the loss in comprehension and participation against the potential gains in outcomes and procedural

fairness [Danaher J., 2016: 257]. However, it is more technical question then social. The role and utility of the people may change dramatically, and AI will just represent this reality. The legal status of the people can be reviewed in favor of the less equal and guaranteed rights to the more flexible system. Thus, this problem will be the object of the ideology of the concrete society.

2.2. Ideology and Ethics

The engineering of social institution, including the social institutions based on AI, needs the ideological background. In different times the role of ideology had been played by different things: the religion, science etc. The basic question of the AI decision making system is "Why people should obey the decisions?". We distinguished that people for voluntarily obeying need the explanation. The ideology tries to explain it. If we take any ideology, they propose the model of ideal or most appropriate society.

AI is a technological tool for the institutional changes. However, there is no preliminary understanding which institutional changes AI performs. These changes can be completely different according to the ideology of society and the creators of the concrete AI. In spite of the significances of the mathematician methods and openness of the information, it is important to input the social believes and the values. The example of the easiest ideology it is easy to find in Azimov's Laws [Azimov I., 1950]. Even very democratic approach for the creation of the AI may face with differences in humans' cultures and values. Of course, there are plenty of values, which are supported by the overwhelming majority of planet's population. However, AI "learning process" based on the decisions made by people. Thus, the same technological product will evolve in two different AI, for example, in China and France. The source of the AI decisions would be the answers of the concrete population, and the values of Chinese and French people in some important topics can be even opposite.

Geiger argues algorithms cannot be divorced from the conditions under which they are developed and deployed [Geiger S., 2014: 346–347]. Moreover, the implementation of the "foreign" AI may provoke the resistance of the people. The creation of the AI involves dozens of social and material practices that are culturally, historically and institutionally situated [Napoli P., 2013].

Here it is crucial to understand that the trust and belief do not eager the western democratic institutions. The level of trust in authoritarian countries may be much higher. For example, the approval of the Social Credit System within the Chinese populace is high [Kostka G., 2019]. However, the implementation of the same system in European's countries would face with tremendous opposition. Some commentators have emphasized that the Social Credit System may be positively received as a response to the perception of moral decline in China, and a concomitant desire to build greater trust [Roberts H. et al., 2021: 67]. That is why the main factors of the trust availability are cultural features and marketing tools. Thus, Robin Li, co-founder of Baidu, stated, "the Chinese people are more open or less sensitive about the privacy issue. If they are able to trade privacy for convenience, safety and efficiency, in a lot of cases, they are willing to do that"[3]. That is why the level of trust within Chinese society can be much higher than in western democracies. However, the democratic institutions are very attractive for general population and inclusive function, which is provided by increase the chances of the higher trust within society. Democratic institutions help to grow the population confidence in foreseeability and that AI system will be under their control [Scherer M., 2015: 378–379]. However, the trust is more complex phenomenon and the trust to some people is exit without foreseeability and control (trust to parents, trust to family partner etc.).

For example, the EU tries to increase the trust with a development of human-centric approach on AI. This approach makes both: put humans at the center of AI developments and design a Trustworthy AI. The legal regulation keeps the human as a responsible person. Even if AI has a certain amount of autonomy, a human operator should always be accountable for its actions. Section 5 of the EU White Paper on Artificial intelligence named "An Ecosystem of Trust: Regulatory Framework for AI", stresses on the need of creating a unique "ecosystem of trust". A version of this solution is already part of the law in the European Union. According to article 15 of the European Directive 95/46/EC (the Data Protection Directive), there must be human review of any automated data-processing system that could have a substantial impact on an individual's life. The Directive does, however, allow for certain exceptions to this rule. Specifically, it allows for people to voluntarily contract themselves out of this right and for governments to override it

[3] Are Chinese People 'Less Sensitive' About Privacy? // Available at: URL: https://www.sixthtone.com/news/1001996/are-chinese-people-less-sensitive-about-privacy%3F (accessed: 25.01.2025)

so long as other measures are taken for protecting the individual's "legitimate interests"[4].

That is why there are no universal ideology, which may answer the upcoming challenge. According to the valuation of the concrete phenomena, AI may perform different decisions. It is difficult for AI to resolve opposite goals, such as social equality and maximization of the productivity. The ideology has to provide the hierarchy of the values, which is the cornerstone for such kind questions. David Easton, one of the leading figures in political systems theory, conceptualized the political system as a "black box". Easton famously defined politics as the authoritative allocation of values for a society [Easton D., 1979: 32]. It is obvious that the AI decisions of the same problem in socialistic and capitalistic country can be different, but the "authoritative allocation" will exist anywhere.

There is no doubt AI and digital world in general changing the human culture. The crucial changes may provoke the ideological vacuum, where no ideologies already existed may match the new society. Thus, some authors try to examine the ideas of the personhood and classic rationality. S. Mhlambi argues that rationality and dehumanization are linked and the implementation of the AI demands to rethink the idea of personhood in more "collective" way [Mhlambi S., 2020: 11]. This self-similarity is reflected in ubuntu's commonly cited aphorisms "I am because you are," and "a person is a person through other persons" [Mbiti J., 1970: 138–142]. However, it is just the one of the possible scenarios.

Thus, utilitarianism and principled ethics pushed AI to make completely different choice working with the same information. C. Djeffal explains that actions detrimental to one person but beneficial for the majority could be regarded as ethical from a utilitarian perspective, they would be regarded as unethical from a principled point of view [Djeffal C., 2019: 274]. However, it is problematic to be sure that AI make a moral choice, the decision of the AI is predictable in the concrete situation. In such a setting, there is no room left for choice. This problem is tied to the question whether machines can actually think, which has attracted contentious reflection from Turing to Searle.

To build the ideological background for the AI we need to answer for Baum's questions [Baum S., 2017: 543–551]:

[4] Directive 95/46/EC, Art. 15.3.

Standing: Who or what is included in the group to have its values factored into the AI?

Measurement: What procedure is used to obtain values from each member of the selected group?

Aggregation: How are the values of individual group members combined to form the aggregated group values?

Some researches believe the concept of "algocracy" has enough ideological background. However, the concept needs the additional explanations. The absolute monarchy usually explained through the religion and customs. Algocracy has a huge advantage in rational explanation: the system in which power is (increasingly) exercised by automated systems is more fruitful for society [Yeung K., 2018: 512–514]. The term algocracy is mostly used in a critical manner [Danaher J., 2016: 246].

However, the algocracy is not entire ideology, it is more applicative to the ideologies, which explain the source of the public power in society. A frame that is complementary to algocracy would not exclusively look at the fact that decisions are delegated, but at *how* they are delegated and *who* controls and influences the automated systems. One example would be to empower voters through targeting and profiling candidates. A smart search engine could help to identify information concerning how parties or candidates think about some issues [Djeffal C., 2019].

The author has to agree that algocracy bases on the same provisions as epistocracy does. The justification of the algorithm governance correlated with epistocracy. Thus, epistocratic systems of governance embody set of epistemic elites over the broader public [Estlund D., 2003: 55–57]. It is even possible to reuse Lenin's famous definition of socialism, "Soviets plus electrification" to the algocracy, "Epistemic elites plus AI".

Estlund points out that if we assume that legitimacy-conferring outcomes are more likely to be achieved by those with better epistemic abilities, then the following argument seems compelling [Danaher J., 2016: 246–251]:

There are procedure-independent outcomes against which the legitimacy of public decision-making procedures ought to be judged. (Cognitivist thesis)

In any given society, there will be a group of people with superior epistemic access to these procedure-independent outcomes. (Elitist thesis)

If there are people with superior epistemic access to these procedure-independent outcomes, then procedures are more likely to be legitimate if those people are given sole or predominant decision-making authority.

Therefore, in any given society, decision-making procedures are more likely to be legitimate if authority is concentrated in an epistemic elite. (Authority thesis).

The AI authority can be justified through different ways. The most appropriate way to legitimate the AI authority is to make it legal. However, the legal basis must be founded on a sort of ideology. From society to society this ideology can be different, but the common core of the justification is laying in the epistocracy provisions. Liberal democracy as a dominant ideology faces the most difficult challenge in upcoming changes.

3. The New Structure of Power

The modern structure of state power was developed with consideration of human nature and the balance of interests among different social groups. There is no doubt that the implementation of AI will not alter the fundamental dynamics of interest balancing, as public authorities will continue to strive for societal stability. However, AI will necessitate a rethinking and simplification of the present-day structure to enhance governance effectiveness.

In assessing how to respond to the emerging phenomenon of algocracy — defined as governance by algorithms — it becomes essential to weigh the potential losses in comprehension and citizen participation against the possible gains in procedural fairness and decision-making outcomes [Danaher J., 2016: 257]. The future structure of governance will be shaped precisely by this balancing act.

The balance of interests between the state, business, and academia differs significantly in China, the United States, and the European Union. As a result, the new structure of power may also vary. Many social constructs surrounding AI systems play a crucial role [Stamper R., 1988: 14–15], and the structure of state power can be fundamentally different even when the same technology is used. As discussed in the previous chapter, the consequences of AI in public management depend on the individuals who create it and the specific features of the algorithms

involved. People may use AI as a tool to replace traditional social institutions. AI may perform the same functions as the legislative, executive, and judicial branches of government.

The most influential idea regarding the structure of power is that of the separation of powers, based on the concept of checks and balances. Criticisms of this idea serve as an excellent case study for examining AI's influence. The triumphalism surrounding the Western, especially American, export of public law and governmental structures extends far beyond its borders [Calabresi S., 1998: 22]. The implementation of algocracy simplifies the system of governance by eliminating unnecessary functions within large and decentralized government systems. The separation of powers is a complex system that emerged due to the intricacies of social organization and high transaction costs associated with trust among individuals in society. It is necessary to agree with B. Ackerman, who emphasizes the separation of powers in favor of three principles: democracy, professional competence, and the protection and enhancement of fundamental rights [Ackerman B., 2000: 639–640].

Democracy, as a value of modern society, is not absolute but offers advantages for sustainable governance, including easier legitimization of authority and shared responsibility in the decision-making process. However, the separation of powers presents certain challenges that do not necessarily support democratic trends. Deadlocks between different branches or fragmentation of political views are issues that may be resolved but require strong and effective institutions. This is why Montesquieu's dictum has led to the erosion of democratic foundations in many countries, particularly in Latin America.

Additionally, new technologies may ensure the same level of citizen participation without relying on parliamentarism or legislation. Blockchain systems can organize analogs of elections or referendums without the need for bureaucrats or specialized electoral bodies. Transformations in transaction and agency costs through blockchain interventions [Sun R. et al., 2020: 9–13] reshape the institutional framework of democratic societies. They foster forms of direct democracy, shifting its applicability from the local to the national level. AI and blockchain will drastically reduce the transaction cost of trust in public governance. The machine-learning process accounts for the "vote" of each individual participating in the process. The real question here concerns people's willingness, their competence, and their trust. Voter absenteeism remains a problem even in modern representative democracies; however,

given everyday routines, people may logically refuse to engage in all public matters. On the other hand, the activism of uninformed individuals may lead to unprofessional and harmful decisions in public administration. This highlights the importance of delegation, which could potentially be directed towards AI rather than human representatives. Thus, AI may fully reflect the essence of vox populi, or at least lead the policy of the majority or consensus-based.

Professional competence serves as another supporting argument. It is logical that a monopoly on power may lead to the degradation of social mobility. The system of checks and balances, however, is not a "magic potion" capable of overcoming this regression. Historical evidence shows that authoritarian and totalitarian regimes have summoned high-level bureaucrats to serve for public purpose. More complex social institutions demand higher levels of social science knowledge from the population. In non-democratic societies, the elite carefully monitors the limitations of an incompetent leader, while the public remains susceptible to the ruler's propaganda [Guriev S., Treisman D., 2019: 101]. Nevertheless, parliaments, as representative bodies, often lack expertise in specific areas and rely on input from executive bodies or private companies.

The primary argument in favor of the liberal democratic system is the protection and promotion of fundamental rights. The situation concerning the protection of human rights becomes predictable once the actual balance of power in society is determined. Centralizing authority in AI poses risks to minorities and vulnerable groups. Even current implementations of AI in social networks exemplify the suppression of freedom of expression and assembly. The separation of powers may aid in protecting human rights by preventing the concentration of power in one entity. AI as an actor might centralize power, but the decision-making process is more intricate and involves individuals advocating for human rights protection. The «new structure of power» must embody the processes of algorithm creation, oversight of their implementation, and the correction of any flaws.

Consequently, one of the core principles of liberal democratic constitutionalism faces threats. Simplification of the system appears to be an inevitable path forward. Current challenges already prompt states to rebalance authority and delegate more power to specific executive bodies. Regarding AI-related issues, new institutions are emerging in various countries: For instance, the United Arab Emirates appointed a

minister for AI, while the German government established an agency for "innovation leaps" among others.

Upcoming changes directly affect constitutional law regulation. Even transferring competencies to AI within the traditional structure of power requires serious justification. It would be intriguing to apply the logic of the German Federal Constitutional Court in this context. In its famous Lisbon judgment, the court permitted the transfer of competencies but also required institutional arrangements within the German legal order, enabling the legislature to actively participate in European politics[5]. The same provision could apply to delegating competencies to AI. The expectation is clear: constitutional bodies (legislative, executive, and judicial branches) must possess strong tools to influence AI.

However, we delve deeper into a discussion about the relevance of the modern structure of power in general. The main critique will focus on legislative power and parliamentary bodies. The implementation of AI and other technologies, such as blockchain, renders parliament increasingly obsolete. Today, legislators lack the flexibility and operational efficiency of the executive branch, leading to substantial transfers of responsibilities from legislators to executive bodies. This trend is partly explained by the relative lack of expertise in emerging technologies. Agencies typically employ experts with specialized knowledge in relevant fields, whereas legislators generally rely on committee hearings and interactions with lobbying groups to access expert opinions on proposed legislation. There is no doubt that agencies possess a clear advantage over legislatures and courts in terms of institutional flexibility [Viscusi W., 1989: 73-74]. Hence, the trend of transferring responsibilities to these specialized agencies is both logical and reasonable.

Despite these concerns legislatures remain the institutions best suited to make policy decisions involving significant ethical considerations and those prioritizing democratic legitimacy [Scherer M., 2015]. This is because legislators are elected at regular elections and maintain greater openness to the general public. Consequently, legislative enactments carry more democratic legitimacy than agency rules or court decisions [Pound R., 1978: 400].

Economic development and the actual balance of power within society shape the necessity and function of public authority, including par-

[5] BVerfG. Judgment of the Second Senate of 30 June 2009—2 BvE 2/08—para. (1-421) // Available at: URL: http:// www.bverfg.de/e/es20090630_2bve000208en.html para 273ff (accessed: 16.02.2025)

liamentary institutions. Within this context, we can understand the emergence of legislative bodies in ancient times. These developments were closely tied to specific patterns of economic growth in early societies.

For example, in ancient Greece labor productivity increased significantly in urban and rural economies where feudal forms of dependency were absent. This led to the expansion of commodity production, trade, and shipbuilding — economic activities that empowered broader segments of the population. Consequently, the role of the common people — the demos — grew, particularly among those engaged in trade, crafts, and maritime commerce.

However, this rising social group encountered resistance from the traditional aristocracy — the eupatridae — who clung to inherited political, economic, and social privileges. The resulting tensions between these classes necessitated new mechanisms of governance and conflict resolution.

As society became more complex, so did its internal relationships, especially concerning property rights and legal disputes. Matters previously settled according to ancestral customs began requiring more formal, publicly recognized regulations. Laws thus emerged as structured methods to regulate social relationships and ensure fairness — laying the foundation for early legislative and judicial institutions.

Parliament continues to lose its significance in the modern system of governance. Other institutions assume parliamentary functions, such as providing a platform for public discussion and civil control. It is crucial to recognize that parliament is not an indispensable institution within the "new structure of power". Modern political developments in many countries — even in Western democracies — make the ideas of the German legal scholar Georg Jellinek increasingly relevant. He regarded parliament as the central element of parliamentarism but did not consider it among the most critical state institutions. In authoritarian states, parliament has become a tool for the executive, while even in democratic states, parliament cannot claim independence, as it represents the will of certain groups whose actions may not directly impact the state or its citizens [Jellinek G., 2004: 425–428].

In recent years, several Western scholars have argued that the state in modern Western societies is increasingly transforming into a technical or bureaucratic mechanism, marked by a growing tendency toward the depoliticization of governance. This transformation signifies a shift

away from traditional models where political power dominated decision-making toward systems where administrative expertise holds sway [Crouch C., 2004: 73–75].

From this perspective, public authority today is no longer directly linked to property ownership, nor does economic wealth necessarily translate into political influence. Instead, power is perceived as concentrated within a professional political elite — comprising bureaucrats, state officials, and technocrats — who operate with a notable degree of autonomy. Access to information has become the primary indicator of power, and the implementation of AI further reinforces this point.

Michel Crozier highlight how bureaucratic organizations develop their own internal logic, often resisting external control, including from political and economic actors. He noted that once established, such organizations tend to generate and maintain their own power independently of those who originally created them [Crozier M., 1964: 184–188]. This insight supports the view that state institutions can function autonomously from the public will, representing the "black box" even without AI.

Similarly, Gianfranco Poggi emphasized the institutional autonomy of the modern state, arguing that it has become an entity in its own right, pursuing goals that may diverge from those of dominant social groups [Poggi G., 1978: 127–137]. His analysis reinforces the idea that state action is not always aligned with economic elites but follows its own institutional imperatives. It was not the new idea to focus on the institutional autonomy of the bureaucracy within the state. For example Ernst Fraenkel distinguish "normative state" as an administrative body endowed with elaborate powers for safeguarding the legal order as expressed in statutes, decisions of the courts, and activities of the administrative agencies, and the normative state survived even in the Third Reich [Fraenkel E., 1941: 60–63].

Colin Crouch, building on these ideas, introduced the concept of "post-democracy", describing "a model, while elections certainly exist and can change governments, public electoral debate is a tightly controlled spectacle, managed by rival teams of professionals expert in the techniques of persuasion, and considering a small range of issues selected by those teams" [Crouch C., 2004: 4]. As he observes, power is increasingly exercised by officials and experts who are not accountable in the traditional democratic sense. The key conclusion here is that the modern system is ready for the integration of AI, with changes likely to be less noticeable to the general public.

This does not imply that fundamental concepts of parliament (such as Dicey's principle of parliamentary sovereignty) will vanish. "The right to make or unmake any law whatever" [Dicey A., 1985: 3-4] may persist, but the understanding of parliament will evolve. It is essential to establish common rules for all members of society, which is difficult to achieve due to human nature and the desire to avoid Locke's notion of the "war of all against all". It is important to have a body capable of making final decisions on crucial questions, whether through consensus (e.g., a democratic parliament) or authority (e.g., a dictator).

It is understandable for the author that criticism of parliament is not a new topic; however, AI technology may catalyze a shift in social development and redistribute the classical functions of parliament to other entities or transform parliament itself. Parliament is not the entirety of the state; it is merely an organ through which certain state functions are executed [Jellinek G., 2004: 431]. Even twentieth-century views on parliamentary functions appear somewhat reluctant. Accountability and criticism, two primary parliamentary functions, have migrated to other platforms. Media, expert councils, and NGOs sometimes play a more active and impactful role in fulfilling these functions. While some traditional institutions face crises, they are often accompanied by the rise of new forms of engagement, such as grassroots democracy, diverse civic initiatives (not always politically oriented), and decentralized communication networks. However, it would be premature to completely dismiss old institutions — especially before effective alternatives have been developed [Petukhov V., Petukhov R., 2015: 32]. The only barrier to fully implementing these functions lies in accessing the necessary information for members of parliament.

Consequently, a democratic society, where the people are the source of power, requires representatives who can access confidential information and protect public interests — or at least the interests of the social group they represent. Parliament is not necessarily the optimal tool or universal platform for this purpose. The evolution of expert councils around the executive branch appears more efficient than using parliament as a universal collective body composed of individuals who either understand or may understand any regulatory topic and are sincere in their commitment to protecting public interests. Therefore, it is challenging to support the view that democracy is impossible without parliamentary democracy [Kerimov A., 2018: 30]. Democratic governance is ensured through two main elements: electoral procedures and the decentralization of power. The former helps express and account for

the public will, while the latter prevents the erosion of this will. Decentralization of power is achievable through the separation of powers at one level and across different levels. The separation of powers does not conflict with the idea of a unified state authority. Rather, it entails distributing roles and powers among various branches of government while maintaining the need for cooperation. The unity of the state's power structure and the prevention of dictatorial control are achieved through balanced interaction among all governmental branches, ensuring that no single branch holds absolute authority.

In an AI world, it is far more effective to coordinate between centers of expertise than between the traditional legislative, judicial, and executive branches of government. The level of expertise required to integrate AI into decision-making must be high, and only a few individuals may grasp the nuances of specific cases. This does not mean that society no longer needs supreme bodies; quite the contrary — the control over AI is even more critical than over humans. However, if the public grants AI diffuse support, believing it better represents their interests than people do, it will be challenging to establish sustainable oversight over it.

Conclusion

The integration of AI into public administration signifies not merely a technological evolution, but a profound reconfiguration of the foundational principles that underpin modern governance. As explored throughout this article, the deployment of AI disrupts traditional power structures, challenges established conceptions of legitimacy, and necessitates a re-evaluation of the relationships between states, citizens, and technology. Empirical evidence increasingly supports the hypothesis that AI transforms the classical architecture of state power. Algorithmic governance systems are progressively replacing or augmenting human decision-making in areas such as law enforcement, regulatory compliance, social welfare, and even legislative drafting. However, this transformation is neither ideologically neutral nor universally beneficial. It raises urgent questions regarding accountability, ethical design, and the ideological frameworks guiding the implementation of AI within the public sphere. Navigating these complexities requires societies to critically engage with the interplay of technical feasibility, political will, and moral responsibility, ensuring that AI functions as a tool for empowerment rather than a mechanism of control.

The article identifies three primary challenges associated with AI-driven governance: the legitimacy of AI authority; the opacity of algorithmic decision-making processes; and the potential erosion of human dignity. One plausible conceptualization of AI authority is algocracy — a form of governance wherein decisions are made or enforced by algorithms. Algocracy shares characteristics with epistocracy, a system in which authority is concentrated among individuals or entities possessing superior knowledge. While algocratic systems may offer advantages in terms of efficiency and data-driven rationality, they also inherit the limitations of epistocratic models, particularly concerning democratic legitimacy and inclusivity. Furthermore, the inherent complexity of algorithmic logic necessitates the development of new legal frameworks, the simplification of existing state structures, and the adaptation of ideological narratives to align with emerging governance paradigms.

Central to the argument is the observation that AI implementation destabilizes the traditional functions of core state institutions, particularly parliaments. Drawing upon Jellinek's theory of state organs, it is evident that legislative institutions are not the state itself, but mechanisms through which specific state functions are executed. The rise of algorithmic governance accelerates the marginalization of these traditional organs. Accountability, once a cornerstone of parliamentary oversight, increasingly migrates to opaque technical systems and technocratic elites. While AI promises notable efficiency gains — such as predictive policing reducing crime rates or machine learning optimizing resource allocation — the depoliticization of governance poses significant risks to democratic legitimacy. The tension between procedural fairness and outcome-oriented efficiency becomes especially acute when algorithms—often designed with embedded biases or operating as "black boxes" — make life-altering decisions in domains such as credit scoring, immigration, and criminal justice.

The ethical implications of AI governance are deeply intertwined with the ideological frameworks that guide its deployment. This article's analysis of epistocracy reveals a paradox: although AI systems may surpass human capabilities in processing information and minimizing errors, their legitimacy ultimately depends on societal acceptance of technocratic rule. This dynamic manifests differently across geopolitical contexts. In China, for instance, the Social Credit System leverages AI to enforce social conformity, reflecting a collectivist ideology that prioritizes stability over individual autonomy. In contrast, the European Union's human-centric AI strategy emphasizes transparency, fairness,

and respect for fundamental rights, mirroring liberal democratic values. These divergent approaches underscore the absence of a universal ethical framework for AI governance, highlighting the need for context-sensitive regulatory and normative responses.

Consequently, the emerging structure of power is likely to revolve around the evolving process of lawmaking. Understanding how AI reshapes legislative practices requires further empirical and theoretical research. Initial engagement with AI in lawmaking demands high levels of IT expertise and sociological insight. Moreover, the outcomes generated by AI systems are contingent upon the quality and nature of the data used, which can significantly influence final decisions. Technological advancements will likely simplify certain parliamentary functions, shifting some responsibilities toward decentralized citizen networks and others toward specialized executive bodies. Rather than opposing current trends, AI is expected to amplify them. Centralized power will increasingly reside within expert-led executive agencies supported by AI, enabling more efficient and specialized decision-making in complex policy domains. Such bodies may be better equipped to address interdisciplinary issues — such as those involving agriculture, taxation, and environmental regulation — than traditional legislative assemblies.

Contemporary lawmaking already relies heavily on expert input, yet it often involves numerous intermediaries whose roles remain ambiguously defined and largely disconnected from substantive public interest representation. Thus, the long-term transformation of the power structure may reinforce liberal constitutionalism, albeit requiring a rethinking of the classical doctrine of separation of powers. The core principle — preventing the concentration of power — will remain intact, though its practical realization will shift from the traditional tripartite model (executive, legislative, judicial) to a decentralization based on spheres of knowledge or regulatory domains. AI, as a technology that diminishes the role of intermediaries, embodies the tools of algocracy. It redistributes power not according to functional branches of government, but along thematic lines of expertise — redefining the very architecture of governance in the digital age.

References

1. Ackerman B. (2000) The New Separation of Powers. *Harvard Law Review*, vol. 113, pp. 633–729.

2. Anderson C.W. (2011) Deliberative, Agonistic, and Algorithmic Audiences: Journalism's Vision of its Public in an Age of Audience. *Journal of Communication*, no. 5, pp. 529–547.

3. Andrejevic M. (2014) The Big Data Divide. *International Journal of Communication*, no. 8, pp. 1673–1689.

4. Azimov I. (1950) Me, Robot. New York: Doubleday, 218 p.

5. Barocas S., Hood S., Ziewitz M. (2013) Governing Algorithms: A Provocation Piece. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2245322

6. Baum S. D. (2017) On Promotion of Safe and Socially Beneficial AI. *AI and Society*, no. 32, pp. 543–551.

7. Calabresi S.G. (1998) An Agenda for Constitutional Reform. In: W.N. Eskridge (ed.). Constitutional Stupidities, Constitutional Tragedies. N.Y.: New York University Press, pp. 22–27.

8. Chomsky N. (2010) The Chomsky Reader. N.Y.: Knopf Doubleday,

9. Citron D., Pasquale F., (2014) The Scored Society: Due Process for Automated Predictions. *Washington Law Review*, vol. 86, pp. 1–33.

10. Coglianese C., Ben Dor L.M. (2021) AI in Adjudication and Administration. Faculty Scholarship at Penn Law. 2118, pp. 791–838.

11. Coglianese C., Lehr D. (2019) Transparency and Algorithmic Governance. *Administrative Law Review*, no. 71, pp. 1–56.

12. Crouch C. (2004) *Post-Democracy*. Cambridge: Polity Press, 144 p.

13. Crozier M. (2017) *The Bureaucratic Phenomenon*. L.: Routledge, 320 p.

14. Dicey A.V. (1985) Introduction to the Study of the Law of the Constitution. Indianapolis: Liberty Fund Press, 436 p.

15. Danaher J. (2016) The Threat of Algocracy: Reality, Resistance and Accommodation. *Philosophy & Technology*, no. 29, pp. 245–268.

16. Danaher J., Michael J. et al. (2017) Algorithmic Governance: Developing a Research Agenda through the Power of Collective Intelligence. *Big Data and Society*, no. 4, pp. 7–14. Available at: URL: https://doi.org/10.1177/2053951717726554

17. Ding J. (2018) Deciphering China's AI Dream. Future of Humanity Institute. Oxford: University Press, 44 p.

18. Djeffal C. (2019) AI, Democracy and the Law. In: A. Sudmann (ed.) The Democratization of Artificial Intelligence: Net Politics in the Era of Learning Algorithms. Bielefeld: Transcript Verlag, pp. 255–284.

19. Dodge M., Kitchin R. (2007) The Automatic Management of Drivers and Driving Spaces. *Geoforum*, no. 2, pp. 264–275. https://doi.org/10.1016/j.geoforum.2006.08.004

20. Domingos P. (2015) *The Master Algorithm: How the Quest for Ultimate Machine Learning Will Remake Our World.* N.Y.: Basic Books, 330 p.

21. Easton D. (1975) A Re-Assessment of the Concept of Political Support. *British Journal of Political Science*, no. 4, pp. 435–457.

22. Easton D. (1979) A Systems Analysis of Political Life. Chicago: University of Chicago Press, 507 p.

23. Estlund D. (2003) Why not Epistocracy? In: N. Reshotko (ed.) Desire, Identity, and Existence. S. L.: Academic Printing and Publishing, pp. 53–69.

24. Fraenkel E. (1941) The Dual State. Oxford: University Press, 248 p.

25. Gaus G. (2010) *The Order of Public Reason: Theory of Freedom and Morality in Diverse and Bound World.* Cambridge: University Press, 964 p.

26. Geiger S.R. (2014) Bots, Bespoke, Code and the Materiality of Software Platforms. *Information*, *Communication and Society*, vol. 17, no. 3, pp. 342–356.

27. Guriev S., Treisman D. (2019) Informational Autocrats. *The Journal of Economic Perspectives,* no. 33, pp. 100–127.

28. Ho D.E., Engstrom D.F. (2021) Artificially Intelligent Government: A Review and An Agenda. In: R. Vogl (ed.) Research Handbook on Big Data Law. Northampton: Edward Elgar Publishing, pp. 57–86.

29. Jellinek G. (2004) *General Theory of the State.* Saint Petersburg: Juridicheskiy Center Press, 599 p. (in Russ.)

30. Kalhan A. (2013) Immigration Policing and Federalism Through the Lens of Technology, Surveillance, and Privacy. *Ohio State University Law Review,* no. 74, pp. 1106–1165.

31. Kerimov A.A. (2018) *Institute of Parliament in Legitimation of Political Power in Modern Russia.* Ekaterinburg: Ural State University Press, 343 p. (in Russ.)

32. Kitchin R. (2017) Thinking Critically about and Researching Algorithms. *Information*, *Communication and Society*, no. 20, pp. 14–29.

33. Kostka G. (2019) China's Social Credit Systems and Public Opinion: Explaining High Levels of Approval. *New Media and Society*, no. 21, pp. 1565–1593. https://doi.org/10.1177/1461444819826402

34. Kushner S. (2013) The Freelance Translation Machine: Algorithmic Culture and the Invisible Industry. *New Media and Society*, no. 8, pp.1241–1258.

35. Mbiti J. (1970) African Religions and Philosophy. N.Y.: Anchor Books, 288 p.

36. Metz C., Satariano A. (2020) An Algorithm that Grants Freedom, or Takes it Away. *The New York Times*. June 2.

37. Mhlambi S. (2020) From Rationality to Relationality: Ubuntu as an Ethical and Human Rights Framework for AI Governance. Carr Center Discussion Paper Series, no. 9, pp. 1–27.

38. Miyazaki S. (2012) Algorhythmics: Understanding Micro-temporality in Computational Cultures. *Computational Culture,* no. 2, pp. 1-20.

39. Musiani F. (2013) Governance by Algorithms. *Internet Policy Review*, no. 2, pp. 1–8. https://doi.org/10.14763/2013.3.188

40. Morozov E. (2019) The Real Privacy Problem. https://www.technologyreview.com/2013/10/22/112778/the-real-privacy-problem/

41. Napoli P.M. (2013) The Algorithm as Institution. Fordham University Schools of Business Research Paper. https://dx.doi.org/10.2139/ssrn.2260923

42. Olsen H.P., Slosser J.L. et al. (2019) What's in the Box? The Legal Requirement of Explainability in Computationally Aided Decision-Making in Public Administration. Courts Working Paper Series. №162. http://dx.doi.org/10.2139/ssrn.3402974

43. O'Neil C. (2017) Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. N.Y.: Broadway Books, 272 p.

44. Peter F. (2017) Political Legitimacy. In: E.M. Zalta (ed.) Stanford Encyclopedia of Philosophy. Available at: https://plato.stanford.edu/entries/legitimacy/.

45. Petukhov V.V., Petukhov R.V. (2015) Participatory Democracy: Institutional Crisis and New Prospects. *Polis*=Polis, no. 5, pp. 25–48 (in Russ.)

46. Poggi G. (1978) *The Development of the Modern State.* Stanford: University Press, 171 p.

47. Pound R. (1908) Common Law and Legislation. *Harvard Law Review*, vol. 21, pp. 383–407.

48. Relyea H.C., Hogue H.B. (2004) A Brief History of the Emergence of Digital Government in the United States. In: Digital Government. Principles and Best Practices. A. Pavlichev, G. Garson (eds.). London: Idea Group Inc., pp. 16–33.

49. Roberts H., Cowls J., Morley J. et al. (2021) The Chinese Approach to Artificial Intelligence: Policy, Ethics, and Regulation. *AI and Society*, no. 36, pp. 59–77. https://doi.org/10.1007/s00146-020-00992-2

50. Scherer M.U. (2015) Regulating Artificial Intelligence Systems: Risks, Challenges, Competencies, and Strategies. *Harvard Journal of Law and Technology,* vol. 29, pp. 353–400. http://dx.doi.org/10.2139/ssrn.2609777

51. Scipione J. (2020) Artificial Intelligence and Europe: Risks, Developments and Implications. SSRN Electron J. http://dx.doi.org/10.2139/ssrn.3598543

52. Shirky C. (2009) A Speculative Post on the Idea of Algorithmic Authority. http://www.shirky.com/weblog/2009/11/a-speculative-post-on-the-ideaof-algorithmic-authority/

53. Stamper R. (1988) Pathologies of AI: Responsible Use of AI in Professional Work. *AI and Society*, no. 2, pp. 3–16.

54. Sun R., Garimella A. et al. (2020) Transformation of the Transaction Cost and the Agency Cost in an Organization and the Applicability of Blockchain — A Case Study of Peer-to-Peer Insurance. *Frontiers in Blockchain*, no. 3, pp. 1–15.

55. Turing A.M. (1950) Computing Machinery and Intelligence. *Mind*, no. 236, pp. 433–460.

56. Viscusi W. (1989) Toward a Diminished Role for Tort Liability: Social Insurance, Government Regulation, and Contemporary Risks to Health and Safety. *Yale Journal on Regulation*, no. 6, pp. 65–107

57. Weber M. (1947) The Theory of Social and Economic Organization. New York: Free Press, 436 p.

58. Yeung K. (2018) Algorithmic Regulation: A Critical Interrogation. *Regulation & Governance*, no. 12, pp. 505–523.

59. Zarsky T. (2012) Automated Predictions: Perception, Law and Policy. *Communications of the Association for Computing Machinery,* vol. 15, no. 9, pp. 33–35.

60. Zarsky T. (2013) Transparent Prediction. *University of Illinois Law Review,* no. 4, pp. 1503–1570.

61. Zuboff S. (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* N.Y.: Public Affairs, 704 pp.

Information about the author:

V.A. Nizov — Candidate of Sciences (Law).

The article was submitted to editorial office 20.01.2025; approved after reviewing 03.03.2025; accepted for publication 30.05.2025.

Legal Issues in the **DIGITAL AGE**

AUTHORS GUIDELINES

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

Article Length

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

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

Article Title

The title should be concise and informative.

Author Details

The details about the authors include:

- Full name of each author
- Complete name of the organization — affiliation of each author and the complete postal address
- Position, rank, academic degree of each author
- E-mail address of each author

Abstract

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

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

Keywords

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

References

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

A reference list should be attached to the article.

Footnotes

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

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