ISSN 2713-2749

Legal Issues in the DIGITAL AGE

Вопросы права в цифровую эпоху

2/2024

Volume 5

Legal Issues in the DIGITAL AGE

2/2024

ISSUED QUARTERLY VOLUME 5

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

ARTIFICIAL INTELLIGENCE AND LAW

A.A. Kartskhiya
Legal Horizons of the New Artificial Intelligence Paradigm 4

I.G. Ilyin
Progress in Natural Language Processing Technologies:
Regulating Quality and Accessibility of Training Data 36

IT, INDUSTRIES, LAW: TELEMEDICINE

B.A. Edidin, A.V. Bunkov, K.V. Kochetkova
The Use of AI in Medicine: Health Data, Privacy Risks and More . . . 57

G. Rousset
The French Telemedicine System: Challenges, Procedures
and Difficulties. 80

IT, INDUSTRIES, LAW: MEDIA

Reeta Sony A.L., Shruti Chopra
Regulating Digital Era: A Comparative Analysis
of Policy Perspectives on Media Entertainment 97

E- GOVERNMENT

E.V. Burdina, V.N. Kornev
Technologies Versus Justice: Challenges of AI Regulation
in the Judicial System 113

A.A. Doschatov
Digitizing Law-Making at Federal Executive Agencies 127

REVIEW

L.K. Tereschenko, O.E. Starodubova, N.A. Nazarov
Digital Platforms in the Focus of National Law 148

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. 2024. Vol. 5. No. 2.
Вопросы права в цифровую эпоху. 2024. Том 5. № 2.

Artificial Intelligence and Law

Research article
УДК: 340, 347
DOI:10.17323/2713-2749.2024.2.4.35

Legal Horizons of the New Artificial Intelligence Paradigm

Aleksandr Amiranovich Kartskhiya

National Oil and Gas Gubkin University, 65 Lenin Avenue, Moscow 119991, Russian Federation, arhz50@mail.ru, Web of Science Researcher ID AAZ-1083-2020, ORCID 0000-0002-8041-0055, Scopus ID 57217114108, AuthorID 771380

Abstract

Modern society is undergoing a structural transformation of the world economy. This is as a result of the transition to a new technological base through the introduction of artificial intelligence, cutting-edge information and communication technology, energy technology, biotechnology and nanotechnology. Artificial intelligence has the ability to significantly change the economy and social relations in society, and its newly discovered capabilities are transformational and global in nature. At the same time, the extraordinary capabilities of artificial intelligence technologies involve risks that can threaten stability and undermine human values. In order to eliminate possible threats and risks and mitigate potential dangers, it is crucial to develop systemic legal measures and ways to regulate AI technologies and models on a national and international scale and to define the legal status of AI, which must include protection of humans from the uncontrolled influence of AI and the inviolability of guarantees of human rights and freedoms. With this in mind, and in order to mitigate potential dangers and ensure the controllability and sustainability of AI technologies based on the concept of trusted (responsible) AI, it is necessary to agree on universal international guidelines for the development and application of AI technologies and models. Furthermore, it is necessary to create a universal code of conduct for AI developers, who together can create a basis for a uniform framework of legal regulation within the national legislation of each country on the principles of human rights protection, privacy and data protection, transparency and explainability, fairness, accountability and safety of artificial intelligence, adequate human oversight and ethical standards for the creation and application of AI models.

Keywords

intellectual property; cybersecurity; trusted intelligence; neural networks; liability; artificial intelligence security; technological sovereignty; threats and risks of AI technology; sustainable development.

For citation: Kartskhiya A.A. (2024) Legal Horizons of the New Artificial Intelligence Paradigm. *Legal Issues in the Digital Age,* vol. 5, no. 2, pp. 4–35. DOI:10.17323/2713-2749.2024.2.4.35

AI has hacked the operating system of human civilization
Yuval Noah Harari. The Economist, April 23th 2023

For humanity's sake, regulation is needed to tame market forces
Helen Toner and Tasha McCauley, former OpenAI Board member.
The Economist, May 26th 2024

Introduction

Over the past decades, scholars have dissected the manifold ways in which artificial intelligence (AI) systems and digital technologies impact pillars of the law in fields such as human rights law, constitutional law, criminal law, tortious liability and contracts, administrative law, international humanitarian law, and more [Barfield W., Pagallo U., 2020: 25]. According to the European Commission High-Level Expert Group (2018)[1], the challenges brought forth by AI in the legal domain depend on the complexity, opacity, openness, autonomy, predictability, data-drivenness, and vulnerability of computers that mimic human intelligence.

A recent survey showed that the most common AI technologies are ChatGPT, Microsoft CoPilot, Character AI for text and code; Midjourney, Stablefusion, and Dalle3 for image generation; and Parker AI, Runway, and Google Gemini for multi-models (which can combine text, images, and video)[2].

Such technologies require significant financial outlays and technical development. According to the Economist, as an example, Elon Musk's

[1] High-Level Expert Group on Artificial Intelligence Draft Ethics Guidelines for Trustworthy AI, European Commission. 2018. Available at: https://www.euractiv.com/wp-content/uploads/sites/2/2018/12/ AIHLEGDraftAIEthicsGuidelinespdf.pdf (accessed: 10.04.2024)

[2] Available at: https://bristolcreativeindustries.com/(accessed: 10.04.2024)

start-up had raised $6bn. The investors, such Silicon Valley stalwarts as Sequoia Capital and Andreessen Horowitz, two venture-capital giants, and an investment fund with ties to the Saudi royal family put AI's financial firepower in the big league, alongside model-builders such as OpenAI, the creator of ChatGPT, and Anthropic (see Fig. 1)[3].

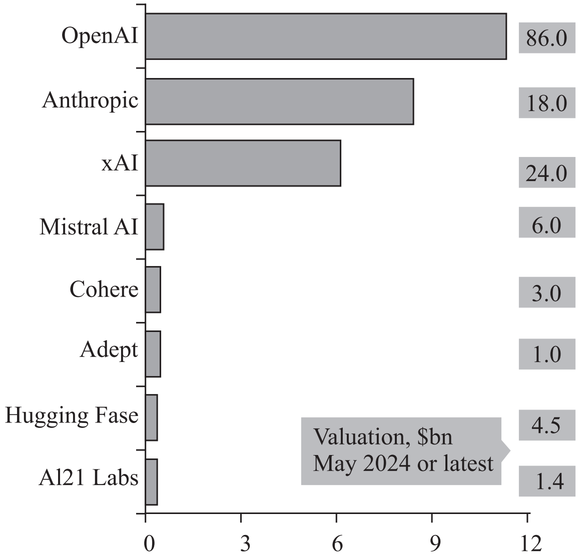


Fig. 1. The money isn't artificial. Al startups, cumulative capital raised, $bn

Source: The Economist. May 30th, 2024.

Moreover, the rumoured Apple-OpenAI deal represents a significant collaboration between two tech giants, promising to integrate OpenAI's advanced generative AI technology into Apple's software ecosystem. Apple is poised to enter the AI landscape in June 2024, and people think the announcement of the Apple OpenAI deal will be made that day alongside new iOS[4].

The fact the development and regulation of artificial intelligence is relevant is also evident on the international agenda. Thus, in November 2023 several countries, including the United States, China, the European

[3] Can Elon Musk's x AI take on Open AI? The Economist. May 29, 2024.Available at: https://www.economist.com/ business/2024/ 05/29/can-elon-musks-xai-take-on-openai (accessed: 11.04.2024)

[4] Available at: https://dataconomy.com/2024/05/31/chatgpt-apple-openai-deal/ (accessed: 11.04.2024)

Union, the United Kingdom, France, Italy, India, Brazil, Japan, Saudi Arabia, United Arab Emirates (Russia did not participate) held the first international summit and have approved a Declaration on the Artificial Intelligence Safety (The Bletchley Declaration on AI Safety)[5]. The declaration expresses a shared understanding of the opportunities and risks associated with artificial generative intelligence and states the urgent need to recognise and collectively manage the potential risks of AI through a new collaborative global effort to ensure the safe and responsible development and deployment of advanced AI. The participating countries agreed that significant risks could arise from potential intentional misuse or unintentional difficulties with control over advanced AI. Cyber security, biotechnology and disinformation risks are of particular concern in this connection. The Declaration notes the potential for serious, even catastrophic harm, intentional or unintentional, arising from the most significant capabilities of AI technologies and models. Among the main risks that the Declaration highlights are bias and breach of confidentiality in the application of AI.

The so-called Hiroshima Process organized by a number of Western countries was another important international event in the world of AI in recent times. On 30 October 2023 in Hiroshima, Japan, the G7 group of countries has approved a joint G7 Leaders' Statement on the Hiroshima AI Process, which proclaimed the International Guiding Principles on Artificial Intelligence and recommended a Code of Conduct for AI developers containing a set of rules AI developers are encouraged to follow on a voluntary basis to mitigate risks throughout the AI lifecycle.

By signing the Declaration, the parties have agreed that the risks posed by AI are inherently international and can be best addressed through international co-operation. The signatories agreed to co-operate in an inclusive manner to ensure the creation of a human-centred, trustworthy and responsible artificial intelligence.

The International AI Safety Summit and Declaration mentioned focused on "Frontier Artificial Intelligence" (Frontier AI) — highly capable general-purpose AI models that can perform a wide variety of tasks and match or exceed the capabilities present in today's most advanced model." Frontier AI is a subset of AI focused on highly advanced general purpose AI models, including foundation models that may have capabilities equal to

[5] Available at: https://www.gov.uk/government/news/countries-agree-to-safe-and-responsible-development-of-frontier-ai-in-landmark-bletchley-declaration (accessed: 11.04.2024)

or greater than the most sophisticated modern systems (e.g., narrower than the scope of the EU AI Act). Today, the most advanced general-purpose language models for large languages are, e.g., OpenAI GPT-4 and Google PaLM 2.

It has been declared that advanced AI (Frontier AI) systems pose significant security risks, especially in areas such as cybersecurity and biotechnology. Concerns arise from the potential for misuse, control issues and increased risks such as misinformation. However, the crucial difference between narrow models and general purpose models is that the latter are often made available through "broad deployment" via sector-agnostic platforms such as APIs, chatbots or open sourcing, and as such "can be integrated into a large number of diverse downstream applications possibly including safety critical sectors."

The Bletchley Declaration on AI Safety is not legally binding and is more a symbol than a detailed roadmap. Yet, the conference participants have agreed that it is necessary to:

Identify the security risks associated with AI, develop a common, evidence-based understanding of those risks, and maintain that understanding as opportunities arise in the context of a broader global approach to understanding the influence of AI on society, and

Based on the risks identified, develop appropriate policies in their countries to ensure secure countering of such risks: increased transparency accompanied by the adoption by private companies of advanced AI capabilities, appropriate assessment indicators, security testing tools, and the development of appropriate public sector capacity and research.

The proposed Code of Conduct contains a non-exhaustive list of recommendations for entities developing the most advanced artificial intelligence systems. These entities must operate on the basis of risk assessment at all stages of the lifecycle, including the design, development, deployment, and use of advanced AI systems. The AI development process consists set of actions, namely:

(a) identify, assess and mitigate risks throughout the AI lifecycle;

(b) develop and implement an AI and risk management policy based on a risk-based approach;

(c) develop and implement robust mechanisms for authenticating content and its origin, including watermarks or other methods that allow users to identify content created by AI;

(d) prioritise the development of advanced AI systems to address the world's most important challenges, in particular the global climate agenda, health, education, and others;

(e) implement appropriate measures to protect intellectual property and personal data.

Along with these international acts, the Resolution on Artificial Intelligence "Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development"[6] (hereinafter "the Resolution") passed on 21 March 2024 by the UN General Assembly is of key significance. Supported by more than 120 member states, the Resolution aims to encourage countries to protect human rights, safeguard personal data and monitor AI for risks on a non-legally binding basis. Though the UN does not have the ability to pass laws or regulations regarding AI or its implementation, the UN Charter gives to the General Assembly the power to initiate studies and make recommendations to promote the development and codification of international law. The main purpose of the document is to ensure "safe, secure and trustworthy AI systems" on a global level. It encourages all 193 Member States and multi-stakeholders from all regions and countries (private sector, international and regional organizations, civil society, the media, academia and research institutions and technical communities and individuals) to develop and support regulatory and governance frameworks.

The Resolution claims improper or malicious design, development, deployment and use of artificial intelligence systems, e.g., without adequate safeguards or in a manner inconsistent with international law, pose risks that could hinder progress towards the achievement of the 2030 Agenda for Sustainable Development and its Sustainable Development Goals and undermine sustainable development in its three dimensions — economic, social and environmental; widen digital divides between and within countries; reinforce structural inequalities and biases; lead to discrimination; undermine information integrity and access to information; undercut the protection, promotion and enjoyment of human rights and fundamental freedoms, including the right not to be subject to unlawful or arbitrary interference with one's privacy; and increase the potential risk for accidents and compound threats from malicious actors.

[6] Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development. UN General Assembly. March 2024. Available at: https://ai.gov.ru/ (accessed: 11.04.2024)

At the same time, while the Resolution does not define “Artificial Intelligence,” it does set out provisions of secure and “trustworthy artificial intelligence systems” which refers to artificial intelligence systems in the non-military domain, whose life cycle includes the stages: pre-design, design, development, evaluation, testing, deployment, use, sale, procurement, operation and decommissioning. The systems are referred to as human-centric, reliable, explainable, ethical, inclusive, in full respect, promotion and protection of human rights and international law, privacy preserving, sustainable development oriented, and responsible. According to the Resolution, such AI systems have the potential to accelerate and enable progress towards the achievement of all 17 Sustainable Development Goals and sustainable development in its three dimensions — economic, social and environmental — in a balanced and integrated manner; promote digital transformation; promote peace; overcome digital divides between and within countries; and promote and protect the enjoyment of human rights and fundamental freedoms for all, while keeping the human person at the centre.

However, as the Resolution has no enforcement powers on its Member States, there are no regulators under the Resolution, nor does the Resolution stipulate how the Member States should regulate AI systems in their own jurisdictions. As the Resolution is not legally binding, it does not confer enforcement powers or give rise to any penalties for non-compliance.

At the same time, as the Foreign Policy Concept of the Russian Federation[7] notes, mankind is going through an era of revolutionary change. This is primarily due to a structural transformation of the world economy as a result of the transition to a new technological base through the introduction of AI, cutting-edge information and communication technology, energy technology, biotechnology and nanotechnology. Other reasons include the growth of national identity, cultural and civilizational diversity and other objective factors that accelerate the redistribution of development potential to new centres of economic growth and geopolitical influence, and contribute to the democratisation of international relations.

It seems reasonable to agree with the view that the advent of Generative AI marks a paradigm shift in the AI landscape, the complexity and emergent autonomy of AI models introduce challenges in predictability and legal compliance [Novelli C., Casolari F., 2024: 1–2].

[7] Decree of the President of the Russian Federation of 31 March 2023 No. 229 “On Approval of the Concept of the Foreign Policy of the Russian Federation” // Collection of Laws of the Russian Federation. 03 April 2023. No. 14. P. 2406.

However, as AI capabilities become more powerful, the growing use of AI systems, as analysts believe [Brundage M. et al., 2018: 5–6], could lead to changes in the threat landscape, which can be categorised as follows: the scalable application of AI systems to perform tasks previously performed by humans — as a result, we see an expansion of existing threats; new threats posed by evolving technologies and AI models; the increasing use of artificial intelligence systems for malicious purposes significantly expands the range of AI applications, types of threats and risks. Three areas of security for AI systems can be distinguished here:

Digital security. The use of AI offers the potential to significantly increase the efficiency of cyber-attacks, which will create new threats by exploiting human vulnerabilities in the form of phishing, speech and image synthesis (deep fakes) or data leakage.

Physical security. The use of unmanned aerial, surface and underwater vehicles and other automated systems (including autonomous weapon systems, microdrone swarms, etc.), as well as attacks on cyber-physical systems (in transportation and industry) or critical infrastructure.

Political security. The use of AI to collect and analyse data for targeted propaganda or manipulation of consciousness and public opinion by violating privacy or analysing and manipulating people's behaviour, attitudes and beliefs on the basis of available data.

It is noteworthy that the National Strategy of the Russian Federation for the Development of Artificial Intelligence for the period until 2030[8] proclaims that the goals of AI development are to ensure the growth of welfare and quality of life of the population, ensure national security and law and order, and achieve sustainable competitiveness of the Russian economy, which includes global leadership globally in the field of AI. According to the National Security Strategy of the Russian Federation,[9] in order to ensure and protect the national interests of Russia from external and internal threats, including unfriendly actions of foreign states, the Russian Federation should more efficiently use its achievements and competitive advantages with account for long-term global trends. In order to solve the tasks in

[8] Decree of the President of the Russian Federation of 10 October 2019 No. 490 On Development of Artificial Intelligence in the Russian Federation // Collection of Laws of the Russian Federation, 14 October 2019, No. 41. P. 5700.

[9] Decree of the President of the Russian Federation of 02 July 2021 No. 400 On the National Security Strategy of the Russian Federation // Collection of Laws of the Russian Federation, 05 July 2021, No. 27 (Part II). P. 5351.

the sphere of national security, AI is used as a tool to ensure information security based on the application of advanced technologies. This includes AI and quantum computing technologies as a means of upgrading industrial enterprises and infrastructure, digitalisation to improve labour productivity and boost development of Russia's scientific and technological base, nanotechnology, robotics, medical, biological, genetic engineering, information and communication, big data processing, energy, laser, additive, creation of new materials, cognitive, and nature-like technologies.

In this situation the importance of comprehensive research into the development of AI and its new paradigm, including legal issues of the application of AI technologies in the digital economy, increases [Naumov V.B. et al., 2023].

1. Modern Legal Aspects of Artificial Intelligence Technologies

The current understanding of artificial intelligence gains particular importance at this time. E.g., the OECD definition contained in the OECD AI Principles 2019 built on the conceptual view of AI detailed in paper "Artificial Intelligence: A Modern Approach" by S. Russell and P. Norvig [Russell S., Norvig H., 2009]. It reads: "An AI system is a machine-based system that can, for a given set of human defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy."

This is in line with the updated definition of AI given in the OECD Memorandum 2023[10], which was formulated with the aim to harmonise and provide legal certainty for universal application. The updated definition reads as follows: «an AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment».[11]

The text above is replaced with the following updated definition: An AI system is a machine-based system that can, for a given set of human-de-

[10] OECD. Explanatory memorandum on the updated OECD definition of an AI system. OECD Artificial Intelligence Papers. 2024. No. 8. Available at: https://doi.org/10.1787/623da898-en. (accessed: 11.04.2024)

[11] Explanatory memorandum on updated OECD definition of AI System. Paris, 2024. Available at: http://www.oecd.org/termsandconditions . (accessed: 10.04.2024)

fined explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical real or virtual environments. Different AI systems are designed to operate with varying in their levels of autonomy and adaptiveness after deployment.

The earliest example of generative AI is a much simpler model known as the Markov chain. The method was named in honour of Andrei Markov, a Russian mathematician who in 1906 introduced this statistical method for modelling the behaviour of random processes. In machine learning, Markov models have long been used to predict the subsequent word, similar to the autocomplete function in an email programme. In text prediction, the Markov model generates the next word in a sentence by looking at the previous word or several previous words. The current basic AI models underlying ChatGPT and similar systems work in much the same way as the Markov model. But ChatGPT is much bigger and more complex: it has billions of parameters and is trained on huge amounts of data, mostly publicly available content on the Internet. In this huge body of text, words and sentences appear in sequences with certain dependencies. This helps the AI model to understand how to break the text into statistical chunks that have some predictability. AI learns the patterns of such blocks of text using this knowledge to suggest a particular solution [Zewe A., 2023].

The concept of Artificial Intelligence or Artificial Intelligence Systems usually includes categories of methods such as machine learning, and knowledge-based approaches and applications such as computer vision, natural language processing, speech recognition, intelligent decision support systems, intelligent robotic systems, and the application of these tools in various domains. Artificial intelligence technologies are advancing at a rapid pace, and additional methods and applications may be created in the future.

Usually, Generative AI (“GAI”) uses neural networks and other algorithms to create, through machine learning, new data or content similar to the original data.

The Generative AI model refers to generative modelling that is instantiated with a machine learning architecture (deep neural network) and, therefore, can create new data samples based on learned patterns. A generative AI system encompasses the entire infrastructure, including the model, data processing, and user interface components. The model serves as the core component of the system, which facilitates interaction and application within a broader context. Deep neural networks are particularly well suited for the purpose of data generation, such as diffusion probabilistic models for text-to-image generation or the transformer architecture and (large) language models (LLMs) for text generation. Generative AI is a branch of AI that can create new content such as texts, images, or audio that increasingly often cannot be distinguished anymore from human craftsmanship [Feuerriegel S. et al., 2024: 112-113].

Large generative AI models that can model output in and across specific domains or specific data types in a comprehensive and versatile manner are oftentimes also called foundation models [Bommasani R. et al., 2021: 4-5].

Generative AI is of immense importance for various industries such as media, arts, entertainment, advertising, and education. That said, it may also pose certain threats due to copyright infringement, dissemination of false or discriminatory information, and loss of control over the content created. Generative AI will have significant economic implications across various industries and markets. Generative AI can increase efficiency and productivity by automating many tasks that were previously performed by humans, such as content creation, customer service, code generation, etc. This can reduce costs and open up new opportunities for growth and innovation [Eloundou T. et al., 2023: 5].

Unlike GAI, descriptive AI based on machine learning is used to analyse, classify and make predictions from raw data, and to identify the data structure, dependencies and trends without creating new data. Descriptive AI can be used for various purposes such as: (a) classification, i.e., dividing data into groups based on their characteristics or attributes (classification of electrocardiograms into normal and abnormal, diagnosis of diseases, etc.); (b) regression, i.e. predicting unknown values based on known data (weather forecast, stock quotes, etc.); (c) clustering, i.e. dividing data into groups based on similarities between elements (business process modelling, etc.); (d) trend analysis, i.e. identifying trends and dependencies in data to provide information about future events or changes. Descriptive AI is the basis for many modern technologies such as recommender systems, automatic sound and image processing systems, quality control systems, and risk management systems. Although descriptive AI does not generate new data, it can provide important information and knowledge that can be used for decision making, planning and strategic planning.

Particular attention is paid to the definition of a conceptual approach to trusted artificial intelligence. In particular, the OECD[12] documents

[12] OECD 2023. Recommendation of the Council on Artificial Intelligence. OECD/LEGAL/0449. Available at: https://legalinstruments.oecd.org/en/instruments/OECD-

outline the principles of responsible governance of trustworthy AI, which complement each other and should be considered as a whole. These include, inter alia:

inclusive growth, sustainable development and well-being that involve engaging in responsible governance of trustworthy artificial intelligence to enhance human capabilities and creativity, promote inclusion, reduce economic, social, gender and other inequalities, and protect the environment, thereby promoting inclusive growth, sustainable development and well-being;

respect for the rule of law and human rights (freedom, dignity and autonomy, privacy and data protection, non-discrimination and equality, diversity, integrity, social justice and internationally recognised labour rights) and, to this end, the implementation of appropriate mechanisms and safeguards that are relevant to the context and in line with the state of the art;

transparency and lucidity, i.e., responsible disclosure of meaningful information about AI systems that is relevant to the context and consistent with the prior art;

reliability and security of the AI throughout its life cycle so that it functions properly and does not pose an unreasonable risk to safety under conditions of normal use, foreseeable use or misuse, or other adverse conditions;

accountability: AI agents should be responsible for the proper functioning of AI systems and for complying with the above principles based on their roles, context, and in accordance with the prior art.

At the same time, a new kind of AI self-developing artificial intelligence has already been developed. According to researchers at the Massachusetts Institute of Technology and the University of California (Fox News)[13], AI subsystems can be created without human assistance. Larger AI models like those used by ChatGPT can build on the “parent” algorithm to create smaller, specific AI applications that can be used, for example, to improve hearing aids, control oil pipelines, or monitor endangered wildlife.

But artificial AI technology continues to improve, and we see agentic AI models emerge. In comparison with General AI, a new model of AI,

LEGAL-0449; OECD.2020. Digitalization and Responsible Business Conduct: Stocktaking of policies and initiatives. Available at: https://www.oecd.org/daf/inv/mne/ publicationsdocuments/ reports/2/ (accessed: 11.04.2024)

[13] Available at: https://vfokuse.mail.ru/article/uchenye-zayavili-o-vozmozhnosti-ii-vosproizvoditsya-bez-uchastiya-cheloveka-59040575/ (accessed: 11.04.2024)

agentic AI is a more flexible system that could enable increased automation and worker productivity in certain types of industries and assist those who lack digital literacy. Large Action Model (LAM) adopts a learning-by-demonstration approach, observing human interactions with interfaces and replicating these actions reliably. AI systems that understand digital interfaces typically designed for humans and learn to execute human actions autonomously within these digital environments. AI agent might be able to interact with apps or websites, add items to a shopping cart and check out in accordance with pre-registered preferences and payment options, fill out and submit a form, or RSVP to an event. As an example, the recently released Humane AI Pin is attached to the user's shirt and acts as an AI-based digital assistant that responds to touch and voice and shows a laser projection on the user's palm; various smartphones and other hi-tech equipment are now equipped with an AI assistant [Pathirannehelage H. et al., 2022: 2]; [Aggarwal R., Singh H., 2024: 3].

Among other things, artificial intelligence, offering innovative solutions and analytical insights, has a great potential to shape the sustainable development model, revolutionise environmental and social processes, and scale the ESG model of corporate governance. There are many ways to realise the potential of AI to advance ESG-based sustainable development, offer innovative solutions to complex economic and governance challenges, and apply socially responsible practices. AI helps in developing strategies and planning scenarios for risk assessment and mitigation and customised risk management solutions tailored to specific industries and their unique challenges, including ESG risk mitigation. AI's ability to process complex data, predict trends and offer useful analytics is key to improving productivity and creating new business models for corporate governance. By harnessing the power of AI, companies can not only comply with regulations, but also introduce innovations, be competitive and comply with ethical business practices. However, AI should be seen as a complement to, not a replacement for, humans in their decision-making process. Successful integration of AI into management practices depends on a synergistic interrelation between technology and human understanding where AI acts as an enabler or catalyst for more informed, ethical and sustainable business decisions.

Generative Artificial Intelligence (GAI) has experienced dramatic growth recently and is accompanied, among other things, by growing challenges to the protectability of AI results in the intellectual property realm. Currently, a legal regime of artificial intelligence authorship and patent protection for AI technologies is being actively developed in various countries [Ivliev G.P., Egorova M.A., 2022: 32–46]; [Tikhomirov Yu.A. et al., 2019]; [Rozhkova M., 2021: 14–22]; [Morhat P., 2018: 1–8]; [Kharitonova Y., Savina V., 2020: 524–549].

The rapid development of artificial intelligence, and generative AI in particular, has created a whole maze of new copyright issues. These questions are primarily related to the way in which AI models are trained and whether the results of the development of these models constitute independently protectable subject matter so that they would be eligible for copyright protection.

The main question is whether works created by AI possess enough creativity to qualify for copyright protection. There is an opinion that works that are created with textual prompts and do not require any additional creative input from a human user, as in the case of generative AI tools, are not protected by copyright because these prompts are more like instructions for the commissioned artist.

The judicial practice in this area is not yet extensive, but is also interesting. E.g., in 2023, the court in Washington in *THALER vs. US Copyright Office* has ruled that only works with human authors can receive copyrights as human authorship is a "bedrock requirement of copyright" based on "centuries of settled understanding." According to the judgement copyright has never stretched far enough "to protect works generated by new forms of technology operating absent any guiding human hand, as plaintiff urges here. Human authorship is a bedrock requirement of copyright." The USA Copyright Act of 1909 explicitly provided that only a 'person' could 'secure copyright for his work' under the Act. Similarly, 9th Circuit appeals court ruling in 2018 that a monkey who took a selfie "could not sue under the Copyright Act for the alleged infringement of photographs this monkey had taken of himself, for 'all animals, since they are not human' lacked statutory standing under the Act."[14] Thaler was not able to point to any case "in which a court has recognized copyright in a work originating with a non-human".[15]

Likewise, in India it was decided that a work must involve a minimum degree of creativity and not be a product of only skill and labour. There-

[14] U.S. Copyright Office. Compendium of U.S. Copyright Office Practices § 101. 2021. Available at: https://copyright.gov/ comp3/ (accessed: 09.04.2024)

[15] Available at: https://www.digitalmusicnews.com/wp-content/uploads/2023/08/thaler-perlmutter-copyright-generative-AI-aug-2023.pdf (accessed: 11.04.2024)

fore, output produced by AI may not satisfy the requirement of ‘‘creativity’’ required for copyright protection, if viewed as a collection of data compiled from already existing sources without any infusion of creativity. In this sense, Indian and US copyright law agree that a certain class of AI-generated works would not qualify for copyright.

Interestingly, the Beijing Internet Court’s decision in *Li v. Liu* (China) makes a distinction between ‘‘straightforward’’ AI generated output where the human author simply takes and uses the output “as is” without any creative involvement and AI generated output where the human author keeps experimenting and adding various prompts, including negative prompts and tech parameters, until they receive a satisfactory result. In the later scenario, the Beijing court determined that such “AI-assisted work” (meaning output where aesthetic choices were exercised and there was personal judgement in the final rendition) would be eligible for copyright protection [Hill M., Hackworth A., 2023].

Another problem of artificial intelligence machine learning is related to algorithmic fairness that aims to address and rectify biases often embedded in machine learning systems. These biases can lead to discrimination in automated decision-making processes. Certain principles such as transparency, explainability and accountability are fundamental to developing artificial intelligence applications if the aim is to turn existing risks of discrimination into an opportunity for increased equality and these principles are respected along the entire algorithmic design chain [Xenidis R., Senden L., 2020: 160].

At the same time, in the process of building AI technologies, developers train the models by providing a huge amount of content to improve the model’s predictive abilities. But much of this content is copyrighted, and training a model on copyrighted material is itself a copyright infringement, even if the model does not reproduce the exact text as part of its output.

GAI raises copyright infringement concerns in several ways. Firstly, there is the problem of content created by artificial intelligence, or GAI itself, which possibly violates copyright on licensed use. Granting copyright to works created by AI has been widely debated because copyright laws traditionally protect only human-created works. Some experts believe that the content created by AI lacks human creativity and therefore does not fulfil the criteria of copyright. According to another viewpoint, copyright can be granted for GAI model creators generating such content. Another problem arises from the use of copyrighted data to train GAI models (so-called training data). The information sources that AI models use for training are copyrighted: text, images, and music. Arguments in defence of this practice are that using copyrighted data to train GAI models is fair use, while others argue that it constitutes infringement [Ivliev G., Egorova M., 2022: 46]; [Kirsanova E., 2023: 36-46].

The creation of new content and branding by AI based on compiled datasets or data stores, including visual elements such as logos, illustrations and textual elements such as image tags are assessed for legitimacy of using this data in new content. Previously, AI developers and vendors have disclaimed liability for any disruption resulting from their AI-based platforms. The key issue here is to determine who is liable for the content created by artificial intelligence: the AI user or the AI owner (provider). Generative AI companies usually publish disclaimers for the results of their AI platforms.

Recently, however, there has been a positive development: large AI vendors in some cases allow liability in the form of compensation for AI-generated content. But even those companies that have begun to offer compensation limit the protection by granting such rights generally to high-paying subscription tiers to the relevant AI applications. Amid growing scepticism about the use of AI, key industry players have formulated policies to ensure copyright protection for their users. E.g., Microsoft has introduced the CoPilot Copyright Commitment[16] where the company assumes liability for potential consequences arising from Microsoft's use of AI, the services of the second pilot and their outcomes. In addition, Microsoft commits to protecting its users from any third-party claims arising from such use.

Legal protection and defence of developments and technical patentable results created using GAI models is a problem in its own right. E.g., at the EPO, inventions involving AI are considered "computer-implemented inventions." Computer-implemented inventions are treated differently by patent offices in different jurisdictions, and in Europe, computer programs "as such" are excluded from patent protection. Nonetheless, software-related inventions remain eligible for patentability provided they exhibit a discernible technical character.

Over the years, the European case law has established a stable and predictable framework for the patentability of computer-implemented inventions, including inventions related to AI [Voller K., 2024].

[16] Available at: https://www.microsoft.com/en-us/licensing/news/microsoft-copilot-copyright-commitment (accessed: 12.04.2024)

An example is the case of Designation of inventor/DABUS (case J 0008/20) concerned two patent applications filed at the EPO (namely EP18275163 and EP18275174) where the applicant, Stephen Thaler, the inventor was noted to be "DABUS" — an AI created by Thaler himself. The EPO had rejected both applications on the grounds that the designated inventor, DABUS, did not meet the requirements for an inventor, that being a need for them to be a 'natural person'. Thaler subsequently appealed both decisions to the Board of Appeal with the opposition — whether an AI can be an inventor of a patent. The Board firmly rejected this point, as "under the European Patent Convention the designated inventor has to be a person with legal capacity". Further, Article 61 of the EPC notes that "[t]he right to a European patent shall belong to the inventor or his successor in title" (the latter being a legal successor in the title of the rights), and the rights of any employee, if they are the inventor, will be determined by the national legislation where they are employed. The Board clearly set out that "designating a machine without legal capacity can serve neither of these purposes" [17].

The UK Supreme Court has also firmly rejected the idea that a machine with AI can be recognised as an inventor under the UK Patents Act 1977. Addressing the ownership of inventions generated by DABUS, the court concluded that Dr. Thaler failed to establish a legal basis for claiming patent rights based on his ownership of the AI machine. It affirmed that Dr. Thaler had no independent right to obtain a patent for technical advances made by DABUS. The court judgement stipulated that, "it is not and has never been Dr. Thaler's case that he was the inventor and used DABUS as a highly sophisticated tool. Had he done so, the outcome of these proceedings might well have been different." The ownership of AI generated inventions is thus likely not an issue, provided a human inventor is identified, per the formal requirements[18].

Usually, the application of AI (including machine learning ("ML") and specific technical implementations of AI can be patented in Europe. However, fundamental algorithmic or mathematical level AI innovations typically fall outside the scope of patentability.

[17] The EPO Are Not 'Board' of AI Yet — EPO Board of Appeal Weighs in on Whether Artificial Intelligence Can Be an Inventor. 2022. Available at: https://www.ipiustitia.com/2022/08/the-epo-are-not-board-of-ai-yet-epo.html (accessed: 11.04.2024)

[18] The UK Supreme Court Judgement, December 20, 2023, Thaler vs Comptroller-General of Patents, Designs and Trademarks UK. Available at: https://www.supremecourt.uk/cases/uksc-2021-0201.html (accessed: 11.04.2024)

For AI to qualify for patent protection it must leave the abstract realm. This can be achieved in two ways. Firstly, the AI serves a technical purpose by addressing a technical challenge within a particular technology field, demonstrating its application in solving a specific technical problem. Secondly, the invention is directed to a specific technical implementation of AI motivated by technical considerations of the internal functioning of a computer, for example a specific technical implementation of neural networks by GPUs.

Generally, AI inventions are sensitive to the choice of network architecture, input representation, and training data. Since a specific technical purpose or implementation of the AI must be demonstrated, fundamental AI/ML improvements are generally not patentable. General purpose AI or generic AI with algorithmic efficiency are also not patentable.

The leading countries in the field of AI development relying on the active government support are rapidly developing national AI technologies. After the development of Deep Mind and the launch of the US-based Open AI ChatGPT in November 2022, public launches of similar LLM-based technologies in other countries followed. In November 2023, a government-backed AI company AI71 was launched in Abu Dhabi, UAE, to commercialise the LLM Falcon AI model. In December of the same year, the massive funding for the French AI Mistral was announced. India has been developing the models LLM Krutrim and Sarvam. States and private companies in the US, China, UK, France, Germany, India, Saudi Arabia and the UAE have massively funded AI development and expanded national production of graphics processing units and other elements necessary for AI development[19]. Russia has worked in a similar area and has certain achievements in neural networks, e.g., SBER (RuGPT-3).

These days, artificial intelligence finds more and more applications. E.g., AI arbitration is a relatively new concept that involves the use of artificial intelligence (AI) in the process of resolving disputes that exploits algorithms to analyse data related to the dispute and make recommendations on how it should be resolved. The use of AI can help to speed up the process

[19] Welcome to the era of AI nationalism. The Economist. January 1, 2024. Available at: https://www.economist.com/ business/2024/01/01/welcome-to-the-era-of-ai-nationalism?utm_content=article-link-2&etear=nl_ today_2&utm_ campaign=a.the-economist-today&utm_medium=email.internal-newsletter.np&utm_source=-salesforce-marketing-cloud&utm_term=1/1/2024&utm_id=1840347 (accessed: 01.04.2024)

of resolving disputes therefore, as the algorithms can analyse large amounts of data quickly and make recommendations in a timely manner that can be done through the use of smart contracts wherein the terms of the agreement and dispute resolution written directly into lines of code. However, there are also potential challenges to using AI in arbitration. One concern is that the algorithms may not be able to fully account for all of the nuances and complexities that can arise in legal situations. Additionally, there may be legal and regulatory issues that need to be addressed before AI arbitration can be widely adopted. For example, there may be concerns about the accountability and transparency of the algorithms used, and how breaches or damages would be handled. AI, which refers to the ability of machines to perform tasks that would normally require human intelligence, can be used to analyse data, make decisions, and optimise processes, as well as, secure a wide range of transactions, including those related to supply chain management, financial instruments, and identity verification.

The problem of AI risks and threats in the field of cyber security takes a special place.

2. Legal Aspects of Current Regulation of Artificial Intelligence

2.1. Legal AI Regulating in Russia

In Russia, the National Strategy for the Development of Artificial Intelligence for the period until 2030 (the “Strategy”), approved in 2019 and substantially extended in 2024, stipulates the following goals of AI development: ensure the growth of welfare and quality of life of the country’s population; ensure national security, law and order; achieve sustainable competitiveness of the Russian economy, including its leading positions in the world in the AI area.

The concept of artificial intelligence has been clarified in the new version of the Strategy, where AI is defined as a set of technical solutions that allow imitating human cognitive functions (including search for solutions without a predetermined algorithm) and obtaining results comparable to or exceeding the results of human intellectual activity when performing specific tasks. The set of technical solutions includes information and communication infrastructure, software (including software that uses machine learning methods), and processes and services for data processing and solution search. The Strategy defines the artificial intelligence model. It is a computer programme (a component of such a programme), which is designed to perform intellectual tasks at a level comparable to or exceeding the results of human intellectual activities and uses algorithms and data sets to deduce patterns, make decisions or predict results.

The Strategy contains new concepts, including:

large generative models of AI that are capable of interpreting (providing information based on queries, e.g., about objects in an image or about a text) and creating multimodal data (texts, images, videos and the like) at a level comparable to or superior to the results of human intellectual activity;

large fundamental models, i.e., AI models that (1) are the basis for creating and refining various types of software, (2) have been trained to recognise certain types of patterns, (3) contain at least 1 billion parameters, and (4) are used to perform a large number of different tasks;

promising AI methods, i.e. methods aimed at creating fundamentally new scientific and technical products, including the development of universal (strong) AI (ability to solve various problems independently, automatic design of physical objects, automatic machine learning, algorithms for solving problems based on data with partial partitioning and (or) insignificant amounts of data, information processing based on new types of computing systems, interpreted data processing, and other methods);

trustworthy AI technologies that meet safety standards, are developed with due regard for the principles of objectivity, non-discrimination, ethics, and rule out any possibility of harm to human beings and violation of their fundamental rights and freedoms, or damage to the interests of society and the state.

The Strategy notes that artificial intelligence is one of the most important technologies available to man today: thanks to AI, the world economy is growing already, innovation in all fields of science is accelerating, the quality of life of the population, availability and quality of medical care, quality of education, labour productivity and quality of recreation are improving. AI technologies are an area of international competition. Technological leadership in AI can enable states to attain meaningful results in key areas of social and economic development. In the late 2010s governments in developed countries began to focus on the development of AI technologies. To date, more than 60 countries have developed and approved their own national strategies for the development of artificial intelligence.

As the new version of the Strategy states, between 2022 and 2023, the world saw a new leap in the development of AI technologies owing to the improvement of large generative models in the fields of language, images (including video images) and sound. Large fundamental models are already capable of writing software codes according to technical tasks, composing poems on a given topic, giving precise and clear answers to test questions of various levels of complexity, including those from educational programmes. AI models create images on any topic in a matter of seconds based on a given text description or sketch. This poses a threat of the dissemination of prohibited information, copyright infringement and the generation of erroneous information.

AI will significantly impact the global economic growth. According to expert estimates, further development of large generative models can bring about a surge in labour productivity, which will lead to an annual increase in the global GDP by 1–2 percent and increase the remuneration of specialists in all sectors of the economy by increasing the volume of production (goods, works, services) and improving its quality.

At the same time, according to the National Security Strategy of the Russian Federation, in order to ensure and protect the country's national interests from external and internal threats, including unfriendly actions of foreign states, it is necessary to increase the efficiency of the use of the achievements and competitive advantages of the Russian Federation with account of long-term global trends. In order to solve the tasks set in the sphere of national security, AI is used as a tool to ensure information security based on the application of advanced technologies, including AI and quantum computing technologies as a means of upgrading industrial enterprises and infrastructure, digitalisation to improve labour productivity, and to boost the development of Russia's scientific and technological base, nanotechnology, robotics, medical, biological, genetic engineering, information and communication, big data processing, energy, laser, additive, creation of new materials, cognitive, and nature-like technologies.

At the same time, the Russian Federation Concept for the Development of Regulation of Relations in the Field of Artificial Intelligence and Robotics of 2020 (hereinafter the Concept[20]) developed in order to determine the main approaches to the transformation of the regulatory system in the Russian Federation so as to create conditions for creation and application of such technologies in various spheres of the economy while respecting the rights of citizens and ensuring the safety of individuals, society and the state,

[20] Collection of Laws of the Russian Federation. 31 August 2020. No. 35. P. 5593.

proceeds from the premise that the development of AI and robotics requires the creation of a regulatory environment comfortable for safe development and implementation of these technologies, based on a balance of interests of the individual, society, the state, companies developing AI and robotics systems, as well as consumers of their goods, works and services.

The Concept refers in Para 5 to technologies based on the use of AI: computer vision; natural language processing; speech recognition and synthesis; intelligent decision-making support; promising methods of AI. Promising AI methods include: ability to solve various problems independently, automatic design of physical objects, automatic machine learning, algorithms for solving problems based on data with partial partitioning and (or) insignificant amounts of data, information processing based on new types of computing systems, interpreted data processing, and other methods).

The Concept notes that the growing degree of AI and robotics systems autonomy, decreasing human control over the process of their application, and a not fully transparent decision-making process create a public demand for regulatory restrictions on the use of AI and robotics systems. At present, there are no unified approaches to regulating artificial intelligence and robotics technologies worldwide. This is due to the existence of a number of problems that have no clear solution.

The Concept outlines the Russian legal model for AI regulation in accordance with the National Strategy for the Development of Artificial Intelligence for the period up to 2030. It stipulates the following main areas for the creation of a comprehensive system for regulation of public relations arising in connection with the development and implementation of AI technologies:

ensuring a favourable legal environment (including the establishment of a pilot legal regime) for access to predominantly anonymised data, including data collected by public authorities and health care providers;

ensuring special conditions (regimes) for access to data, including personal data, for the purposes of academic research, creation of AI technologies and development of technological solutions based thereon;

creating legal conditions and establishing procedures for simplified testing and implementation of technological solutions developed on the basis of AI, as well as delegating to AI-powered information systems the ability to make certain decisions (except for decisions that may infringe upon the rights and legitimate interests of citizens). This includes the performance of state functions by state bodies (except for functions aimed at ensuring the security of the population and the state);

eliminating administrative barriers to the export of civilian products (works, services) created with AI;

creating unified systems of standardisation and conformity assessment of solutions developed on the AI basis, developing Russian Federation's international cooperation on standardisation issues and ensuring the possibility of certification of products (works, services) created on the AI basis;

encouraging investments by improving mechanisms for joint participation of investors and the state in projects related to the development of AI technologies, and providing targeted financial support to entities engaged in the development and implementation of AI technologies (provided that the introduction of such technologies will result in significant positive effects for the Russian economy);

developing ethical rules for human interaction with artificial intelligence.

The above areas must become the main landmarks in establishing a comprehensive system for regulation of public relations arising in connection with the development and implementation of AI technologies and robotics.

The Concept stipulates that, given the economic and social significance of AI and robotics technologies in various fields, their development and operation should not be confined to regulatory measures (except in cases involving a high risk of harm to human life and health). It is also unacceptable to use AI and robotics that pose a clear threat to the defence of the country and the state security.

For developing particular regulatory solutions it is necessary to use a risk-based approach based on an assessment of the amount of potential harm to these values, taking into account the likelihood of risk compared to the potential positive effect of the introduction of AI and robotics technologies, and the need to take measures to minimise the relevant risks.

The mere fact that AI systems and robotics are used should not be a basis for regulatory restrictions.

It is necessary to support the development of regulation developed and enforced by market participants (self-regulation), including the adoption and use of documents of the national standardisation system, ethical codes (sets of ethical rules) and other documents of self-regulatory organisations, as well as other instruments.

In view of the fundamental complexity of this sphere of legal relations, the development of a regulatory regime for artificial intelligence and robotics technologies requires the active involvement of representatives of corporate developers of AI and robotics systems and R&D organisations in the process of expert elaboration of the relevant laws and regulations.

In the future, some norms of law may also need to be clarified in order to provide normative legal regulation of new types of legal relations.

2.2. Legal AI Regulating in Europe

In March 2024, the European Parliament has passed a law regulating artificial intelligence that will come into force in June 2024 (EU Artificial Intelligence Act).[21] It applies to AI technology providers and users of AI-based technologies in the private and public sectors. The purpose of the Act is to improve the functioning of the internal market and the functioning of the internal market and promote the uptake of human-centric and trustworthy artificial intelligence, while ensuring a high level of protection of health, safety, fundamental rights, including democracy, the rule of law and environmental protection, against the harmful effects of AI systems in the Union and supporting innovation.

As with other EU data-related legislation, the Act also applies extraterritorially to companies and organisations outside the EU. The AI Act applies to:

providers placing on the market or putting into service AI systems or placing on the market general-purpose AI models in the EU, irrespective of their location or establishment;

deployers of AI systems that have their place of establishment or are located within the EU;

providers and deployers of AI systems that have their place of establishment or are located outside the EU, where the output produced by the AI system is used in the EU;

importers and distributors of AI systems;

product manufacturers placing on the market or putting into service an AI system together with their product and under their own name or trademark;

authorized representatives of providers which are not established in the EU.

[21] Available at: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai (accessed: 01.04.2024)

The AI Act establishes a legal framework for the application of AI based on the assessment of risks (as the combination of the probability of the occurrence of harm and the severity of that harm) associated with the use and placing on the market the following categories of artificial intelligence systems: prohibited artificial intelligence practices, high-risk artificial intelligence systems, systems with transparency requirements, and general purpose artificial intelligence models.

In the Act, an AI system' means a "machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments» (Article 3). In addition to many other important definitions, the Law also contains a definition of "deep fake" that means AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful".

According to the definition, a key characteristic that distinguishes "AI systems" from traditional software is that an AI system derives conclusions for the output from the input ("infers, from the input it receives, how to generate outputs"). This is intended to emphasise the ability of AI systems to derive models and/or algorithms from input data. By contrast, the EU wanted to exclude systems that are based on rules that are defined exclusively by natural persons in order to carry out automatic processes from the scope of the AI Act. By definition, the capabilities of AI systems should go beyond basic data processing operations and be understood more as learning, reasoning or modelling.

The definition in the AI Act also assumes that AI systems are "designed to operate with varying levels of autonomy". Accordingly, there must be a certain degree of independence of the system's actions from humans. In other words, the system must be able to operate without human intervention.

The characteristic of "adaptiveness" is intended to express the ability of an AI system to (continue to) learn itself and thus constantly change.

Prohibited Artificial Intelligence Practices are:

the placing on the market, the putting into service or the use of an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective, or the effect of materially distorting the behaviour of a person or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause that person, another person or group of persons significant harm;

the placing on the market, the putting into service or the use of an AI system that exploits any of the vulnerabilities of a natural person or a specific group of persons due to their age, disability or a specific social or economic situation, with the objective, or the effect, of materially distorting the behaviour of that person or a person belonging to that group in a manner that causes or is reasonably likely to cause that person or another person significant harm;

the placing on the market, the putting into service or the use of AI systems for the evaluation or classification of natural persons or groups of persons over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics, with the social score leading to either or both of the following:

detrimental or unfavourable treatment of certain natural persons or groups of persons in social contexts that are unrelated to the contexts in which the data was originally generated or collected;

detrimental or unfavourable treatment of certain natural persons or groups of persons that is unjustified or disproportionate to their social behaviour or its gravity;

the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement, unless and in so far as such use is strictly necessary for one of the following objectives;

AI systems to infer emotions of a natural person in the areas of workplace and education institutions, except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons;

the use of biometric categorisation systems that categorise individually natural persons based on their biometric data to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation; this prohibition does not cover any labelling or filtering of lawfully acquired biometric datasets, such as images, based on biometric data or categorizing of biometric data in the area of law enforcement.

The Act identifies high-risk AI systems that pose a potentially high risk to human rights and freedoms and differentiates them into two high-risk AI groups. The first group includes AI systems that pose a risk when the AI system is intended to be used as a safety component of a product, or the AI system is itself a product, covered by the Union harmonisation legislation (e.g., Regulation (EU) 2017/745 on medical devices or Regulation (EU) No 167/2013 on agricultural and forestry vehicles); the product whose safety component pursuant to point (a) is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment, with a view to the placing on the market or the putting into service of that product pursuant to the Union harmonisation legislation.

It may apply to AI systems used in, among other things, cars, toys, lifts, equipment and safety components for use in medical devices and *in vitro* diagnostic medical devices, products related to civil aviation, marine equipment, products related to railway systems, and various types of vehicles.

General-purpose artificial intelligence models represent an independent type of AI systems. According to the EU AI Law, a general-purpose AI model is trained on large amounts of data using scalable self-monitoring that demonstrates significant generality, is capable of competently performing a wide range of individual tasks, and can be integrated into a variety of downstream systems or applications, including serving as the basis for general-purpose AI systems.

In addition, the AI Act also introduces a category of general purpose AI models with a systemic risk for more advanced general purpose AI models to be defined by the European Commission. General-purpose AI (GPAI) models with a systemic risk will be subject to additional obligations for model evaluation and testing, risk mitigation, security, and incident reporting.

GPAI models are subject to a range of obligations fostering technological deployment and ensuring adequate safeguards, including the provision of detailed technical documentation to the competent authorities, the provision of information to downstream providers, the implementation of policies to protect copyright and the publication of a summary of the content used for training the GPAI model. Providers that release GPAI models under a free and open-source licence are subject to certain exemptions of these obligations.

GPAI models are considered to have a systemic risk if they have high impact capabilities, e.g., if they have great computing power (currently when the computation used for its training is greater than 10^25 FLOPS and subject to future amendments by the Commission). Furthermore, GPAI models can be classified as having systemic risk in case of a decision of the

Commission (either ex officio or following a qualified alert from a scientific panel of independent experts). The provider of such GPAI model needs to:

perform model evaluation in accordance with standardised protocols;
conduct systemic risk assessments and mitigate systemic risks;
report incidents to authorities; and
ensure adequate cybersecurity protection, including the physical infrastructure of the model.

The EU AI Law follows a risk-based approach taking into account the risks of AI to natural persons. The AI Act therefore distinguishes between prohibited AI practices, high-risk AI systems, AI systems with transparency risk and GPAI models with/without a systemic risk. Before placing an AI system on the market, putting it into service, deploying, distributing, importing or otherwise using it, it must be carefully ruled out that such system does not entail an "unacceptable risk" within the meaning of the AI Act.

2.3. US Legal AI Regulating

On 30 October 2023, US President Biden has issued a new Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence.[22] It sets new standards for AI safety, provides a set of measures and directs government agencies to implement specific policies to address areas of concern in national security, data protection, labour relations and social health. The Order stipulates an obligation for companies developing the most powerful AI systems to report the results of AI safety tests and other important information to the US. government. Under the Defense Production Act, the Order requires developers of AI foundation models that potentially pose a serious threat to national security, national economic security, or national public health to notify the federal government when training an AI model about the results of all pen-tests (red-team) to assess the cyber security of the AI model before companies make those results public.

The Order includes more than a hundred policy directives related to AI security to more than twenty federal agencies, tasking them with policies to address problem areas such as national security, data protection, workplace bias, and public health. It also imposes obligations on private companies developing powerful AI systems that could pose a threat to national security

[22] Available at: https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/ (accessed: 11.04.2024)

or public health, requiring them to share safety test results and methods and other sensitive information with the U.S. government. Most of the directives issued by President Biden under this Order must be implemented within 2024.

In March 2023, the President has approved a new version of the National Cybersecurity Strategy[23] establishing protected US critical infrastructure has become one of the national security priorities. The initiative seeks to shift some of the burden of cyber security risk mitigation from end users and critical infrastructure operators to private sector enterprises that are best positioned to make meaningful progress on security and resilience. The Strategy also highlights the need to change incentives in favour of long-term private sector investment. The strategy is based on five pillars: protecting critical infrastructure; identification and destruction of threat actors; establishing market mechanisms to improve security and resilience; investing in a sustainable future; and building international partnerships to achieve common goals. Each pillar contains specific strategic objectives that build on previous programmes and guide the implementation efforts of government and private sector entities.

2.4. China AI Legal Regulating

China has achieved significant success in its efforts to become a technology superpower over the past few years, making continuous efforts to establish itself as the world's leading IP producer.

The country has transformed from a low-wage economy to a high-tech country. In fact, according to the World Intellectual Property Organisation (WIPO), China accounted for 47% of all patent applications worldwide in 2023. On 13 July 2023, the Chinese government has published regulations on generative artificial intelligence, Interim Measures for the Administration of Generative Artificial Intelligence Services (hereinafter Interim GAI Measures; Measures)[24] came into force on 15 August 2023. The Measures aim to regulate generative AI that is primarily intended for content creation. They are the latest addition to the emerging system of AI regulation in China, which already includes a number of AI-specific and local laws.

[23] Available at: https://www.whitehouse.gov/wp-content/uploads/2023/03/National-Cybersecurity-Strategy-2023.pdf (accessed: 01.04.2024)

[24] Available at: http://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm (accessed: 11.04.2024)

The Chinese government has been supporting its AI industry on a national level since the beginning. In its 13th Five-Year Plan (2016-2020), China identifies AI as key to achieving economic growth goals. In 2017, the Chinese government presented its vision for AI development in its Next-Generation AI Development Plan. The Plan presents Beijing's comprehensive strategy to focus AI on the country's socio-economic development efforts (AI industry), which will make China a global AI leader by the year 2030.

GAI Interim Measures differ from other laws in that they specifically regulate the use of generative AI defined as "models and related technologies that have the ability to generate content such as text, images, audio and video," so as to provide content generation services to the Chinese public. Compared to the provisions of Deep Synthesis, the generative AI covered by GAI Interim Measures encompasses more than algorithm-based generative technologies, and includes rules-based models and systems.

The Measures apply to generative AI service providers defined as legal entities and individuals that use generative AI to provide generative AI services, including the provision of such services through application programming interfaces (APIs). Also, GAI Interim Measures cover the provision of generative AI services to the public indirectly through business arrangements. On the other hand, institutions that develop and apply generative AI technology but do not provide generative AI services to the public do not fall under this regulation.

In addition, the GAI Interim Measures establish an extraterritorial scope by specifying that they apply to the provision of services to the public in the PRC mainland, potentially extending their application to individuals and organisations outside of China that provide generative artificial intelligence services to individuals in the PRC. This nuance is complemented by another provision stating that, if generative AI vendors outside the PRC fail to comply with the Measures and other laws, this will entail notification to the relevant agencies to take technical measures and other necessary measures to deal with the perpetrators.

Conclusion

Exponential improvements in artificial intelligence and other advanced technologies in recent years have led to a surge in interest (academic, commercial, military, etc.) and financial investment in artificial intelligence.

It is obvious that the rapid progress in the development and practical application of AI technologies is driven by their expected potential to increase productivity, encourage innovation and entrepreneurship, provide solutions to global problems, including social problems such as improving healthcare and helping to solve the climate crisis, as well as to achieve the Sustainable Development Goals. At the same time, this process also generates new threats and challenges for the human civilisation, which is a special factor that must be taken into account when promoting the development of artificial intelligence.

References

1. Aggarwal R., Singh H. (2024) Overcoming Limitations of Ai Agents: Integrating Tacit Knowledge Through Inferred Latent Themes. Available at: SSRN: https://ssrn.com/abstract=4843878 (accessed: 10.04.2024)

2. Amelin R.V., Channov S.E. (2023) *Evolution of Law under Influence of Digital Technologies.* Moscow: Norma, 280 p. (in Russ.)

3. Antonova N.V. et al. (2019) *The Legal Concept of Robotization.* Moscow: Prospekt, 240 pp. (in Russ.)

4. Barfield W., Pagallo U. (2020) Advanced Introduction to Law and Artificial Intelligence. *Law in Context*, vol. 37, no.1. Available at: https://books.google.ru/books?id=7MgBEAAAQBAJ&printsec=frontcover&hl=ru#v=onepage&q&f=false (accessed: 11.04.2024)

5. Bommasani R. et al. (2021) Opportunities and Risks of Foundation Models. Available at: https://doi.org/10.48550/arXiv.2108.07258 (accessed: 11.04.2024)

6. Brundage M. et al. (2018) The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. Available at: https://www.eff.org/files/2018/02/20/malicious_ai_report_final.pdf (accessed: 01.04.2024)

7. Cowgill B., Tucker C. (2020) Algorithmic Fairness and Economics. Columbia Business School Research Paper. Available at: SSRN: https://ssrn.com/abstract=3361280 (accessed: 11.04.2024)

8. Eloundou T., Manning S. et al. (2023) GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models. Available via license: Creative Commons Attribution-ShareAlike 4.0 International (accessed: 11.04.2024)

9. Feuerriegel S. et al. (2023) Generative AI. *Business & Information Systems Engineering,* vol. 66, no. 2, pp. 111–126.

10. Hill M., Hackworth A. (2023) Copyright in the Age of AI. Available at: https://www.charlesrussellspeechlys.com/en/insights/quick-reads/102j7d1-copyright-in-the-age-of-ai/#page=1 (accessed: 11.04.2024)

11. Ivliev G.P., Egorova M.A. (2022) On Legal Status of Artificial Intelligence and Products Created by Artificial Intelligence Systems. *Zhurnal rossiyskogo prava*=Journal of Russian Law, no. 6, p. 32–46 (in Russ.)

12. Kharitonova Y. S., Savina V. S. (2020) Artificial Intelligence Technology and Law: Challenges of Modernity. *Vestnik Permskogo universiteta*=Bulletin of Perm University, no. 49, pp. 524–549 (in Russ.)

13. Kirsanova E.E. (2023) Review of the Main Theories of Determining the Legal Regime of Objects Created by Artificial Intelligence. *Zakon*=Law, no. 9, pp. 36–46 (in Russ.)

14. Morhat P.M. (2018) Legal Personality of an Electronic Person. *Pravovye issledovania*=Legal Studies, no. 4, pp. 1–8 (in Russ.)

15. Naumov V.B. et al. (2021) Legal Aspects of Using Artificial Intelligence. Available at: https://www.hse.ru/mirror/pubs/share/480106412.pdf (accessed: 11.04.2024) (in Russ.)

16. Novelli C., Casolari F. et al. (2024) Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cyber security. Available at: SSRN: https://ssrn.com/abstract=4694565 (accessed: 10.04.2024)

17. Pathirannehelage H. et al. (2022) Design Principles for AI-augmented Decision-Making: an action design study. Available at: SSRN: https://ssrn.com/abstract=4071519 (accessed: 11.04.2024)

18. Rozhkova M.A. (2021) Will Artificial Intelligence Become an Independent Subject of Law? *Khozyaistvo i pravo*=Economy and Law, no. 6, pp. 14–22 (in Russ.)

19. Russell S., Norvig P. (2009) Artificial Intelligence: A Modern Approach. Available at: http://aima.cs.berkeley.edu/ (accessed: 11.04.2024)

20. Voller K. (2024) Generative AI — not so Great at Generating European patents. Available at: https://www.gje.com/resources/generative-ai-not-so-great-at-generating-european-patents/ (accessed: 11.04.2024)

21. Xenidis R., Senden L. (2020) EU Non-Discrimination Law in the Era of Artificial Intelligence: Mapping Challenges of Algorithmic Discrimination. In: U. Bernitz et al. (eds.) General Principles of EU Law and the EU Digital Order. Available at: SSRN: https://ssrn.com/abstract=3529524 (accessed: 10.04.2024)

22. Zewe A. (2023) Explained: Generative AI. Massachusetts Institute of Technology News. Available at: https://news.mit.edu/2023/explained-generative-ai-1109 (accessed: 11.04.2024)

Information about the author:

A.A. Kartskhiya — Doctor of Sciences (Law), Professor.

The article was submitted to editorial office 23.06.2024; approved after reviewing 28.06.2024; accepted for publication 28.06.2024.

Legal Issues in the Digital Age. 2024. Vol. 5. No. 2.
Вопросы права в цифровую эпоху. 2024. Том 5. № 2.

Research article
УДК*:* 340, 347
DOI:10.17323/2713-2749.2024.2.36.56

Progress in Natural Language Processing Technologies: Regulating Quality and Accessibility of Training Data

Ilya Gennadievich Ilyin

Saint Petersburg State University, 7/22 Liniya V.O., Saint Petersburg 199106, Russia, i.g.ilin@spbu.ru, orcid: https://orcid.org/0000-0003-1076-2765

Abstract

Progress in natural language processing technologies (NLP) is a cardinal factor of major socioeconomic importance behind innovative digital products. However, inadequate legal regulation of quality and accessibility of training data is a major obstacle to this technological development. The paper is focused on regulatory issues affecting the quality and accessibility of data needed for language model training. In analyzing the normative barriers and proposing ways to remove them, the author of the paper argues for the need to develop a comprehensive regulatory system designed to ensure sustainable development of the technology.

Keywords

personal data; data regime; generative neural network; artificial intelligence; natural language processing; large language models; data access; copyright.

Acknowledgments: the paper is published within the project of supporting the publications of the authors of Russian educational and research organizations in the Higher School of Economics academic publications.

For citation: Ilyin I.G. (2024) Progress in Natural Language Processing Technologies: Regulating Quality and Accessibility of Training Data. *Legal Issues in the Digital Age,* vol. 5, no. 2, pp. 36–56. DOI:10.17323/2713-2749.2024.2.36.56

Background

The technology of natural language processing (NLP) is associated with mathematical linguistics and artificial intelligence and allows computers to understand and generate natural language [Hirschberg J., Manning C.D., 2015: 261–266]. As applied to information technologies, language and speech help to promote the engagement between man and computer as exemplified by digital products for processing and analysis of texts (spelling, grammar, duplication, readability checking services, etc.), text translators, voice assistants and other interactive response technologies (chat bots, automated client support systems etc.).

Progress in natural language processing is crucial both from the economic perspective as a key factor for development of artificial intelligence [Feng Z., 2023: 7–8, 25] with a potential for innovative digital products, and also from the social perspective in view of the importance to develop and preserve the natural language as a major aspect of the national and cultural identity.

Meanwhile, despite the innovative nature and socioeconomic value of the technology under discussion, the existing legal framework cannot fully support its sustainable development, a key trouble being normative barriers for access to training data with qualitative and quantitative parameters needed to achieve progress.

From the technical perspective, the urgency of the problem follows from the methods of natural language processing. The technology relies on generative neural networks to create large language models (LLM) [Glauner P., 2024: 24–34]. These models are trained on large data arrays including those structured as a linguistic corpus — a database containing numerous texts (books, transcriptions, translations etc.) and audio files (audio books, broadcasting recordings, podcasts and other audio content) — something that allows them to study the structure of natural language and “understand” different language contexts.

Large language models assume the use of not only available data but also those generated by the neural network on their basis. Such approach, on the one hand, considerably expands the amount of training data but, on the other hand, makes it more difficult to correct algorithmic errors and defects. For instance, if training data contained defects that could affect the functioning of the algorithm, these defects would corrupt the data generated by the model. In this situation, removing corrupt data is technically difficult. One example of large language models is BERT[1], GPT-3[2] and the underlying digital products like Google Assistant or ChatGPT.

From the regulatory perspective, the issue has been identified in the relevant strategic planning documents, with the 2030 National AI Development Strategy[3] (hereinafter Strategy) as one of the key documents in the field. In the Strategy, normative barriers and a lack of methodological framework for support of AI systems with reliable data are referred to as obstacle for the development of artificial intelligence in Russia.[4] The Strategy calls to develop a comprehensive regulatory system for social relations related to the development and application of AI technologies[5], in particular, to remove excessive normative barriers and create an enabling regulatory environment for development and introduction of AI technologies[6], remove regulatory barriers for development and introduction of large generative models to be trained on large data arrays[7], and provide for regulatory support of AI developers' access to different types of data.[8]

[1] Generative Pre-trained Transformer (GPT) is a line of deep learning models developed by OpenAI (United States) and based on the Transformer architecture. Trained without a "trainer", it does not need to be adapted and can be used for a variety of tasks. For detail on GPT see: Yenduri G. et al. Generative pre-trained transformer: A comprehensive review on enabling technologies, potential applications, emerging challenges, and future directions // arXiv preprint arXiv:2305.10435. 2023. For detail on the Transformer architecture see: Vaswani A. et al. Attention is all you need //Advances in neural information processing systems. 2017. Vol. 30.

[2] Bidirectional Encoder Representations from Transformers (BERT) is a deep learning model designed by Alphabet Inc. (United States). Based of the Transformer architecture, it is trained on bidirectional context meaning an ability to analyze and understand contexts both from left to right and vice versa. For more detail on BERT see: Devlin J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding//arXiv preprint arXiv:1810.04805. 2018.

[3] The 2030 National Artificial Intelligence Development Strategy approved by Presidential Decree No. 490 "On the Development of Artificial Intelligence in Russia" of 10 October 2019 ("2030 National AI Development Strategy"). Here and elsewhere all references to documents, regulations, legal practice are taken from SPS Consultant Plus.

[4] Para 17(16) (e), 2030 National AI Development Strategy.

[5] Ibid. Para 24 (f).

[6] Ibid. Para 24 (f).

[7] Ibid. Para 51(11) (c).

[8] Ibid. Para 51(11) (b).

In view of the objectives set by the Strategy, the paper purports to provide a conceptual analysis of the problem to regulate the quality and accessibility of training data, and to identify and propose ways to address the underlying legal constraints.

In terms of its subject matter, the paper has three parts in addition to the background and conclusion. The first part explores the legal aspects related to the impact of data parameters on language models to be developed. The second part is focused on the existing legal arrangements that support the required data quality. The third part is devoted to the issues of accessibility of training data, analysis of normative barriers and discussion of the ways to remove them.

1. Data Parameters: Aspects of Impact on Language Models under Development

1.1. Data and Language Models: Interrelation and Technical Parameters

Progress in natural language processing technologies is largely hinged on the efficient language models developed for a particular language. These models are crucial for subsequent operation of available digital products and affect to what extent a computer is able to “understand” and process texts. A language model is created through a series of consecutive stages.

At first, training data are put together: this stage involves a large amount of textual and other language data from a wide range of sources. Training data for language models will normally include textual data (for instance, written texts, speech transcriptions and annotated lists), speech data (audio recordings, phonetic and prosodic annotations) and multimodal data (image-text, video-text and audio-text pairs) [Dash N.S., et al., 2018: 291].

Once collected, the data is pre-processed. This stage involves removal of noise (for instance, irrelevant information, errors, duplicates), text normalization (bringing to a common format), breaking a text into sentences and words, stop word removal, lemmatization (grouping together inflected forms) and stemming (stripping words down to their stems [Khyani D. et al., 2021: 350–357]. The purpose of pre-processing is to prepare data for mining and language model training [Goldberg Y., 2017: 65–76].

The next stage is training of the language model itself, with regularities, dependencies and peculiarities of the data in question identified through the use of machine and deep learning algorithms. Language models could be trained to address a number of tasks: text classification, tonality analysis, named entity recognition, machine translation, etc. [Zhou M. et al., 2020: 275–290]. After the training, language models are evaluated on text data to check for efficiency and accuracy. A language model can be fine tuned and optimized depending on the evaluation's results.

Finally, the introduction of a language model assumes its accessibility for integration into the respective digital products. This process will require ongoing monitoring of its functioning with changes and improvements to be made as necessary, for example, to take account of technological innovations and user feedback. Due to ongoing improvement of the model, the stage of introduction is time consuming.

1.2. Functional Errors of Language Models: Legal Defects and Quality Defects

Quality and diversity of training data will directly impact the ability of a language model to be trained and to interpret texts in a given natural language. The structure of data including their arrangement and format, representativeness, amount and other parameters will affect the training process and accuracy of understanding a text's semantics and context. The use of data below the required qualitative/quantitative parameters will hinder further progress of the technology, only to result in negative implications in both technical terms — algorithmic errors due to falsely identified correlations and regularities — and legal terms like illegitimate restriction of rights and liberties (algorithmic discrimination), violation of privacy, personal and family secrets, occurrence of harm etc.

Training data defects could be regarded from two perspectives: firstly, incompatibility with specific technical criteria and metrics (quality defects) such as those of representativeness, amount, purity etc.; secondly, violation of the applicable legal regime (legal defects) such as personal data protection when data are processed as part of a language model.

It has a sense first discuss in more detail the implications of training data quality defects. It should be noted above all that quality defects will not inevitably bring negative outcomes. For example, a minor inaccuracy, insufficiency, irrelevance of training data, while not having a major bearing on common dependencies to be identified, could impact the findings of data analysis with regard to specific individuals [Hacker P., 2021: 260, 263]. The set of required quality parameters and respective metrical values should apparently differ in technical terms depending on the purpose of a given language model and its area of application.

In general terms, the data quality defect as applied to the natural language processing technology could contribute to the digital divide [Lythreatis S. et. al, 2022: 1—11] and cause language discrimination.

Digital divide is a kind of social inequality identified as impossibility for individuals or social groups to have equal access to information and communication technologies, as well as equal level of skill to use them [Rogers S.E., 2016: 197—199]. The urgency to address this problem has been underlined at the national[9] and international level.[10] In terms of law, the problem of digital divide will primarily affect the relations of constitutional law, in particular, the legal status of individuals, human and civil rights and liberties guaranteed by the state [Mushakov V.E., 2022: 69—73] including equal civil and human rights and liberties irrespective of the language.

Digital divide can manifest itself as language discrimination resulting in limited access of specific social groups to a technology due to impossibility to use it in a native language (limited choice of supported languages) or incorrect functioning due to specific dialect and peculiarities of the language spoken by the social group in question.

Article 2 of the Universal Declaration of Human Rights (1948)[11] prohibits discrimination including on the basis of language. A similar provision is set by Article 1 (3) of the UN Charter[12] as also reflected in paragraph 2, Article 19 of the Russian Constitution[13] whereby the state guarantees equal civil and human rights and liberties irrespective of language.

[9] Federal Government Resolution No. 313 "On approving the Information Society public program of the Russian Federation" of 15 April 2014 // SPS Consultant Plus.

[10] United Nations Declaration of Principles Building the Information Society of 12 December 2003. Available at: https://www.un.org/ru/events/pastevents/pdf/dec_wsis.pdf (accessed: 19.04.2024); UN Tunis Agenda for the Information Society of 15 November 2005. Available at: https://www.un.org/ru/events/pastevents/pdf/agenda_wsis.pdf (accessed: 19.04.2024)

[11] Universal Declaration of Human Rights (passed by the UN General Assembly 10.12.1948). Available at: https://www.un.org/ru/documents/decl_conv/declarations/declhr.shtml (accessed: 10.06.2024)

[12] United Nations Charter (passed in San Francisco 26.06.1945). Available at: https://www.un.org/ru/about-us/un-charter/full-text (accessed: 10.06.2024)

[13] Constitution of Russia (approved by universal vote on 12.12.1993 as amended in the course of all-Russia popular vote on 01.07.2020).

Progress in natural language processing technologies adds up a new form of discrimination where it occurs through inadequate digitization of languages rather than someone's guilty action.

A language model to be developed will require access to training data in a given language. Meanwhile, digital data for development of robust and accurate language models are not available for all languages. For example, if the training data set was limited and did not cover all dialects of a language, the functioning of the language model may be incorrect or inaccurate or fail altogether when processing a natural language incorporating such dialects. The differences of pronunciation, vocabulary and grammar can result in defective text or speech recognition and analysis. Moreover, such problems will not arise for a language with a high level of digitization and therefore high representativeness. A similar issue is also observed in respect of minor languages. Thus, while digitization of specific major languages (like English, Russian) is high, many digital products are still not available to speakers of minor languages, for example, Udmurt, Buryat, Tuvan. For this reason, technical and legal support of access to the relevant linguistic corpuses is critical for digitization of the said languages and thus for development of the technology in question and elimination of digital divides.

Data quality can be undermined both for objective reasons (for instance, insufficient digitization level) and because of wishful action to corrupt training data and thus change the language model's training outcomes. In practice, such action is called data poisoning [Russo A., Proutiere A., 2021: 3234–3241]. False examples introduced into the training data set could result in wrong outcomes produced by the model like corrupt and incorrect translation of documents by automatic translation systems, only to affect the accuracy and meaning of the information to be transmitted. In chat bots, this can result in wrong answers to user queries to dump down user experience, undermine trust in the technology and bring about related legal implications, such as violation of consumer rights to quality products/services[14], right to information[15], etc. Errors in text analysis systems can result in wrong interpretation of text tonality or content, something especially critical in analysis of public opinion or monitoring of social networks and fraught with major implications including wrong legal qualification of one's

[14] Article 4 of Federal Law No. 2300-1 "On Protecting Consumers' Rights" of 07 February 1992 (hereinafter "Law on Protecting Consumers' Rights").

[15] Ibid. Article 8.

actions that could be wrongly qualified as incitement of hatred or humiliation of human dignity.[16]

Where natural language processing is used in critically important sectors such as medicine, the implications of data poisoning can be especially harmful and cause considerable damage, for example, through a wrong diagnosis due to wrong interpretation of medical data, thus jeopardizing human life and health.

Meanwhile, correcting technical errors and removing poor quality training data from language models will cause the issue of algorithmic shadow left by such data [Li T.C., 2022: 480–505]. In the general sense, this problem means that even removed data will still impact the created language models. Thus, for example, removing personal data from a training data set does not fully prevent their further influence on the language model: algorithmic shadow will be still observed in its operation. This is fraught with violating the data subject's rights and questions the operational legitimacy of such model as a whole.

Algorithmic destruction — elimination of data through special algorithms — is among technological solutions advanced in modern studies of this domain to address the algorithmic shadow problem [Rahman A., 2020: 575–577]; [Schneier B., 2015: 448]. Some researchers believe technology can be successfully applied to deal with algorithmic shadow to guarantee the data removal right to data subjects, for example, as regards personal data processing [Li T.C., 2022: 505]. However, it is worth noting that the development of specific algorithms to remove corrupt data will come at a significant economic and technological cost. It suggests that using this method across the board to deal with algorithmic shadow, just as making it legally binding is premature and requires further study from both legal and technological perspectives.

2. Legal Mechanisms of Data Quality Assurance

2.1. Dualism of Approaches

Extreme importance of qualitative data parameters and potential impact on operation of language models suggest the need to assure these parameters in legal and technological terms. In view of the discussed regulatory

[16] Article 282 of the Criminal Code of Russia No. 63-FZ of 13 June 1996; Article 20.3.1, Administrative Code of Russia No. 195-FZ of 30 December 2001.

methods, two approaches to address this task can be proposed: normative approach based on imperative (centralized) method; and contractual approach based on dispositive (decentralized) method.[17]

Normative approach assumes that quality parameters will be established and assured via legally guaranteed mandatory technical requirements, standards, certification and control procedures, as well as directly by law. This will put in place general rules for all parties involved in AI development thus allowing to introduce stricter control. A downside of this approach may be its insufficient flexibility to adapt to changes, something likely to become critical in the context of rapid advance of information technologies.

Contractual approach, in its turn, relies on decentralized relations between the parties, with consensual data quality standards to enhance flexibility and adaptivity to varying demands and situations. However, that approach requires more complex engagement between the parties to legal relationships and cannot invariably guarantee that their interests are mutually observed (such as in case of an inadequate counterclaim under a paid service agreement, abuse by a stronger contracting party, etc.). With both approaches having upsides and downsides, the problem is likely to be efficiently addressed through a comprehensive solution combining certain elements of the approaches. It is useful discuss each of them in detail.

2.2. Normative Approach: Data Accuracy Principle

The number of regulations governing data quality is currently extremely limited, one regulatory source to be considered being the Federal Personal Data Law.[18] It establishes the principle of "data accuracy"[19] whereby data should be accurate, adequate and relevant for processing purposes. Moreover, the data that fall short of these criteria should be either deleted or corrected. This principle is echoed by the data subject's right to correct the underlying data.[20] Meanwhile, implementation of the said principle is problematic.

[17] The issue of qualification of regulatory methods is beyond the scope of the paper. Meanwhile, it should be noted that classification of regulatory methods is a subject of debate in doctrine. For example, the following methods are proposed: incentives and punishment, authorization (licensing), prohibition and enforcement.

[18] Federal Law No. 152-FZ "On Personal Data" of 27 July 2006 (as amended on 06 February 2023) (hereinafter "Federal Personal Data Law").

[19] Ibid. Para 6, Article 5.

[20] Ibid. Para 1, Article 14.

Firstly, the law does not specify to what extent personal data could fail to meet the criteria mentioned. Moreover, as was told above, a minor inaccuracy, inadequacy or irrelevance of data will not have a major impact under certain conditions.

Secondly, it is not clear how one can assess and measure the accuracy, adequacy and relevance of personal data with regard to processing purposes. For example, other countries' law will sometimes establish stricter requirements to data depending on processing purposes. Thus, Germany's Data Protection Act has a special provision on personal data processing for scoring — assessment of creditworthiness in the financial sector — that allows to use and process only the data obtained through a "scientifically acknowledged procedure of mathematical statistics".[21]

Implementation of this principle should apparently rely on the risk-oriented approach to allow for possibility to process in some cases the data that do not fully meet the required criteria while in other cases, on the contrary, specify and introduce stricter criteria for data processing.

Normative definition of data quality parameters through the said principle is also restricted by its inapplicability to all types of data since the Personal Data Law applies only to personal data processing.[22] Therefore, the said principle is applicable only to personal data processing. Moreover, now data cannot be invariably and unambiguously qualified as personal data, with difficulties concerning both the form of expression and qualification likely to arise at some processing stage. Overall, the issue is that the current definition of personal data[23] assumes a binary approach, that is, data can be either personal or otherwise. This approach does not take into account data for different individuals can be identifiable to a variable extent, for example, due to accessibility of other datasets [Oostveen M., 2016: 306], and that the current progress in information and computer sciences reveals different level of possible identifiability and related sets of risk [Kolain M., Grafenauer C., Ebers M., 2021:174]. In addition, it is noteworthy that data being processed could lose and acquire the relevant identifiability markers, that is, be dynamic rather than static. Therefore, data can be qualified as personal only at a specific stage of the language model's development. The

[21] § 31(1) Federal Data Protection Act (BDSG). Germany. Official English translation is available at: https://www.gesetze-im-internet.de/englisch_bdsg/englisch_bdsg.html#p0256 (accessed: 10.06.2024)

[22] Para 1, Federal Personal Data Law.

[23] Ibid. Para 1, Article 3.

data accuracy principle is thus applicable only to the data qualified as personal at the given stage than to all data processed at different stages of the language model's development.

2.3. Contractual Definition of Data Quality. Application of GOST

Regulating data quality through contractual terms is another approach. In this case, qualitative parameters could be described either explicitly with the help of the chosen technical criteria and specifications or with reference to the corresponding standards like GOST, or else via another applicable technical regulation.

Two types of contracts can be identified in the proposed context: those entered to settle the relationships with regard to data accessibility and use (such as a licensing agreement to deposit or use a database) and those not explicitly aimed at regulating the use of data but whose qualitative parameters are likely to impact significantly the relationships in question (such as a licensing agreement with the end user of a digital product).

In the first case, the parties will explicitly set the qualitative parameters of data in the relevant agreement. Thus, in order to deposit language data in the Common Language Resources and Technology Infrastructure (CLARIN)[24], the depositor will sign a licensing agreement describing qualitative parameters and forms of data to be uploaded, assigning responsibilities and also establishing the terms of payment and distribution of data based on sample licenses designed by the organization [Kelli A., Vider K., Lindén K., 2016: 13–24].

In the second case, the described qualitative parameters, terms of use and distribution will normally apply not to data but the underlying digital products. For example, before starting to use Yandex Speech Kit[25], users are required to accept the terms defining the procedure of use.[26] This situation will raise the question of whether the data (including qualitative pa-

[24] International infrastructure for support of research in the area of humanities and social sciences by providing access to various language resources and tools. For detail see: https://www.clarin.eu (accessed: 10.06.2024)

[25] A Yandex service allowing to transform text into speech (speech synthesis) and vice versa (speech recognition). See: URL: https://yandex.cloud/ru/services/speechkit?utm_referrer=https%3A%2F%2Fwww.google.com%2F (accessed: 10.06.2024)

[26] Speech Kit terms of use / Yandex Speech Kit. Available at: URL: https://yandex.ru/legal/cloud_terms_speechkit/ (accessed: 10.06.2024)

rameters) is relevant for the underlying digital product. Will the data lose independence, only to become its qualitative parameter? Who will be then responsible for the product's defects caused by questionable data: model developer or product developer? The answer to these questions is likely to be of principal importance both for performance under the said contracts and generally for the problem of contractual assurance of training data quality. Further, it should be noted that the question of assigning responsibility for harm caused by AI systems is debatable among researchers. It is generally proposed, firstly, to design risk minimization mechanisms already at the stage of AI system development; secondly, more specifically define who can assume responsibility for such harm; and, thirdly, apply a concept similar to that of "major hazard" in respect of AI systems [Kharitonova Yu.S., Savina V.S., Panyini F., 2022: 683-708].

One way to define quality data via contractual terms is to apply relevant technical standards such as intergovernmental standards (GOST). With regard to data, the fundamental document is GOST R ISO 8000-100-2019 Data quality[27] as well as GOST R ISO/MEK 20546-2021 Information technologies. Big data. Overview and glossary.[28] Key requirements to data quality such as accuracy, adequacy, relevance and consistency are defined in GOST R ISO 8000-100-2019 while GOST R ISO/MEK 20546-2021 provides an extensive overview and unification of the terms related to big data, something that helps to standardize the data processing approaches and establish a common conceptual framework for regulating the relations involved in language model training.

Technical Committee for Standardization No. 164 Artificial Intelligence (TK164) is currently in charge of developing relevant GOST applicable to AI and data.[29] The Committee is crucial for the development of regulatory framework for AI technologies in Russia, in particular, the rules that allow researchers and developers to have access to the required amount of data for efficient training of models and lower risk of unauthorized use of information. One standard under development in the discussed domain is Data quality for analytics and machine learning. The draft standard consisting of

[27] Rosstandard Order No. 836-st «On approving a national standard of Russia" of 29 October 2019.

[28] Rosstandard Order No. 632-st "On approving a national standard of Russia" of 13 July 2021.

[29] Set up by Rosstandard Order No. 1732 "On establishing the technical committee for standardization Artificial Intelligence" of 25 July 2019/ SPS Consultant Plus.

several parts follows ISO/IEC series 5259 international standard that equally consists of several parts that describe the principal concepts, terms and examples of defining data quality for analytics and machine learning, propose data quality model, measurement methodologies and guidance on data quality reports, and outline the data quality management process including risk management aspects and ways to meet the requirements to quality.[30]

Development and approval of the above GOSTs are supposed to greatly facilitate the issue of defining quality data in terms of both data parameters themselves and the applicable metrics and conceptual framework. Meanwhile, it should be noted that GOSTs will often contain rigid and detailed requirements that can be inappropriate or excessive in a particular case, only to complicate the adaptation of contractual terms to the parties' specific needs.

3. Regulating Access to Training Data

3.1. Personal Data

While defective data quality is normally related to technical shortcomings, legal defects primarily involve the problem of compliance with a legal regime of using data for training. Moreover, the problem itself is expressed in the form of conflict between the interests of developers critically in need of more or less free access to large amounts of data and the third-party interests protected by specific legal regime constraining such access.

The issue of compliance with data regime while developing AI models applicable, particularly, to personal data and other restricted information, as well as protection of intellectual property rights is recognized in the Strategy as one of the "challenges" faced by Russia in the area of AI development.[31]

As the issue of implications of personal data regime for language models to be developed and marketed was explored by the author in detail in previous studies, this paper will present only the main findings. One way to determine the extent of impact of personal data regime is to analyze physical, time-bound and territorial scope of the underlying regulation. Under this approach, physical impact can be determined in respect of different development stages of digital products and the extent of personal data use at each stage [Kelli A. et. al., 2021:154–159], while time limits by the effec-

[30] For detail on ISO/IEC series 5259 standard see: https://www.iso.org/ru/search.html?PROD_isoorg_ru%5Bquery%5D=ISO%2FIEC%205259 (accessed: 10.06.2024)

[31] Para 17(16), subparagraph (g), 2030 National AI Development Strategy.

tive term of data subjects' right to personal data protection, and territorial scope by national jurisdictions where the respective models are developed or marketed.

It is noteworthy that this approach reveals a number of shortcomings. For example, attempts to determine physical impact via different development stages show that data could lose and, on the contrary, acquire identifiability markers at some stage, only to considerably complicate their qualification as personal data.

Determination of time limits raises the issue of effective term of deceased persons' right to personal data protection. While not establishing such term, the law only prescribes that data in this case can be processed only if consented by successors where such consent was not given during the person's lifetime.[32] In absence of such term, no time limits of the underlying legal regime could be determined. It is equally noteworthy that, apart from the term problem, there is no order of priority in respect of successors who could give such consent, only to cause legal uncertainty in situations where some successors will withhold it while others not.

As for the territorial scope, the problem is in the need to comply with different national regimes at a time which is often impractical as, for example, in the case of the General Data Protection Regulation[33] and Russian personal data protection law. Meanwhile, it could become necessary due to both exterritorial effect of regulation itself, specific and related fields, and because of technical necessity to collect and process data in the national territory of other countries.

3.2. Protecting Intellectual Assets in Designing and Training Language Models

Whereas the impact of personal data regime on development of language models was explored by the author in detail in previous studies, the issue of the underlying use of intellectual assets received less attention. The urgency of this problem is confirmed by numerous cases of litigation be-

[32] Para 7, Article 9, Federal Personal Data Law.

[33] Regulation (EU) 2016/679 of the European Parliament and the Council of 27 April 2016 of 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 force since 25 May 2018. Available at: URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj (accessed: 10.06.2024)

tween language model developers and authors such as, for example, claims against OpenAI[34], language model developer for Chat GPT. Let us discuss the problem in more detail.

In the context of intellectual property law, data used for language models such as texts and audio files can be represented as items of copyright and related rights. Meanwhile, it is noteworthy that they will not be protectable across the board. Thus, copyright protectability criteria include creative components and objective form of a work,[35] with related rights exercisable to the extent that the copyright to the work used to create the item of related rights was observed[36] etc.

Depending on extent of protectability of copyright and related rights, the data used to develop a language model could be divided into three groups: “unprotected” (such as acts of legislation, official documents etc.), “safe” (such as manuals, technical specifications, expert opinions etc., all those generally not subject to protection) and works subject to copyright and related rights [Truyens M., Van Eecke P., 2014: 153–170]. A functional language model will require the components from all the three groups: the use of only “unprotected” and “safe” groups will not suffice. Meanwhile, it is technically problematic to draw a line for associating specific components with a particular group. Thus, though not all language model data will be subject to copyright and related rights, one cannot exclude the use of protected items for sure.

One caveat is in order regarding the concept of “using” the said items to develop a language model. Some researchers believe that the use of works for data mining — a stage of the model’s development — does not involve copyright since it protects the creative form of expression while in data mining works are viewed as a database and are thus outside the available remedies [Kolsdorf M., 2021: 142–164]. This assumption is, in our view,

[34] See, for example, collective lawsuit, case No. 1:24-cv-00084, Nicholas Gage v Microsoft, OpenAI, United States District Court for the Southern District of New York. Available at: https://fingfx.thomsonreuters.com/gfx/legaldocs/klvydkdklpg/OPENAI%20COPYRIGHT%20LAWSUIT%20basbanescomplaint.pdf (accessed: 10.06.2024). Lawsuit, case 1:23-cv-11195, the New York Times company v. Microsoft, OpenAI, United States District Court for the Southern District of New York. Available at: https://nytco-assets.nytimes.com/2023/12/NYT_Complaint_Dec2023.pdf (accessed: 10.06.2024)

[35] Para 80–82, Federal Supreme Court Plenum Resolution No. 10 “On the application of Part IV of the Civil Code of Russia” of 23 April 2019.

[36] Para 3, Article 1303, Civil Code of Russia.

questionable. In the first place, the idea that copyright protects only the creative form is open to debate. Despite that this approach is explicitly reflected in law[37], research papers invoke a need to extend the scope of copyright to the work's content [Gavrilov E.P., 2009: 31–38] or else attempts to use the existing legal mechanisms to overcome the said constraint, for instance, by delineating the concepts of external and internal forms of a work [Kashanin A.V., 2010: 68–138]. Moreover, the use of works for data mining should be considered in conjunction with other related operations including those preceding mining such as copying, collection, transmission and classification of data. Except for temporary copying required for technological process and not amounting to the use of works[38], the said operations can involve intellectual property rights. The above is equally applicable to language models where data mining is just a development stage.

Using the items of copyright and related rights to design language models will require to comply with the author's personal non-property rights as well as the underlying exclusive rights. As such, the use of copyrighted items can rely on two patterns, the first based on the author's (other copyright holder's) prior consent (in the form of licensing agreement or that for assignment of exclusive rights), the second (doctrine of free use) restricting the author's (other copyright holder's) rights. While neither of the said patterns fully satisfies the industry's needs, they involve risks related to illegitimate use of intellectual property assets meaning violation of copyright and related rights.

The first pattern based on the author's prior consent to use copyrighted items for linguistic resources is apparently the least risky in terms of violation of copyright and exclusive rights. However, it raises an issue primarily related to identification of the author or other copyright holder who is often impossible to identify. It is further complicated by the question of how to go about the works created automatically or with minimum human involvement. Another trouble is that of time and cost of negotiations to conclude the respective agreements.

It is worth noting that large technological dotcom companies providing a wide range of digital services will often resort to such pattern. For example, the licensing agreement for Alisa voice assistant allows Yandex to

[37] Para 5, Article 1259, Civil Code of Russia (Part IV) No. 230-FZ of 18 December 2006 ("Civil Code of Russia").

[38] Para 2, Article 1270, Civil Code of Russia.

use voice prints borrowed not only from the application but also from the company's numerous other services.

While the second pattern based on the free use doctrine does not involve any time or cost in terms of author's consent and payment of royalties, its use in Russia is restricted to specific cases listed in law.[39] As developing a language model — language digitization and data mining — is potentially important for both science and culture, the model's use for "information, research, education and culture" is likely to be the most suitable of all cases of free use established by law.[40] However, the following analysis will reveal a number of complications to apply this exception.

With the invoked purposes to fit together, free use for research, education or culture also requires to specify the author and a borrowing source, allowing to use a work to the extent that fits the citing purposes [Gracheva D.A., 2023: 50]. These criteria are impractical to meet with regard to a language model.

Firstly, as was noted above, it is not always possible to exactly identify the authors of all works being used and thus make the respective references.

Secondly, the law does not say how to determine whether a work is used within the extent of the respective citing purposes. A functional language model will require a considerable amount of data to inevitably include protected works, with their number and extent of their citing likely to differ depending on the underlying technology and purposes. In absence of the criteria to determine the extent of possible citing, there will always be risk that in a given case the use of a work may be recognized as excessive in relation to purposes.

Thirdly, development of a model does not always serve only scientific and cultural purposes. In this particular case, the issue lies in the ratio of business and scientific/cultural purposes. The natural language processing technology has a scientific and social value, something that does not rule out its high economic potential. In this regard, the question is whether one could rely on the doctrine of free use to develop models for subsequent commercialization. It is logical to assume that where such model was originally developed by a business entity, this doctrine would generally be of no avail. Meanwhile, this situation is causing the number of potential produc-

[39] Sub-para 1, para 2, Article 1270, Articles 1273–1280, Civil Code of Russia.

[40] Article 1274, Civil Code of Russia.

ers to dwindle by excluding those without enough resources to rely on the first pattern based on prior consent, only to constrain the development of the entire sector.

Thus, the application of both first and second patterns to use protected works for designing language models and developing NLP technologies is now thwarted.

3.3. Data Dissemination: Repositories and Re-use

A possible solution to the above issue is to encourage higher education institutions to engage in the development and creation of language databases (linguistic corpuses) for further dissemination via a licensing agreement system. As language digitization has a high social value, the involvement of universities in this process appears logical and reasonable. There are examples of partnership between business entities and universities for development of natural language processing technologies such as ABBY chair at the Moscow Institute of Physics and Technology (MIFT), or the joint academic program of the Tsentr Rechevykh Tekhologiy company group and the ITMO National Research University. However, the development of linguistic corpuses on the basis of universities, while addressing the problem of targeted use of data to directly create linguistic corpuses and develop language models, leaves out the issues of their further dissemination. Thus, what will happen if a university loses interest in further dissemination of a linguistic corpus for some reason or other, or does not have enough funds to do it? On the contrary, can a university create a linguistic corpus through free use of works and then commercialize the outcomes relying on the concept of entrepreneurial university through a spin-off company? All these questions are currently open and urgent and require further in-depth study and analysis from the perspective of both jurisprudence and other sciences.

Another possible solution to the data accessibility problem is to make the data at state information systems (SIS) available to developers, that is, allow to re-use the already accumulated data. Re-use of SIS data for designing language models can considerably expedite the process of development and introduction of new technologies as well as enhance their effectiveness and adaptivity to various areas of application. As stated by the Federal Accounting Chamber in an analytical report[41], there were over 800 federal SIS

[41] Available at: https://ach.gov.ru/upload/pdf/Оценка%20открытости%20ГИС%202020.pdf (accessed: 10.06.2024)

in Russia in 2020 for support of information exchanges between public authorities in various social spheres. These systems contain data ranging from statistics to education, health and other socially important sectors. The use of SIS data in the interest of technological development thus appears to be quite promising.

Despite a varying degree of maturity of such systems, there is a reason to assume that SIS data will be of sufficient quality while their diversity will ensure representativeness. This will lay down a robust foundation for designing high quality, comprehensively trained language models capable of addressing widely diverse tasks. However, this will only be possible if the specifics of each type of data and their adequacy for the given purpose are carefully accounted for.

Meanwhile, data re-use is fraught with a number of legal and ethical risks related to both compliance with legal regimes (such as tax secret, personal data) and transparency and safety. Preventing the said risks will apparently require to develop common regulatory principles and approaches to data re-use including clear legal provisions and standards of data protection and data subject rights, as well as generally enhance control and audit mechanisms for the use of data to develop AI systems.

Conclusion

The paper was designed to provide a conceptual analysis of the regulatory problem for quality assurance and accessibility of training data in the context of the Strategy's objectives.[42]

Firstly, with regard to data quality assurance, likely implications of using corrupt data were explored and discussed from the perspective of undermining both technical parameters of data (quality defect) and legal regime (legal defect). Secondly, two approaches to data quality assurance were analyzed: normative and contractual. Despite their inherent downsides, it is feasible to use and apply both approaches in developing relevant regulation.

With regard to data accessibility, the research has allowed to identify and describe a number of constraints to use data for training. These constraints come in the first place from normative barriers that impede access to data due to a need to comply with the underlying legal regimes, as well as from a lack of adequate legal mechanisms to override them. These constraints to

[42] 2030 National AI Development Strategy.

a large extent slow down the process of development and introduction of language models to undermine the technology's progress as a whole as well as digital transformation of various economic and social sectors.

Progress of the technology will largely depend, on the one hand, on cooperation between all of the sector's stakeholders and, on the other hand, on the availability of modern regulation to support its sustainable development.

References

1. Dash N.S., Arulmozi S. (2018) History, features, and typology of language corpora. Singapore: Springer, p. 291.

2. Feng Z. (2023) Formal analysis for natural language processing: a handbook. Berlin: Springer Nature, pp. 7,8, 25.

3. Gavrilov E.P. (2009) Copyright and the content of artistic work. *Patenty i litsenzii*=Patents and Licenses, no. 7, pp. 31–38 (in Russ.)

4. Glauner P. (2024) Technical foundations of generative AI models. Legal Tech — Zeitschrift für die digitale Anwendung, pp. 24–34.

5. Goldberg Y. (2017) Features for textual data. In: Neural network methods for natural language processing. Cham: Springer, pp. 65–76.

6. Gracheva D.A. (2023) Free use of copyright and related rights in the context of development of digital technologies in Russia. *Trudy po intellektualnoy sobstvennosti*=Works on Intellectual Property, vol. 45, no. 2, pp. 44–52 (in Russ.)

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

8. Hirschberg J., Manning C.D. (2015) Advances in natural language processing. *Science*, vol. 349, no. 6245, pp. 261–266.

9. Kashanin A.V. (2010) Development of ideas on the form and content of works in the copyright doctrine. The problem of protectability of research works. *Vestnik grazhdanskogo prava*=Bulletin of Civil Law, vol. 10, no. 2, pp. 68–138 (in Russ.)

10. Kelli A., Vider K., Lindén K. (2016) The regulatory and contractual framework as an integral part of the CLARIN infrastructure. CLARIN Annual Conference. Linköping University Electronic Press, pp. 13-24. Available at: https://helda.helsinki.fi/server/api/core/bitstreams/1f7b8a3c-790c-4e66-9677-f5f9aca785d6/content (accessed: 04.07.2024)

11. Khyani D. et al. (2021) An interpretation of lemmatization and stemming in natural language processing. *Journal of Shanghai University for Science and Technology,* vol. 22, no. 10, pp. 350–357.

12. Kolain M., Grafenauer C., Ebers M. (2021) Anonymity assessment-a universal tool for measuring anonymity of data sets under the GDPR with a special focus on smart robotics. *Rutgers Computer & Technology Law Journal,* vol. 48, p. 174.

13. Kolzdorf M.A. (2021) Free use of the items subject to copyright and related rights in Big Data processing. *Zakon*=Law, no. 5, pp. 142–164 (in Russ.)

14. Li T.C. (2022) Algorithmic destruction. *Southern Methodist University Law Review,* vol. 75, pp. 480-505. DOI: https://doi.org/10.25172/smulr.75.3.2

15. Lythreatis S. et al. (2022) The digital divide: a review and future research agenda. *Technological Forecasting and Social Change,* vol. 175, pp. 1–11.

16. Mushakov V.E. (2022) Constitutional human rights in the context of addressing the digital divide. *Vestnik Sankt-Petersburgskogo universiteta* MVD=Bulletin of Saint Petersburg University of Interior Ministry, no. 1, pp. 69–73 (in Russ.)

17. Oostveen M. (2016) Identifiability and the applicability of data protection to big data. *International Data Privacy Law,* vol. 6, no. 4, pp. 299–309.

18. Rahman A. (2020) Algorithms of oppression: how search engines reinforce racism. *New Media* & *Society*, vol. 22, no. 3, pp. 575–577. DOI: https://doi.org/10.1177/1461444819876115.

19. Rogers S. E. (2016) Bridging the 21st century digital divide. *TechTrends*, vol. 60, no. 3, pp. 197–199.

20. Russo A., Proutiere A. (2021) Poisoning attacks against data-driven control methods. 2021 American Control Conference (ACC). IEEE, pp. 3234–3241. Available at: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9482992 (accessed: 04.07.2024). DOI: 10.23919/ACC50511.2021.9482992.

21. Schneier B. (2015) *Data and Goliath: the hidden battles to collect your data and control your world.* N.Y.: Norton, 448 p.

22. Truyens M., Van Eecke P. (2014) Legal aspects of text mining. *Computer Law & Security Review,* vol. 30, no. 2, pp. 153–170.

23. Zhou M. et al. (2020) Progress in neural NLP: modeling, learning, and reasoning. *Engineering*, vol. 6, no. 3, pp. 275–290.

Information about the author:

I.G. Ilyin — Postgraduate Student.

The article was submitted to editorial office 21.06.2024; approved after reviewing 29.06.2024; accepted for publication 29.06.2024.

Legal Issues in the Digital Age. 2024. Vol. 5. No 2.
Вопросы права в цифровую эпоху. 2024. Том 5. № 2.

Research article
УДК*:* 340, 34.06
DOI:10.17323/2713-2749.2024.2.57.79

The Use of AI in Medicine: Health Data, Privacy Risks and More

Boris Aleksandrovich Edidin[1], Alexey Vasilievich Bunkov[2], Ksenia Vladimirovna Kochetkova[3]

[1, 2] ANO «IRI, Institute for Digital Content Support & Development», 6/1/2 Kadashevskaya Embankment, Yakimanka Area, Moscow 119017, Russia,

[3] MGIMO University, 76 Prospect Vernadskogo, Moscow 119454, Russia,

[1] b.edidin2018@gmail.com

[2] bunkov.a@iri.center

[3] kochetkova.k@iri.center, Author ID: 1033155, ORCID ID: 0000-0002-6254-9539, Scopus Author ID: 57223024821

Abstract

In the era of advancements in artificial intelligence (AI) and machine learning, the healthcare industry has become one of the major areas where such technologies are being actively adopted and utilized. The global health care sector generated more than 2.3 zettabytes of data worldwide in 2020. Analysts estimate that the global market for artificial intelligence (AI) in medicine will grow to $13 billion by 2025, with a significant increase in newly established companies. Artificial intelligence in medicine is used to predict, detect and diagnose various diseases and pathologies. The sources of data can be various results of medical research (EEG, X-ray images, laboratory tests, e.g. tissues, etc.). At the same time, there are understandable concerns that AI will undermine the patient-provider relationship, contribute to the deskilling of providers, undermine transparency, misdiagnose or inappropriately treat because of errors within AI decision-making that are hard to detect, exacerbate existing racial or societal biases, or introduce algorithmic bias that will be hard to detect. Traditional research methods, general and special ones, with an emphasis on the comparative legal method, were chosen. For the AI to work it needs to be trained, and it's learning from all sorts of information given to it. The main part of the information on which AI is trained is health data, which is sensitive personal data. The fact that personal data is qualified as sensitive personal data indicates the significance of the information contained, the high risks in case it's leaking, and hence the need for stricter control and regulation. The article offers a detailed exploration of the legal implications of AI in medicine, highlighting existing challenges, the current state of regulation, and proposes future perspectives and recommendations for legislation adapted to the era of medical AI. Given the above, the study is divided into three parts: international framework, that will focus primarily on applicable WHO documents; risks and possible ways to minimize them, where the authors have tried to consider various issues related to the use of AI in medicine and find options to address them; and relevant case-study.

Keywords

health data; AI training; sensitive personal data; privacy; deanonymization; data quality; algorithms; World Health Organization.

For citation: Edidin B.A., Bunkov A.V., Kochetkova K.V. (2024) The Use of AI in Medicine: Health Data, Privacy Risks and More. *Legal Issues in the Digital Age,* vol. 5, no. 2, pp. 57–79. DOI:10.17323/2713-2749.2024.2.57.79

Introduction

AI-powered applications have demonstrated their potential to transform medical diagnosis, treatment plans, drug discovery and patient care. E.g., the use of ChatGPT-like solutions in health care has enormous potential to improve the patient-provider relationship, such as patient clinic letter writing, medical note-taking and consultation, etc. [Chen C.W., Walter P., Wei J.C., 2024].[1]

The top 5 countries in the use of AI in medicine currently are the USA (48%), UK (7%), Israel (6%), Canada (4%), China (3%) [Imameeva R.D., 2021: 34].

Russia is also one of the leading countries in digitalization in medicine, including the use of AI. In 2022 a unique digital library of anonym da-

[1] The use of ChatGPT is being actively discussed in other fields besides medicine, especially in education and law. For example, for three months, experts from ANO IRI tested ChatGPT for its possible use in analytical and legal applications. The neural network was studied "out of the box", i.e. without any additional customizations and technical integrations with other services. Available at: URL: https://ири.рф/news/eksperty-iri-protestirovali-ispolzovanie-chatgpt-v-sfere-yurisprudentsii-i-normotvorchestva/ (accessed: 12.03.2024)

tasets for the evaluation and training of neural networks started functioning in Moscow.[2] In July 2023, unique AI-based medical technologies were presented at the Russia-Africa Summit.[3] Moscow is already using 12 AI systems in healthcare, which are registered and approved by the Federal Service of Roszdravnadzor; most of them are neural networks that assist radiologists.[4] At the moment of writing, smart algorithms are already assisting doctors in finding pathologies in 21 clinical areas.[5]

As of May, 2024 a medical decision support system based on artificial intelligence has facilitated preliminary diagnoses in Moscow hospitals, amounting to 14 million diagnoses.[6] A digital assistant system called "TOP-3" which analyses the patient's health complaints and offers three preliminary diagnoses. The physician may then either concur with one of the proposed diagnoses or formulate an alternative. The service is capable of identifying 95% of the most common diseases. However, neural networks do not replace the role of the doctor; rather, they free up the time and attention of the specialist to examine the patient and communicate with them. Ultimately, the final decision is always at the discretion of the physician.

Yandex and Sechenov University have launched a cloud platform of medical data for scientists in Russia, with 18 million medical documents uploaded to it.[7] With its help, specialists will be able to quickly find relevant medical reports, test results, CT scans, X-rays and other information, the company assures. They specified the platform will help scientists to develop new drugs and treatment methods, and developers — to develop AI in the field of health care.

[2] Moscow opened access to a digital data library for developers of artificial intelligence services in medicine // Available at: URL: https://www.mos.ru/news/item/107729073/ (accessed: 12.12.2023)

[3] Available at: sberbank.ru/ru/sbertv/broadcast/article?video=XB3j0r&listId=2 (accessed: 12.12.2023)

[4] Available at: URL:

[5] https://www.rbc.ru/rbcfreenews/6463777c9a7947472428599e (accessed: 12.12.2023) 5 Ibid.

[6] Rakova claimed to make 14 million provisional diagnoses using AI // Available at: URL: https://www.rbc.ru/rbcfreenews/664cc2e99a7947847c959310 (accessed: 22.05.2024)

[7] Yandex and Sechenov University have created a medical data platform for scientists // Available at: URL: https://www.forbes.ru/tekhnologii/502840-andeks-i-secenovskij-universitet-sozdali-platformu-medicinskih-dannyh-dla-ucenyh?erid=LdtCKapuV (accessed: 22.12.2023)

One of the most recent advancements in the field of AI application in healthcare is an AI-based algorithm to create drugs. Russian researchers at ITMO University have developed AI-based algorithm that will simplify and reduce the cost of creating finished pharmaceutical forms.[8] The new solution will make it possible to generate auxiliary molecules with the desired properties and assemble the basis of a future drug before experiments are conducted.

AI can perform a large number of functions related to the automation of labor-intensive processes and assistance to the medical doctor, such as:

analyzing and processing data (to make possible diagnoses and conclusions, or to come up with personalized treatment (e.g. individual therapy plans, precise selection of drug dosages, etc.);

monitoring the effectiveness of the actions taken (assessment of treatment dynamics);

monitoring of the patient's condition, recording the indicators from body sensors or hospital equipment data;

interacting with patients and their relatives to collect primary information or counseling on standard issues;

exercising various auxiliary functions related to document management and medical staff activities, such as voice recognition systems for filling out medical records, medical histories and other documents.

As a separate direction, robotic surgery can be singled out, in which robots help out during operations both with the participation of a doctor (robots act as assistants) or entirely by themselves without human participation, for more common and "easy" surgical procedures.

Another area is conducting research in pharmacology and development of new drugs and vaccines [Thomas S., Abraham A. et al., 2022]; [McCaffrey P., 2022]; [Boniolo F., Dorigatti E. et al., 2021]. By implementing AI technologies, pharmaceutical companies are able to shorten drug development and clinical trials, thereby reducing the cost of launching new drugs, which also facilitates the production of high-quality drugs with fewer side effects [Alekseeva M.G., Zubov A.I., Novikov M. Yu., 2022: 11].

These are just a few examples of what artificial intelligence is being used for in healthcare (more uses see Fig. 1).

[8] An algorithm for fast drug formation with the help of AI created in St. Petersburg // Available at: URL: https://nauka.tass.ru/nauka/20046793?utm_source=yxnews&utm_medium=desktop&utm_referrer=https%3A%2F%2Fdzen.ru%2Fnews%2Fsearch%3Ftext%3D (accessed: 12.03.2024)

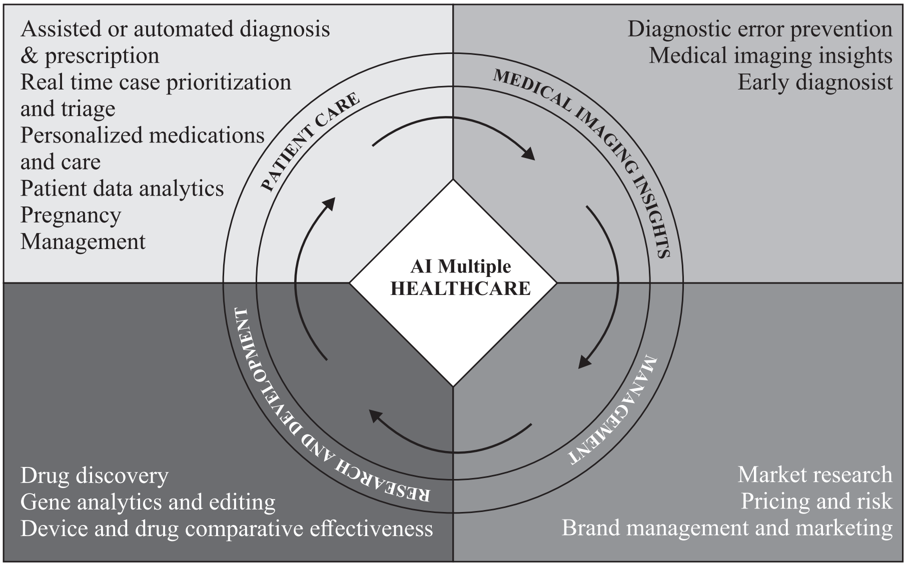


Fig. 1. AI Use Cases in Healthcare Industry in 2024[9]

All models, AI systems, algorithms depend on health data which comes from different sources:

clinical data (laboratory data, patient records and etc.);
genomic data (mostly means genetic testing);
imaging (results from X-ray, MRI, and other radiology diagnostics);
administrative data (not health data per se, but the information connected to a patient like financial statements, insurance, billing info, etc.);
sensors and wearables data.

Being a central element of the above-mentioned technologies is the use of huge amounts of medical data of patients, which is the basis for building any algorithm. And this vast amount of data shall be handled accordingly though its lifecycle (see Fig. 2).

However, the prospect of using AI in healthcare is accompanied by a number of legal[10] and ethical issues, in particular those related to the criticality for human health and life of any errors in decision-making, as well as the collection, storage and use of confidential patient information.

[9] Available at: Top 18 AI Use Cases in Healthcare Industry in 2024 (aimultiple.com), (accessed: 22.05.2024)

[10] The legal and policy issues around privacy and patient data affect both clinical AI and other health care AI systems [McNair D., Price W.N. 2019: 197].

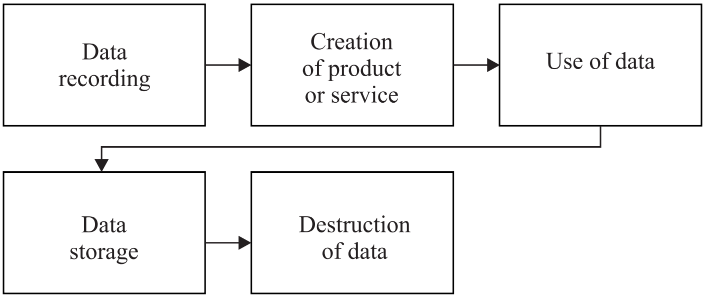


Fig. 2. Health Data Lifecycle

1. International Framework

As AI is a new and emerging technology, so far there is been no framework adopted on the issue. That is to say obligatory framework in the form of an international treaty that would regulate the use of AI. However, there are a number of soft-law documents on the issue.

At this point, the most extensive framework in a form of recommendation was provided by UNESCO as one of the specialized agencies of the UN. The paper "The Recommendation on the Ethics of Artificial Intelligence"[11] serves as an ethical guideline and helps to ensure strict adherence to the rule of law in the digital world. The document focuses on 11 Policy Areas, including "Health and Social Well-Being". As stated in Clause 122 (d): "Member States should pay particular attention in regulating prediction, detection and treatment solutions for health care in AI applications by ensuring effective mechanisms so that those whose personal data is being analyzed are aware of and provide informed consent for the use and analysis of their data, without preventing access to health care".[12]

Concerns about the introduction of AI in healthcare have been raised by another UN specialized agency — the World Health Organization (hereinafter WHO). First, in 2018 WHO has adopted a Resolution on Digital Health.[13] Among other things, the Resolution urges WHO Member States "to develop, as appropriate, legislation and/or data protection policies

[11] Recommendation on the Ethics of Artificial Intelligence // Available at: https://unesdoc.unesco.org/ark:/48223/pf0000381137_eng (accessed: 20.01.2024)

[12] Ibid.

[13] Seventy-first World Health Assembly, Agenda item 12.4 "Digital Health", 26 May 2018 // Available at: https://apps.who.int/gb/ebwha/pdf_files/WHA71/A71_R7-en.pdf (accessed: 20.01.2024)

around issues such as data access, sharing, consent, security, privacy, interoperability and inclusivity consistent with international human rights obligations and to communicate these on a voluntary basis to WHO".

Later, following the Resolution of 2018, WHO Member States agreed on the Global Strategy on Digital Health for 2020–2025 which highlights the importance of AI.[14] As set forth in the Strategy, health data shall be classified as sensitive personal data and be attributed the highest possible safety and security standard (see page 11 of the document). Besides safety, Member States need to make sure the data is complete in its integrity. The Strategy points out that the use of health data to train AI is a secondary use of health data and shall be accompanied with appropriate deanonymization of datasets.

The WHO has also urged caution in using AI-generated large language model (LLM) tools to protect and promote human well-being, safety and autonomy, and to preserve public health. [15] It is noted that hasty implementation of unproven systems could lead to errors made by health professionals, or harm patients thus undermining trust in AI. Among the main concerns, WHO highlighted that the data used to train AI may be biased, creating misleading or inaccurate information.

Recently, in January 2024 the WHO has published framework "Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models".[16] The guide outlines more than 40 recommendations for consideration by governments, technology companies and healthcare providers to ensure the appropriate use of LMM to promote and protect public health. Among other things, the risks and potential benefits of LMM are described, and the following key recommendations are provided for LMM developers to ensure the following:

[14] Global strategy on digital health 2020-2025. Geneva: World Health Organization; 2021 // Available at: https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf (accessed: 20.01.2024)

[15] WHO outlines considerations for regulation of artificial intelligence for health, October 2023 // Available at: https://www.who.int/news/item/19-10-2023-who-outlines-considerations-for-regulation-of-artificial-intelligence-for-health#:~:text=The%20World%20Health%20Organization%20(WHO),manufacturers%2C%20health%20workers%2C%20and%20patients (accessed: 20.01.2024)

[16] Ethics and governance of artificial intelligence for health: Guidance on large multimodal models // Available at: https://www.who.int/publications/i/item/9789240084759 (accessed: 05.02.2024)

LMMs are not only developed by scientists and engineers. Potential users and all direct and indirect stakeholders, including health care providers, academic researchers, healthcare professionals, and patients, should be involved early in the product development process.

LMMs must perform well-defined tasks with the necessary accuracy and reliability to improve the capacity of health systems and protect the interests of patients.

2. Risks and Possible Mitigating Measures

The following key risks associated with the use of AI in medicine can be identified:

a) Errors in AI algorithms

In addition to inaccurate and low-quality data on which AI is trained, errors can also occur in the algorithms themselves, for example, due to incorrect AI programming or failure to take into account any data used, which can lead to incorrect treatment recommendations or diagnoses. The issue here is the transparency of the algorithms and their ethical use.

b) Breach of privacy and data security

Accumulation and use of large amounts of patient medical data increase the risk of unauthorized access, breach of confidentiality and data leaks.[17] The legislation of the Russian Federation provides a number of mandatory requirements for information security in the collection and processing of personal medical data.

c) Risk of data deanonymization

Cleansing patient data from personal information does not guarantee anonymization, as artificial intelligence models can re-identify a person. In order to minimize this risk, an example of a possible practice may be the anonymization of personal data by mixing method (shuffle). In this case, the original field value of one record is replaced by a randomly selected value of the same attribute of another data record within the same dataset.

[17] For example, in the USA the HHS' Office for Civil Rights has reported over 239 breaches in 2023, affecting the health care data of more than 30 million individuals within the U.S. See: B. Lewis. Navigating Health Data Privacy in AI-Balancing Ethics and Innovation. Available at: https://www.lexology.com/library/detail.aspx?g=19c61aa9-3e34-4894-84b4-81d814de926c (accessed: 20.01.2024)

The serious limitation of this method is the unsuitability of “shuffled” data for the search of possible correlations.

d) Data quality, validity and relevance.

A substantial barrier to innovation in healthcare seems to be the availability of high-quality data on which to train AI, which limits the types of users who can successfully innovate [Price W.N., Sachs R., Eisenberg R.S., 2021:39]. To ensure data quality, reliability and relevance, a number of principles and sequences must be followed. In the course of the performance the following should be provided:

definition of the goals and objectives of data collection, the planned ways of their subsequent use and the tools used in this process, the planned volume of data to be collected;

preparing a statement of work or other structured description of the data collection requirements, which includes the requirements for the planned result;

collecting raw data from various sources and determining the sources of information, their reliability, and the type of data to be collected. Algorithm developers need to assemble data from multiple sources to train machine learning algorithms. Those data — as well as data about how the algorithms perform in practice – may then be shared with other entities in the healthcare system for the purpose of evaluation and validation [Price W.N., 2017: 13]. For example, the “Regulations for the preparation of datasets describing approaches to generating a representative sample of data” distinguish the following types of data used in medical AI: medical records, electronic medical records, laboratory data, medical images, genomics, auxiliary data;[18]

data annotation and data markup. Annotation and markup refer to the processing of raw data for the purposes of its use in machine learning, in which the data is assigned a label or tag that allows algorithms to classify the received and processed information. The outcome of partitioning is the presence of fixed patterns in the data and its characteristics. This allows machine learning models to further interpret and sort incoming data. This is one of the key and labor-intensive stages of AI training work. According

[18] Prepared by the State Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Moscow Health Department. Available at: URL: https://telemedai.ru/biblioteka-dokumentov/reglament-podgotovki-naborov-dannyh-s-opisaniem-podhodov-k-formirovaniyu-reprezentativnoj-vyborki-dannyh-chast-1-1 (accessed: 20.01.2024)

to the research by Cognilytica, preparing a dataset can take up to 80% of the total development time of an AI solution: e.g. in video annotation each hour of video requires about 800 man-hours.[19]

The organization of the data annotation and markup process should take into account:

availability of a sufficient number of qualified personnel, taking into account the planned volumes of annotation;

refusal to use outsourcing and crowdsourcing for medical data processing due to increased risks of leaks and possible low quality of the final result;

use of Russian software for data annotation and markup;

prohibition of remote work with data (e.g., when employees work on their personal devices from home). Thus, working in a secure corporate network is essential;

training of employees responsible for data annotation and markup;

feedback between the team of employees responsible for data annotation and markup and the teams that train and use the trained models to clarify and update the procedures for working with data, to quickly take into account necessary changes, to take into account the identified errors and the possibility of their prompt correction.

correct data entry into the system (correct and accurate description, no duplication of data, etc.). This shall be accompanied by reducing the impact of human factor: it's highly recommended to use automated data entry and verification, or independent verification by other specialists);

regular data auditing/verification/updating. In addition to internal verification, an external audit by an independent third party is recommended. One option would be to establish a system of licensed organizations authorized to conduct this type of audit;

data cleansing. Data cleansing involves identifying and correcting any errors or inconsistencies in the data. Data cleansing shall be performed by automated tools. However, in order to ensure data cleansing quality, it is advisable to provide for random checks by specialists. Cleansing should result in: deletion of duplicate data; deletion of data not related to the dataset; identification of missing data (ensuring data completeness); standardization of data — unified standards of data recording, transformation of data

[19] Cognilytica White Paper AI Data Engineering Lifecycle Checklist Following Steps for AI Project Success, 2020 // Available at: https://www.cloudera.com/content/dam/www/marketing/resources/whitepapers/ai-data-lifecycle-checklist-cloudera-whitepaper.pdf?daqp=true%20. (accessed: 20.01.2024)

into selected standards. This point is most critical if the data set is collected from different sources. Data cleansing can be broken down into the following steps: parsing; correcting; standardizing; matching; consolidating [Stöger K., Schneeberger D., Kiesebergc P., Holzinger A., 2021].

Competence Substitution (leading to medical negligence).

The frequent use of AI systems by a specialist creates a technology dependence, which can potentially lead to a loss of skills and abilities among medical staff, and a shift of responsibility for decision-making in the medical system. This can be particularly critical in the event of system failure or malfunction and the need for rapid decision making.

These risks can be minimized both at the regulatory level and at the level of the medical facility itself.

In the first case, it is necessary to legally establish the use of AI in healthcare (for example, by adding Article 36.3 "Peculiarities of Medical Care Provided Using Artificial Intelligence" to the Federal Law "On Basics of Health Protection of Citizens in the Russian Federation"), as well as to establish the criminal liability of a doctor for the final decision made using AI tools.[20] It has been supported by some specialists, who point out that medical doctors and hospitals that use AI bear the ultimate responsibility for its use, however, they need to be trained accordingly [Naik N. et al., 2022].

The second case requires regular training of personnel on how to work with the results obtained, as well as verification of decisions made by the doctor (e.g., random verification of decisions by an independent medical commission on a regular basis).

At the moment there is no unified approach to the issue of liability for errors resulting from the use of AI in medicine. A diagnosis founded upon AI technology offers an array of issues that are difficult to remedy through present concepts of responsibility [Hodge S.D., 2022: 436]. Some researchers assume that certain AI models should be given a unique legal status akin to personhood to reflect its current and potential role in the medical decision-making process, thus clearing who should bear responsibility and for what [Chung J., Zink A., 2018: 1]. The national legislation of some states is trying to find a solution, but it is still at the draft stage. An interesting approach in this context is that of Brazil, where a bill on AI use by doctors,

[20] Federal Law of 21.11.2011 No. 323-FZ // Available at: URL: https://www.consultant.ru/document/cons_doc_LAW_121895/ (accessed: 02.03.2024)

lawyers, and judges is under consideration by the National Congress.[21] The bill allows the use AI systems by doctors if it is under the doctor's supervision and the doctor's autonomy is preserved. The use of AI without such supervision would be considered as illegal medical practice.

The recently adopted European Union AI Act[22] also refers to the use of AI in healthcare. However, it does not directly cover the issues of responsibility for AI decisions and diagnoses by doctors, the Act still has some implications for the health sector. That being said, the Act classifies as high-risk AI systems those that could have a significant impact on human health and safety. AI Act mandates strict compliance for high-risk systems in terms of testing, documentation and transparency. The EU AI Act is founded upon a commitment to the upholding of ethical principles and the protection of fundamental human rights. The Act mandates that AI systems, especially those used in healthcare, are developed and deployed in a manner that respects human dignity, autonomy and privacy.

A similar approach has been effectively adopted by European Union GDPR: its Article 22 regulates decisions based solely on automated processing.[23] As it has been argued, this could be interpreted as to establish a "right to information and explanation, and therefore entail that "black box" systems, which do not allow any "meaningful human control", nor any explanation, should be prohibited.[24]

The use of AI in medicine, in addition to the above, also raises the following questions: Who is responsible for algorithm development? Is phased implementation with testing necessary? Are there market authorization procedures and is there certification? Are there regulations for identifying

[21] Projeto de lei No. 266, de 2024 sobre o uso de sistemas de inteligência artificial para auxiliar a atuação de médicos, advogados e juízes // Available at: https://legis.senado.leg.br/sdleg-getter/documento?dm=9547216&ts=1708613219368&disposition=inline (accessed: 02.03.2024)

[22] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts // Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206 (accessed: 21.05.2024)

[23] The General Data Protection Regulation (Regulation (EU) 2016/679) // Available at: https://gdpr.eu/article-22-automated-individual-decision-making/ (accessed: 10.04.2024)

[24] Verdicchio M., Perin, A. When Doctors and AI Interact: on Human Responsibility for Artificial Risks // Available at: https://link.springer.com/article/10.1007/s13347-022-00506-6#citeas (accessed: 10.04.2024)

and responding to errors and incidents? Who is responsible for identifying errors?

We now propose to consider each issue separately.

a) Who is responsible for algorithm development?

The price of error in the operation of such algorithms is extremely high. Algorithms can be developed and written by both companies and developers. Proper certification of companies and accreditation of developers would be an essential prerequisite. A "Personal license" system may be established to ensure the traceability of specialists' involvement in algorithms development.

b) Is a phased implementation with testing necessary? Are there market authorization procedures?

These two issues are closely interlinked and should be addressed together. Medical devices using AI are subject to mandatory testing and registration. Authorization procedures are available and are generally related to the production of medical devices (hereinafter — MD) and the requirements for their production.

By the Federal Law No. 323 medical devices also include "special software". The criteria for classifying software as a medical device are set out in one the Roszdravnadzor's information letters.[25] Such criteria include the following points:

the software is a computer program or its module regardless of the hardware platform, methods of placing the software and providing access to it;
the software is not an integral part of another MD;
the software is intended by the manufacturer to provide medical care;
the result of the software is interpretated in an automatic mode, including the use of artificial intelligence, and this result influences clinical decision-making.

In the case of qualifying, according to the criteria, MD as software, it is necessary to determine the class of risk to which such MD belongs (see Order of the Ministry of Health No. 4 of 06.06.2012[26]). In accordance with

[25] On Software. Roszdravnadzor letter of 03.02.2020 No. 02I-297/20 // Available at: URL: https://www.garant.ru/products/ipo/prime/doc/73467702/ (accessed: 02.03.2024)

[26] Order of the Ministry of Health No. 4n 06.06.2012 "On Approval of Nomenclature Classification of Medical Devices" (together with Classification of Medical

Section III of Annex 2 to this Order, the following classes of potential risk are distinguished:

Class 1 — low-risk software;

Class 2a — software with medium risk degree;

Class 2b — high-risk software;

Class 3 — the highest-risk software.

According to clause 15.1.1., software with the use of AI technologies belongs to class 3.

After determining the risk class, the developer (manufacturer) must conduct technical and clinical trials regulated by the Russian Ministry of Health.[27] It is worth bearing in mind that these tests are conducted not by the developer, but by third parties that are independently determined by the developer: separately by a testing organization and separately for clinical trials by a medical body. Requirements for medical bodies conducting clinical trials are approved by Order of the Ministry of Health.[28] A list of medical organizations meeting these requirements is available on the website of Roszdravnadzor.[29]

After the tests have been carried out and a full set of documents has been drawn up, the developer must register its AI-based medical device. According to Clause 4 of Article 38 of the Federal Law No. 323, the circulation of registered medical devices is allowed on the Russian territory. Clause 15 of Article 38 establishes a ban on the production of: 1) medical devices not included in the state register of medical devices and organizations (individual entrepreneurs) engaged in the production and manufacture of medical

Devices by Type, Classification of Medical Devices by Class depending on the potential risk of their use") // Available at: URL: https://www.consultant.ru/document/cons_doc_LAW_132477/ (accessed: 02.03.2024)

[27] Order of the Ministry of Health No. 885n of 30.08.2021 "On Approval of the Procedure for Conformity Assessment of Medical Devices in the Form of Technical Tests, Toxicological Studies, Clinical Tests for the Purpose of State Registration of Medical Devices" // Available at: URL: https://www.garant.ru/products/ipo/prime/doc/402937444/ (accessed: 02.03.2024)

[28] Order of the Ministry of Health of the Russian Federation No. 300n of 16.05.2013 "On Approval of Requirements for Medical Organizations Conducting Clinical Trials of Medical Devices and the Procedure for Establishing Compliance of Medical Organizations with These Requirements" // Available at: https://base.garant.ru/70585522 (accessed: 02.03.2024)

[29] List of medical organizations conducting clinical trials of medical devices // Available at: https://roszdravnadzor.gov.ru/services/clinicaltrials (accessed: 02.03.2024)

devices, except for medical devices produced for testing and (or) research; 2) falsified medical devices. Thus, it is prohibited to manufacture medical devices that have not been entered in the state register. Registration of medical devices is carried out by Roszdravnadzor, according to the Resolution of the Government of the Russian Federation No. 1416 of December 27, 2012 "On Approval of the Rules of State Registration of Medical Devices", and information about registered MIs is placed in a special register.[30]

c) Are there regulations for identifying and responding to errors and incidents?

Obviously, it is necessary to develop requirements at the federal level for organizational and technical measures to detect and respond to errors and incidents in organizations using IT solutions based on artificial intelligence. The established requirements should be implemented at the level of each organization that develops and maintains or uses IT solutions based on artificial intelligence.

d) Who is responsible for identified errors?

One of the key principles of decision making, especially in controversial situations, should be the principle of the responsibility of the specific person making that decision. Before registration and production, at the stage of technical and clinical trials, no real decisions should be made. After the registration of a software product or hardware-software complex and its release on the market, there should be mandatory responsibility of an authorized employee at every stage of the system that uses artificial intelligence in its work.

3. National Models

In 2021, AI in healthcare market was worth around 11 billion U.S. dollars worldwide.[31] The Global AI in Healthcare Market was estimated to be 14.41 billion U.S. dollars in 2023 and is expected to reach 51.07 billion U.S. dollars by 2028.[32]

[30] State register of medical devices and organizations (individual entrepreneurs) engaged in the production and manufacture of medical devices // Available at: URL: https://roszdravnadzor.gov.ru/services/misearch (accessed: 02.03.2024)

[31] Artificial intelligence (AI) in healthcare market size worldwide from 2021 to 2030 // Available at: URL: https://www.statista.com/statistics/1334826/ai-in-healthcare-market-size-worldwide/ (accessed: 10.04.2024)

[32] Global AI in Healthcare Market (2023-2028) by Sections, Diagnosis, End user and Geography. IGR Competitive Analysis, Impact of Covid-19, Ansoff Analysis //

Below authors of article offer a closer look at implementation practices in the United States and China, including real-life cases and projects by big tech companies. Jurisdictions were selected based on the worth of AI market in health care and most elaborative framework.

3.1. US Model

It was forecast that the global healthcare AI market would be worth almost 188 billion U.S. dollars by 2030, increasing at a compound annual growth rate of 37 percent from 2022 to 2030.[33]

The US is the leader in terms of AI investment and number of medical databases in the world. US AI in the healthcare market is projected to grow to 51.3 billion U.S. dollars by 2030.[34]

Surely, the number of digitized health data has recently grown, attracting new companies to the health data ecosystem. Technology start-ups, in addition to IT giants such as Google, Apple, and IBM, collect data through the use of apps, their online search platforms, and an ever-expanding array of health technology devices (e.g., sleep trackers, electrocardiograms, smart thermometers, etc.)[35]

For instance, in 2015, Google's DeepMind Health AI entered in partnership with a National Health Services hospital system in the UK and shared 5 years of identifiable medical data on 1.6 million patients. Later, the UK regulator (ICO) concluded that the companies failed to comply with data protection laws, especially considering the sensitive subject matter — health data.

Nuance is an AI-powered voice recognition company that serves healthcare alongside other verticals like security and customer engagement. It works for both telehealth and inperson consultations, and it raised 69.4 million U.S. dollars of investment before being acquired by Microsoft.

Available at: https://www.researchandmarkets.com/reports/5451294/global-ai-in-healthcare-market-2023-2028-by (accessed: 10.04.2024)

[33] Ibid.

[34] US Artificial Intelligence (AI) in Healthcare Market Analysis // Available at: https://www.insights10.com/report/us-artificial-intelligence-ai-in-healthcare-market-analysis/#:~:text=US%20Artificial%20Intelligence%20(AI)%20in%20the%20healthcare%20market%20is%20projected,forecast%20period%20of%202022-30. (accessed: 10.04.2024)

[35] Winter J.S. 2021. AI in healthcare: data governance challenges // Available at: https://jhmhp.amegroups.org/article/view/6448/html (accessed: 10.04.2024)

In the absence in the US a federal law on personal data protection, the main regulation for medical data is encompassed in the Health Insurance Portability and Accountability Act (otherwise known as HIPPA).[36] HIPPA's provisions establish standards for protection of personal health data, or as used in the Act — "protected health information, PHI", collected by covered entities. Covered entities, in turn, include hospitals, clinicians, or insurers, which means that HIPAA protects only this limited area of health information. [Spector-Bagdady K., Armoundas A.A., et al. 2023:1063]

PHI refers to any information in a medical record or data set that can be used to identify an individual and that was created, used, or disclosed in the course of providing a healthcare service, such as diagnosis or treatment. HIPAA and regulations related thereto permit researchers to access and use PHI when necessary. Use of AI in healthcare is not directly mentioned in HIPPA, however as AI becomes increasingly prevalent in healthcare, it is vital to prioritize compliance with the HIPAA. Since HIPAA also establishes the standard for safeguarding medical information in the United States, ensuring the confidentiality, integrity, and availability of all electronic protected health information (ePHI). Given the nature of AI applications, which often involve handling sensitive health data, it is crucial for these applications to adhere to these regulations.

The states have been repeatedly trying to come up with federal privacy law. Today there are separate laws on state level regulating the following: consumer privacy, medical data, genetic data (separate from "general" medical, since it is more sensitive and requires additional oversight), consumer protection. The most recent try is under consideration by the US Congress. The new privacy bill[37] is not unlike previous similar bills and incorporates concepts that are familiar to companies and users from state data protection laws. Most importantly, if enacted, it would repeal and replace all US states data protection laws that have come into force in recent years, such as those of California, Colorado, Virginia, and others.

According to the Bill, personal data ("covered data" in the Bill's wording) includes any information that identifies or is associated, together or in combination with other information, with a natural person or a device that

[36] Available at: https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/combined/hipaa-simplification-201303.pdf (accessed: 10.04.2024)

[37] A Bill to Establish Protections for Covered Data of Individuals, and for Other Purposes // Available at: https://www.commerce.senate.gov/services/files/3F5EEA76-5B18-4B40-ABD9-F2F681AA965F (accessed: 10.04.2024)

identifies or is associated with one or more persons. The Bill also creates a “subcategory” of personal data deemed sensitive, that is subject to additional and heightened requirements. The definition of “sensitive data” is much broader than in other data protection laws and includes: physical or mental health information; genetic information; biometric information.

To summarize, companies that are subject to HIPAA and comply with its rules would be deemed to be in compliance with similar provisions of the proposed Bill. However, if a healthcare company is subject to the Bill, it would also be required to comply with the Bill’s data privacy provisions.

In addition to the above, the US government is actively investing in the private sector. The 2023 Executive Order set a goal to “accelerate grants” awarded to develop AI systems in healthcare.[38] Furthermore, the Order outlines the necessity to improve “healthcare-data quality to support the responsible development of AI tools for clinical care”.[39]

3.2. China Model

According to the recent research, Chinese AI in healthcare market was valued at 0.07 billion U.S. dollars and is expected to grow significantly at a compound annual growth rate (CAGR) of 52.8% from 2020 to 2028.[40] Other statistics shows that the Chinese market was projected to reach 11.91 billion U.S. dollars by 2030.[41] As of 2021, an investment of approximately 60 billion yuan (9 billion U.S. dollars) had only been made in the field of smart medicine in China.

Some of the major players include Google Health, Tencent Trusted Doctors and NERVTEX. Among the world’s top 20 cities in terms of AI companies hosted, Beijing ranks first with 395 companies, and Shanghai,

[38] Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 2023 // Available at: https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/ (accessed: 10.04.2024)

[39] Ibid.

[40] China AI in Healthcare Market Size and Trends to 2031 // Available at: https://www.linkedin.com/pulse/china-ai-healthcare-market-size-trends-2031-cv2ye (accessed: 10.04.2024)

[41] China Artificial Intelligence (AI) in Healthcare Market Analysis // Available at: https://www.insights10.com/report/china-artificial-intelligence-ai-in-healthcare-market-analysis/ (accessed: 10.04.2024)

Shenzhen and Hangzhou are also among the top 20.[42] AI is being deployed across the healthcare industry in areas such as medical imaging devices, diagnostics and drug discovery.[43] .

Over the last 10 years China adopted a range of regulations and guidelines on the topic. It first started with 2016 document of the use of big data for the healthcare industry[44]. The Opinion outlined the following:

the integration of nationwide and provincial healthcare platforms with the online drug tender platform in 2017;

the establishment of a classified open platform for nationwide health information in 2020;

the sharing of basic data on the population, legal persons and geographic location between ministries and regions;

the establishment of 100 clinical data sample centres;

the creation of an electronic health archive for citizens; and

the introduction of a healthcare card.[45]

Then in 2017 the Ministry of Industry and Information Technology issued a three-year AI action and implementation plan[46], fostering the development of smart products in healthcare.

In 2021 State Food and Drug Administration has issued a Guidance for classifying and defining medical software products with artificial intelligence.[47] The document sets out the classification, registration, filing and clinical evaluation requirements for AI medical software products. Moreover, the Guidance defines "medical device data" as information generated

[42] China AI in Healthcare Market Size and Trends to 2031 // Available at: https://www.linkedin.com/pulse/china-ai-healthcare-market-size-trends-2031-cv2ye (accessed: 10.04.2024)

[43] How AI is shaping these three industries in China // Available at: https://www.jpmorgan.com/insights/global-research/artificial-intelligence/ai-transforming-industries-in-China (accessed: 20.05.2024)

[44] Opinion on the Promotion and Standardisation of Application and Development of Big Data for the Healthcare Industry, 2016 // Available at: https://www.gov.cn/zhengce/content/2016-06/24/content_5085091.htm (accessed: 10.04.2024)

[45] China Releases New Opinion to Promote Big Data in Healthcare // Available at: https://cms.law/en/chn/publication/china-releases-new-opinion-to-promote-big-data-in-healthcare (accessed: 10.04.2024)

[46] Available at: https://www.cac.gov.cn/2017-12/15/c_1122114520.htm (accessed: 10.04.2024)

[47] Available at: https://www.nmpa.gov.cn/ylqx/ylqxggtg/20210708111147171.html?type=pc&m= (accessed: 10.04.2024)

by medical devices, thus directly creating two categories of data — already gathered medical data and so-called secondary data that is produced by the devices itself (obviously, based on initial data).

There are 3 categories of key stakeholders in development of AI healthcare:

a) Governmental stakeholders:

The Shanghai Hospital Development Centre (SHDC) is currently implementing a Hospital Link Project with the primary objective of establishing a network of interconnected systems that will facilitate the real-time sharing of data and information between all hospitals in Shanghai.

The Chinese Innovative Alliance of Industry, Education, Research and Application of Artificial Intelligence for Medical Imaging has published a number of consensus documents on topics related to AI. One such document is the 2019 White Paper on Medical Imaging Artificial Intelligence in China, which serves as a reference point for understanding market demands and establishing standardised systems in the field of medical imaging, with the objective of facilitating the introduction of AI products.

b) Academic stakeholders

Academic stakeholders — mostly research institutes in the field of engineering, life science and physical science needed to research to solve problems in biomedicine: Med-X, Shenzhen Institutes of Advanced Technology of the Chinese Academy of Sciences, and Shanghai Institute of Materia Medica.

c) Tech companies

Chinese technology companies, such as Tencent and Alibaba, have begun to recognize the challenges facing the healthcare sector as an opportunity to leverage their consumer-oriented approach, which is focused on meeting the diverse demands of consumers across multiple contexts and channels, in order to capture a new market for digital healthcare solutions. Although physicians, hospitals, and other healthcare providers possess greater experience working within a heavily regulated environment and are able to deliver high-acuity care, these technology companies have the advantage of being able to innovate, scale up, and respond rapidly to market demands, as well as benefiting from a deep understanding of consumers. This has positioned them well to meet the basic healthcare needs of a significant proportion of the population.

Tecent is pursuing a strategy to transform hospitals into Smart Hospitals. This strategy enables patients to schedule appointments with specialists, conduct virtual visits, and access personal health information such as diagnostics, imaging reports, and prescriptions.

Alibaba employs its logistics expertise to facilitate the expedient delivery of pharmaceuticals procured from partner pharmacies within a timeframe of less than 24 hours. In order to gain a greater share of the value chain, Alibaba established its Tmall pharmacy division with the objective of distributing over-the-counter drugs and medical devices to consumers.

A prediction model was constructed by Ping An Technology using case reports from participating hospitals, historical data from regional health authorities and meteorological and environmental statistics. This model was designed to predict flu outbreaks with an accuracy rate of over 90%. The company also created "one-minute clinics" — small rooms or booths where patients enter to connect with an AI doctor that in a few minutes offers a preliminary diagnosis of ailments.[48]

Conclusion

The use of AI in healthcare is coupled not only with benefits, but also with a number of pitfalls. In order to erase these pitfalls, mitigate the identified risks and increase the potential benefits, we are in crucial need for legal regulation.

The above-mentioned measures for risk mitigation can be included in strategic documents on the use of artificial intelligence in healthcare, taken into account in local acts of medical and research institutions, as well as developers of systems using artificial intelligence.

Preparing large verified academic datasets is a lot of work and it is expensive to say the least. For this purpose, BRICS and its health committee (BCICH) or the BRICS Academic Forums could serve as a discussion platform for these issues with further elaboration of international regulation, especially considering recent expansion of BRICS members.

[48] Integrated Healthcare: A New Insurance Model in China // Available at: https://group.pingan.com/media/perspectives/Integrated-Healthcare-A-New-Insurance-Model-in-China.html (accessed: 10.04.2024)

References

1. Alekseeva M.G., Zubov A.I., Novikov M. Yu. (2022) Artificial Intelligence in Medicine. *Mezhdunarodnyi nauchno-issledovatel'skii zhurnal*=International Research Journal, no. 7, pp. 10–13 (in Russ.)

2. Boniolo F. et al. (2021) Artificial Intelligence in Early Drug Discovery Enabling Precision Medicine. *Expert Opinion on Drug Discovery*, vol. 16, no. 9, pp. 991–1007. DOI: https://doi.org/10.1080/17460441.2021.1918096

3. Chen C.W., Walter P., Wei J.C. (2024) Using ChatGPT-Like Solutions to Bridge the Communication Gap Between Patients With Rheumatoid Arthritis and Health Care Professionals. *Journal of Medical International Research Medical Education*. Available at: https://doi.org/10.1186/s13643-022-01939-y (accessed: 16.05.2024)

4. Chung J., Zink A. (2018) *Asia Pacific Journal of Health Law & Ethics*, vol.11, no. 2, pp. 51–80.

5. Da Silva M. et al. (2022) Legal Concerns in Health-Related Artificial Intelligence: a Scoping Review Protocol. *Systematic Reviews,* no. 11. Available at: https://doi.org/10.1186/s13643-022-01939-y (accessed: 18.07.2023)

6. Hodge S.D. (2022) The Medical and Legal Implications of Artificial Intelligence in Health Care — An Area of Unsettled Law. *Richmond Journal of Law & Technology,* vol. XXVIII, issue 3, pp. 405–468.

7. Imameeva R.D. (2021) The Risks of Creation and Functioning of Artificial Intelligence in Medicine. *Vestnik Moskovskogo universiteta imeni S. Yu. Vitte. Seriya 2: Yuridicheskie nauki*=Bulletin of Moscow Witte University. Series 2: Legal Science, no. 1, pp. 33–40. DOI: 10.21777/2587-9472-2021-1-33-40 (in Russ.)

8. McCaffrey P. (2022) Artificial Intelligence for Vaccine Design. *Methods in Molecular Biology*, vol. 2412, pp. 3–13. DOI: 10.1007/978-1-0716-1892-9_1

9. McNair D., Price W.N. (2019) Health Care AI: Law, Regulation, and Policy. In: Artificial Intelligence in Health Care: The Hope, the Hype, the Promise, the Peril. M. Matheny et al. (eds.) Washington: National Association of Medicine, pp. 181–213.

10. Naik N. et al. (2022) Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility? *Frontiers in Surgery,* vol. 9. DOI: https://doi.org/10.3389/fsurg.2022.862322

11. Price W.N. (2017) Artificial Intelligence in Health Care: Applications and Legal Implications. *The SciTech Lawyer,* vol. 14, no. 1.

12. Price W.N., Sachs R., Eisenberg R.S. (2021) New Innovation Models in Medical AI. Law & Economics Working Papers. 47 p.

13. Spector-Bagdady K., Armoundas A.A. et al. (2023) Principles for Health Information Collection, Sharing, and Use: A Policy Statement From the American Heart Association. Circulation, vol. 148, pp. 1061–1069. DOI: 10.1161/CIR.0000000000001173

14. Stöger K., Schneeberger D., Kiesebergc P., Holzinger A. (2021) Legal Aspects of Data Cleansing in Medical AI. *Computer Law & Security Review*, vol. 42. DOI: https://doi.org/10.1016/j.clsr.2021.105587

15. Thomas S. at al. (2022) Artificial Intelligence in Vaccine and Drug Design. *Methods in Molecular Biology*, vol. 2410, pp. 131–146. DOI: 10.1007/978-1-0716-1884-4_6

Information about the authors:

B.A. Edidin — Candidate of Sciences (Law), Deputy Director General for Legal Affairs.
A.V. Bunkov — Direction Manager.
K.V. Kochetkova — Candidate of Sciences (Law), Senior Lecturer.

Contribution of the authors: the authors contributed equally to this article.
The authors declare no conflicts of interests.

The article was submitted to editorial office 10.06.2024; approved after reviewing 28.06.2024; accepted for publication 28.06.2024.

Legal Issues in the Digital Age. 2024. Vol. 5. No. 2.
Вопросы права в цифровую эпоху. 2024. Том 5. № 2.

Research article

УДК: 340
DOI:10.17323/2713-2749.2024.2.80.96

The French Telemedicine System: Challenges, Procedures and Difficulties

Guillaume Rousset

Jean Moulin Lyon 3 University, 1C Avenue des Frères Lumière, CS 78242 69372 Lyon Cedex 08, France,

guillaume.rousset@univ-lyon3.fr

Abstract

In France, telemedicine has been developing rapidly for several years, in response to economic, technical and legal challenges. The aim of the article is to present the broad outlines of the system that has been put in place, from a number of angles. The first deals with the problems to which telemedicine is proposed as a response. These problems are essentially what are known as "medical deserts". Telemedicine is presented as a tool for compensating for the absence or shortage of healthcare professionals in a given area. This system would then promote reliable access to healthcare for the population. While this is a laudable objective, a more detailed analysis casts doubt on whether this result will be achieved. The second angle of this reflection concerns the conditions and procedures for implementing telemedicine. This involves looking at the players involved in order to determine what type of person can be mobilized, in terms of both the type of healthcare professional involved and the type of patient concerned. It is also a question of determining where a telemedicine procedure should be carried out, which shows the diversity of possibilities: where can the booths be set up and, more generally, where should the patient be on this occasion? The third and final angle of this contribution deals with the question of the risks that the practice of telemedicine may generate for patients over and above the benefits that can be imagined.

Keywords

France; healthcare system; access to care; medical deserts; healthcare professionals; technology.

For citation: Rousset G. (2024) The French Telemedicine System: Challenges, Procedures and Difficulties. *Legal Issues in the Digital Age,* vol. 5, no. 2, pp. 80–96. DOI:10.17323/2713-2749.2024.2.80.96

Introduction

For several years now France, like many other countries, has been experiencing a boom in telemedicine [Gallois F., Rauly A., 2019]; [Sauer F., 2011].[1] Boosted by the health crisis linked to Covid-19 [Cayol A., 2020]; [Sebai J., El Manzani Y., 2023], this practice is the focus of much attention, including that of the highest public authorities, with the President of the Republic, Emmanuel Macron, again very recently encouraging its increased development to enable doctors to benefit from more available medical time.[2]

However, essential it may be, it is important to define what telemedicine is at the outset, before examining the various ways in which it can be implemented. This point makes it possible to deal with the necessary conceptual delimitation, but also, and more importantly, to determine the scope of the study. The term telemedicine has been directly enshrined and defined since the law of 21 July 2009 no. 2009-879[3], by article L. 6316-1 of the Public Health Code; it states that "telemedicine is a form of remote medical practice using information and communication technologies. It puts a medical professional in contact with one or more healthcare professionals, with each other or with the patient and, where appropriate, with other professionals providing care to the patient. It makes it possible to establish a diagnosis, to ensure, for a patient at risk, preventive monitoring or post-treatment monitoring, to request a specialist opinion, to prepare a therapeutic decision, to prescribe products, to prescribe or carry out services or procedures, or to monitor the patient's condition". It should be noted, however, that in the texts, certain synonyms for telemedicine are used, such as telehealth or telecare.

[1] Dossier La telemedicine. Revue de droit sanitaire et social 2020/3. Available at: https://documentation.ehesp.fr/ (accessed: 24.05.2023)

[2] Macron E. Conférence de presse de M. Emmanuel Macron, président de la République, sur les priorités du nouveau gouvernement en matière d'école, d'ordre public, d'économie, de natalité, d'égalité des chances, d'écologie, de services publics et de santé. Paris, le 16 janvier 2024.

[3] Law no. 2009-879 of 21 July 2009 portant réforme de l'hôpital et relative aux patients, à la santé et aux territoires, JORF n° 0167 du 22 juillet 2009.

Behind the term “telemedicine”, there are, in reality, a variety of concepts: teleconsultation, in which a medical professional consults a patient remotely; tele-expertise, in which a medical professional seeks the opinion of one or more other health professionals remotely on the basis of their training or expertise; remote monitoring, in which a medical professional interprets health data relating to a patient’s follow-up remotely and makes decisions about the patient’s care; remote medical assistance, that enables a medical professional to provide remote assistance to another healthcare professional to carry out a procedure; medical regulation that enables a healthcare professional to question patients remotely in order to organize emergency medical assistance. Certain issues and problems converge or are even identical for these different components. But these convergences are not systematic, if only in the way they are treated by the media, with attention being focused much more frequently on teleconsultation than on the other elements (even if, to further confuse the issue, teleconsultation is seen referred to as telemedicine...).

Telemedicine is one aspect of e-health and telehealth, as defined by the European Commission [Babinet O., Isnard Bagnis C., 2020].[4] These actions are also part of the implementation of health networks and platforms, as well as digital health territories.

On the basis of these elements, two questions appear to be of primary importance, making it possible to ask, firstly, what problems and issues telemedicine is supposed to respond to, and secondly, how the system should be envisaged. These points will enable us to deal successively with the “why” and the “how”.

1. Telemedicine: a Response to What Problems and Issues?

Traditionally, telemedicine has been presented as a response to what are known in France as “medical deserts”. This notion is more of a political slogan than a legal one, so it is essential to clarify it.

1.1. An Appropriate Response to a Misleading Concept?

Medical deserts are areas where there are not enough healthcare professionals to meet the health needs of the resident population. The idea

[4] Dossier E-santé et nouvelles technologies. Les Tribunes de la santé 2010/4 (n° 29). Available at: https://documentation.ehesp.fr/ (accessed: 16.10.2022)

is simple to grasp, but the expression used poses a problem. First of all, in terms of form, we need to look at the two terms that make up the expression, because both are problematic, as we began to write in health.

On the one hand, the word 'desert' conjures up an idea of aridity, referring to elements that are lacking in considerable numbers, which can be misunderstood in the case we are interested in. In this case, the only aridity that exists is in terms of the density of healthcare professionals, without there necessarily also being demographic, social or economic aridity. To put it another way, the expression "medical desert" conjures up the image of the small country village where there is nothing left: after the school, post office and bakery close, the doctor's surgery closes too. Of course, this is part of the reality [Battesti Ch., Delhomme I., 2023], but medical deserts should not be reduced to this, since they also affect economically or socially dense areas that are rich in population. One example is peri-urban areas where there is a high population density, but a potentially impoverished economic fabric and, in our case, an insufficient presence of healthcare professionals [Schmidt N., 2017].[5] Beyond this, it is quite possible to have areas that are very dense socially and economically, but which are poor in healthcare professionals. In urban areas in general, and even in town centres, it is sometimes very difficult to find a doctor in certain specialties, particularly an ophthalmologist, or a doctor with affordable fees (i.e. in sector 1), without the area being arid demographically, socially and economically [Fichaux J., 2022]; [Schmidt N., 2017]. [6]

In view of all these factors, telemedicine must be a response that can be applied throughout the country, i.e. in all areas that can be described as medical deserts. However, it may be difficult to install telemedicine sys-

[5] Some passages are worth quoting: "After rural areas and big cities, the third type of medical desert is the poor suburbs. There are 40 times more specialists per 10,000 inhabitants in the 7th arrondissement of Paris than in La Courneuve, where there are only 1.6 per 10,000 inhabitants. You might say that, after all, the people of La Courneuve only have to take the RER to get to the 7th arrondissement. But there are many reasons why they can't, not least the fact that most of the doctors in the 7th arrondissement charge higher fees than their peers, which limits access to the most disadvantaged".

[6] Once again, some passages are worth reproducing here: "Medical deserts can also be created where we least expect them. This is particularly the case in very large cities like Paris. Because young doctors find it hard to set up practice in the city, to find a practice of sufficient size at affordable prices, a medical desert is created, which is quite paradoxical given France's medical history. This means that when people move to a new town, they have difficulty finding a referral doctor. They have to wait until they have their first child to get back into the medical monitoring circuit, with a gynecologist who will refer them to a colleague, and so on. But for the many young people who move to Paris, particularly to work, it remains more complicated".

tems in rural areas that are also sometimes digital deserts, i.e. areas where internet access is not good enough or at least not sufficient to enable this technological activity [Durat (de) G., 2017]. If it is the case, it means that telemedicine is not an appropriate response for this type of territory.

Secondly, the use of the word "medical" in conjunction with the term "desert" suggests that only the medical profession is affected, whether in the strict sense (doctors) or in the broad sense (doctors, midwives, dental surgeons). In reality, it is all healthcare professions that can be affected by these inequalities, whether medical, paramedical or pharmaceutical. The most emblematic example, beyond that of doctors, but still within the medical professions, is that of dental surgeons, since the latest negotiations on agreements between the Health insurance and the unions in this profession bear witness to a real awareness of the seriousness of these inequalities and the need to introduce corrective tools, such as selective agreements [Duguet S., 2023]; [Manus J.-M., 2022]. As far as the paramedical professions are concerned, the case of physiotherapists is relevant, this profession being newly subject to a stronger form of regulation in response to these issues of professional distribution.[7] Even pharmacists are beginning to be affected, despite the fact that historically this profession has had the fewest demographic and distribution problems [Nayrac C., 2023]. Nor should we forget, beyond the false anecdotal appearance, the animal health professional who is the veterinarian [Manus J.-M., 2020]; [Chabas C., 2019]. Beyond these professions, of course, the medical profession is the focus of much attention, which is logical since the doctor is the prescriber and his presence (or absence) has major consequences for the activity of other professionals (why would a private nurse set up in an area with no doctors?). However, just because this profession is important in this respect does not mean that it should overshadow other professions. For all these reasons, it is reasonable to argue that the expression "medical desert" is more media hype than science, more a slogan (albeit a clear one) than an academic concept. Author of the article prefers the expression "territorial inequalities in health", which, although not enshrined in legislation, at least has the merit of being more precise, constituting an additional term alongside social inequalities and economic inequalities in health.

[7] Arrêté 21 August 2023 portant approbation de l'avenant n° 7 à la convention nationale organisant les rapports entre les masseurs-kinésithérapeutes libéraux et l'assurance maladie signée le 3 avril 2007, JORF n° 0196 du 25 août 2023, texte n° 28.

Whatever expression is used, it is an important indication that the French system of telemedicine, which will be described later, poses a problem because it focuses solely on the medical professions, i.e. doctors, midwives and dental surgeons. By excluding all the paramedical professions such as nurses and physiotherapists, telemedicine can only solve part of the problem in question.

1.2. An Appropriate Response to an Uncertain Concept?

Secondly, what about the substance of medical desert? It certainly does not correspond to a legal concept. Although it features prominently in the various legal texts in the form of measures designed to combat it, the term is not used explicitly by the legislative or regulatory authorities. Based on the analysis of certain authors, the medical desert could be a false concept, a false problem or, at the very least, a problem whose scale is decreasing.

Firstly, a false notion. This is how one author somewhat provocatively and therefore attractively entitled his article, basing his analysis essentially on two arguments. The first relates to the difficulty of determining with certainty the number of healthcare professionals, in this case doctors, as both sources and results vary, which has an effect on the assessment of their distribution across the country and, even more so, on their inadequacy in terms of the extent of the population's healthcare needs. As the author puts it "[...] it is therefore not possible at this time to make real estimates of the "time available to doctors by specialty", which places a heavy burden on estimates of the real capacity to meet demand" [Carlioz P., 2016: 63]. The second argument is based on the recent nature of the phenomenon. It is stated that it was not until 1992 that the first questions appeared about geographical inequalities in health. On the contrary, it was the totally opposite idea of a "medical plethora" that was vigorously denounced for a much longer period of time, i.e. many decades, from 1900 until the end of the 1980s. This suggests that, historically speaking, medical deserts are a minor phenomenon because they are completely new.

Then, a false problem. For another author, this would be the issue of medical deserts, a concept which he says no one really knows what it covers [Vallancien G., 2012]. Rather than using our own words to describe the analysis, we think it more appropriate to quote certain passages, which deserve to be mentioned especially for their clarity, whatever one may think of them: "So we panic, we alert and we complain, but aren't the French used to ringing the general practitioner's bell for a yes or a no? The same people who complain about not having the ambulance service on their doorstep won't hesitate to drive 25 kilometers to *Ikea* at the weekend! A crazy society that no longer knows what its priorities are. Not seeing the doctor's plaque hanging on the village house is seen as a loss and a risk! And yet there's no evidence that living far from a doctor is dangerous, as the islanders themselves are well aware, as they never make the headlines. So, do we need a doctor in every one of France's 36,000 communes to ensure quality prevention and care? Do we have to satisfy every mayor, every member of parliament and every senator who is up for re-election? When will we dare to carry out a fundamental review of the organization of our country's health cover, going beyond hackneyed recipes, ineffective incentives and ill-conceived coercion? Whatever its relevance, this analysis is worth considering for several reasons. Firstly, because it is completely different from the usual approach to the subject, which is interesting in principle. Secondly, and more importantly, because it allows us to reflect on the degree of access to healthcare that the population should have, but also on the type of criteria to be applied: should we take a spatial approach, asking ourselves how far a citizen should be from a healthcare professional? Or a temporal approach that asks how far away a citizen should be from the first healthcare professional? Or is it a combination of both approaches? These issues raise the question of the criteria for determining medical deserts, i.e. the method to be used, a subject we will address later.

Finally, a shrinking problem. Unlike the previous two, this idea is not the brainchild of an author, but of an institution, which is more surprising and, in fact, even more interesting. This institution is the Direction de la recherche, des études, de l'évaluation et des statistiques (DREES), which is the ministerial statistical service responsible for health and social affairs, under the supervision of several ministries, including the Ministry of Health. This relativisation emerges from several of the studies carried out. The first of these dates from 2010 and provides some very clear information: "According to the Gini index for the population, pharmacies and private GPs are very well distributed across mainland France. They are respectively the 1er and 3^{e} facilities and services for which there is the best match with the population, out of the 137 in the database. Their level of relevance to the population is close to that observed for hairdressing salons (2^{e} facilities) or bakeries (4^{e} facilities). The catchment areas or 'cantons-ou-villes' (for large conurbations) are relatively equal in terms of the density of pharmacies and private GPs".[8] The comparison with hairdressers or bakers is interesting,

[8] DREES. Comptes nationaux de la santé 2009. Ministère de l'emploi, du travail et de la santé. Coll. Etudes et statistiques, 2010, p. 42. Available at: https://sante.gouv.fr (accessed: 12.05.2022)

and echoes the comments made earlier by Guy Vallancien. But is this an observation specific to these two healthcare professions, and does it not reveal a general trend for all professions? Surprisingly, the rest of the DREES report is broadly in line with the first quotes: "The other healthcare professionals generally referred to as primary care (masseurs, physiotherapists, dental surgeons and nurses) are ranked, according to the Gini index, between 14^{e} and 24^{e} of the facilities and services best suited to the population. In this sense, they are similar to local services like banks, supermarkets or restaurants. The direct-access specialist doctors studied (in ophthalmology, pediatrics and gynecology) rank between 56^{e} and 66^{e} among the facilities best suited to the population, which is comparable to secondary schools or gendarmeries".[9] The same report draws a clear conclusion: "this phenomenon is rather limited: either there are few medical deserts, or they are of limited size".[10] The second work by DREES is based on a study published shortly afterwards, in 2012, in which the institution states that "inequalities in the geographical distribution of doctors have decreased significantly over the last 20 years, both between regions and between departments within the same region",[11] which is quite clear from the maps presented.[12] Admittedly, this idea does not appear in later editions of the DREES work, and honesty obliges us to point out that several other documents drawn up by DREES, notably at a later date, do not go in the same direction [Polton D., 2021]; [Lapinte A., Legendre B., 2021]. Nonetheless, it is particularly interesting to read, from a public pen, elements that tend to qualify the conventional discourse on the reality and scale of medical deserts. Is this the sign of a divergent analysis or of a cyclical trend that does not reflect a general trend? This would need to be refined, but it is nonetheless instructive. This work is also taken up in an article written by Olivier Véran, a doctor who was a Member of Parliament at the time of writing, but has since become... the Minister for Health, who clearly asks: "Do medical deserts exist? [Véran O., 2013], concluding that "the term 'sandbox' would seem to be more appropriate than 'medical desert', given their size and organization, which is closer to a leopard print than an extensive desert...".

[9] Ibid.

[10] DREES. Comptes nationaux de la santé 2009, Ministère de l'emploi, du travail et de la santé. Coll. Etudes et statistiques, 2010, p. 42. Available at: https://sante.gouv.fr (accessed: 12.05.2022)

[11] DREES. Les médecins au 1er janvier 2012. Etudes et Résultats n° 796, mars 2012. Available at: https://sante.gouv.fr (accessed: 28.07.2021)

[12] Ibid.

It is therefore possible that the public authorities are promoting a solution in response to a problem that does not quite exist, or at least not in the way we think about. Be that as it may, this is a politically appropriate and low-risk approach, since it makes it possible to achieve an objective (developing access to care) without generating hostility from the professionals who provide this care, since freedom of establishment is respected in this case. It is also interesting to note that this tool is sometimes presented as an (almost) magical solution, as the words of the President of the French-speaking Academy of Telemedicine and e-Health illustrate: "As a matter of urgency, our duty is to launch a CALL for General Mobilization (a "Marshall" Plan) for a great cause by decreeing that any citizen, even in the most isolated places, will have an answer, in less than 20 minutes, to his anguished question "But what's wrong with me?" with care worthy of the name. With the deployment of telemedicine, by combining our medical excellence, our unrivalled capacity for innovation and a strong political will, France can achieve its ambitions. The challenge of equitable access to healthcare is now within reach. It's come to turn action into achievements for the greater good of all. Today and for tomorrow's future generations. We can do it, and we owe it to them! [Alajouanine G., 2022].

Beyond this idealistic or even messianic discourse, it is important to realize that telemedicine can be used in a number of ways to respond to medical deserts [Babinet O., Isnard Bagnis C., 2021: 147]. In fact, this technique addresses the two main difficulties encountered in terms of territorial inequalities, which are spatial and temporal. The spatial difficulty means that a medical desert is characterized by an insufficient supply of healthcare professionals in a given area, who then move to other areas [Durupt M., 2016]. In this case, there are no professionals in the patient's place of residence. The temporal difficulty is different, since here the healthcare professional is well established but is causing patients to take too long to make an appointment because of an imbalance between availability and the number of patients. The problem is therefore temporal rather than geographical, since the healthcare professional is well established in the area, but appointment times are excessive. Telemedicine is an interesting response to these difficulties, since it has the force of abolishing the geographical parameter, since the professional can be consulted wherever he or she is based or wherever the patient lives. So, it doesn't matter if a patient doesn't have any healthcare professionals in the area in which they live, they can access care via telemedicine whatever happens. Telemedicine can also solve the problem of time, since various financial incentives enable healthcare professionals to offer a range of services in addition to those available in the traditional way.

2. Telemedicine: What Methods, What Risks?

If the French telemedicine system is to be analyzed as effectively as possible, it is essential to ask two questions: how should the system be set up? and what risks might the use of telemedicine generate?

2.1. Methods of Implementation

There are two important dimensions to these procedures, that are not exhaustive [Bourdaire-Mignot C., 2011]. The first concerns the players involved, in order to determine what type of person may be involved, corresponding to the "who". It covers several dimensions, relating successively to the type of healthcare professional involved and the type of patient concerned.

As far as the type of healthcare professional is concerned, there is technically a wide variety of possibilities, potentially involving both medical professionals (doctors, midwives, dental surgeons) and paramedical professionals (nurses, physiotherapists, speech therapists, chiropodists, etc.) or pharmaceutical professionals, especially dispensing pharmacists. Of course, the relevance of telemedicine varies according to the type of procedure carried out and therefore the type of professional involved. For example, a lot of technical procedures cannot be carried out without the physical presence of the patient and the technical mobilization of the body and the organ in question. However, the French authorities have taken a restrictive approach, restricting the scope of teleconsultation to the medical professions. It means, a contrario, that the paramedical professions, as well as the pharmaceutical professions, are excluded from the scheme and that no telemedicine act can be performed for them. Should this be seen as an exclusion justified by the technical factors explained above, or as a political choice aimed at not developing telemedicine too extensively right away, given the reluctance that this technique may possibly arouse? Everyone will answer according to analytical grid.

In addition, for healthcare professionals authorized to carry out their activities using telemedicine, it is important to ask whether this practice method can be used for all medical procedures or just some. Can or should certain procedures be excluded? Taken to the extreme, the case of certain countries such as Australia shows that telemedicine could be envisaged (but ultimately not adopted) in the context of medical assistance in dying, which is legal in that country.[13] On a related but distinct subject, the question has also arisen as to whether certain prescriptions can be limited when they are carried out as part of a consultation on TV. Two cases were studied. The first, which was finally adopted as part of a bill, would have enabled the Minister for Health to limit or prohibit the prescription of certain drugs by telemedicine in the event of a supply shortage. This approach was censured by the French Constitutional Council, that ruled that these provisions violated the principle of equality before the law in that they could have had the effect of "depriving a patient of the possibility of being prescribed a medicine necessary for his or her state of health on the sole grounds that he or she has consulted a doctor remotely" [Cordier C., 2023]. The second case, which has been retained, limits the prescription of work stoppages to a maximum of three days (initial stoppages and any extensions) when this takes place via teleconsultation, essentially if the prescriber is not the attending physician [Law no. 2023-1250].[14] This is justified by the fact that if a patient requires a longer period of leave from work or its renewal, an in-depth face-to-face examination is necessary to ensure that the correct diagnosis is made.

The second question relating to the "who" concerns the type of patient who can benefit from teleconsultation. Should this dematerialized procedure be envisaged for all types of patients, in other words, for all types of consultation? Or should we restrict this type of practice to certain health needs, certain consultations and therefore certain patients? The French public authorities have opted to open the door to all types of consultation as long as they fall within the remit of the medical professions mentioned above, i.e. doctors, midwives and dental surgeons.

Once these two issues have been addressed, another fundamental question concerns the relationship between these two players through the way in which telemedicine is carried out. Does a consultation on TV, for example, involve a consultation carried out solely remotely, with no physical contact and no professional presence with the patient? Or should a healthcare professional act as a technical intermediary between the patient and the healthcare professional? This intermediary professional would be physically close to the patient, for example to direct the camera, adjust the technical procedures and, more generally, ensure that the patient gets to grips with the

[13] Australie: l'utilisation de la téléconsultation dans le cadre de l'»aide volontaire à mourir» est illégale. *Gènéthique*, 30 novembre 2023.

[14] Law no. 2023-1250 of 26 December 2023 de financement de la sécurité sociale pour 2024, JORF n° 0299 du 27 décembre 2023, texte n° 1.

gondola for an effective consultation. The answer to this question is a major one, because if we accept the presence of a professional as an intermediary, it means that telemedicine is not totally disembodied, that it is not just an exchange at a distance. In this case, telemedicine is another way of being in contact with a healthcare professional, without excluding his or her physical presence, at least in a general and absolute sense. In France, the model adopted is not to choose between these possibilities, but rather to retain both. It is therefore possible to have the consultation both with the presence of an intermediary healthcare professional and without any presence of this type, the meeting being totally virtual between the professional and the patient.

The other important dimension is no longer a question of "who" but of "where". Determining where a telemedicine procedure is to be carried out shows the diversity of possibilities: where can the booths be set up and, more generally, where should the patient be for the procedure? This could be at the patient's home, a potentially pleasant and comfortable situation for the patient, since he or she does not have to travel, solving the problems of transport and time available. However, this does raise the fundamental question of the quality of the patient's computer equipment and Internet connection, as well as the conditions in which the consultation takes place, for example in terms of the brightness required for the professional to capture images and interpret elements correctly, particularly in a dermatology consultation.

The second possibility corresponds to a place of care. Since there are no doctors' surgeries, TV consultation could, as a matter of principle, be carried out, for example, in a health establishment, provided that it is located in the patient's place of residence. The idea of a paramedical practice seems unreasonable, since it is difficult to imagine the presence of a paramedical practice in an area where there is no paramedical practice (why would a nurse, for example, set up practice in an area where there is no doctor, i.e. where there is no prescriber for the procedures she has to carry out?) The only credible location under these conditions is the dispensing pharmacy, if there are any in this area, which, as we have said, is by its very nature subject to territorial inequalities in health.

The third possibility is not a place of care, but a place of public service in the sense, for example, of an administration [Renaudie O., 2013]. This possibility is not frequently mentioned, but it seems an interesting one, given that the territorial coverage of public services, while not perfect, is not yet in the majority of cases deficient. In this category, the avenue most frequently put forward concerns the mobilization of rail and/or bus stations. While not an administration in themselves, they are a form of public service that could be in keeping with the spirit of access to healthcare in under-serviced areas. The geographical distribution of railway stations suggests to the promoters of these solutions that they would be a suitable location for providing the population with good access to telemedicine. However, it should not be forgotten that there are some areas of mainland France, albeit in a minority, that have no railway stations at all, as shown by the case of the Ardèche department (it does have bus stations).

The last avenue is the one that raises the most questions, with the case of retail outlets. Several projects have been envisaged based on the installation of telemedicine booths in supermarkets, especially the Monoprix shop. This would mean the coexistence of places dedicated to mass consumption, operated by profit-making companies on the one hand, and, on the other, gondolas whose purpose is to provide access to care, prevention and treatment, through procedures funded by the social security system with a view to public health. Everyone will judge the relevance of this coexistence, but it is certain that the French medical association (Ordre National des Médecins) saw it as highly problematic and contrary to various provisions of the code of medical ethics, favoring a form of commercialization and consumerization of healthcare, an analysis we share.

Whatever answer is chosen, the question will then arise as to the procedures to be followed, particularly in terms of the players involved, referring back to the question already addressed of whether or not it is necessary to have an intermediary healthcare professional who, in the telecab, will assist the patient so that the teleconsultation can take place under the best possible conditions. In France, everything is still open to discussion, but a major public health agency, the Haute Autorité de Santé, has established four guidelines: to ensure the quality, continuity and safety of care; to promote access to care, by ordering care that complements face-to-face care; to preserve face-to-face care; and to avoid any commercial abuses. On this basis, the agency makes three recommendations in this respect: the location of telemedicine equipment must guarantee accessibility, quality and safety of care; the operator must ensure that the equipment functions properly; and a person responsible for the telehealth equipment must be present on site.

2.2. The Risks Involved

First of all, it is necessary to consider the risks that telemedicine-based care may generate for patients, over and above the benefits in terms of access to care that we have imagined earlier in this contribution. To help us think about this, it is useful to refer to a study carried out by researchers at several British universities (Oxford and Plymouth, with the support of the Nuffield Trust), which indicates that teleconsultations can expose patients to potentially fatal errors or delays in diagnosis [Payne R., Clarke A., Swann N. et al., 2023]. Let's be clear about the work we are doing: we are obviously not saying that telemedicine creates risks and that conventional care does not generate any, since, of course, any form of care involves various dangers. On the contrary, the aim is to demonstrate that telemedicine creates specific or increased risks. This is the case with the elderly and younger people who, because of their greater communication difficulties, are most exposed to a risk linked to an inadequate patient/doctor relationship and inappropriate information gathering.

In France, this theme is clearly important enough to become the subject of a question put by a member of parliament to the Minister of Health.[15] The comments are enlightening and deserve to be reported: "Insurers have in fact increased their civil liability premiums by explicitly mentioning the increase in claims caused by teleconsultation. According to studies carried out by the insurance industry, remote appointments present an increased risk of the practitioner being called into question, necessitating the intervention of the insurer. The most common grounds for dispute include underestimation the seriousness of the patient's state of health from a distance and prescribing inappropriate treatment. While the aim was to improve the supply of healthcare, the increase in premiums risks weakening the professionals who engage in teleconsultations by increasing their costs. Has your ministry identified this problem and, if so, are any measures planned to prevent teleconsultation from affecting doctors' insurance premiums?". The response from then Minister for Health, Agnès Firmin Le Bodo, was reassuring, but the question remains.

This is also what emerges from a recent study carried out in France by Agence de presse médicale, which one commentator analyses severely: "A survey carried out by the health insurance scheme in the Ile-de-France region reveals significant differences between the practices of GPs in private practices and those of doctors working for platforms dedicated to teleconsultation", showing "disproportionate prescriptions, with a large number of

[15] Lemoine P. Question n° 420 relative à l'assurance des professionnels de santé. XVIe législature, session ordinaire de 2023—2024, première séance du mardi 28 novembre 2023.

consultations billed illegally with night or Sunday surcharges”, but also that “doctors, unable to carry out a proper clinical examination, play it safe and prescribe 2.5 times more antibiotics than GPs in their practices”. What’s more, “almost 20% of consultations are followed by a new consultation during the week, oben in person” [Prudhomme C., 2023].

Conclusion

Of course, telemedicine should not be reduced to a source of risks [Vioujas V., 2015] since, according to another study, telemedicine can be virtuous by enabling the development of an activity that is more respectful of the environment. In fact, compared with activities whose traditional organization contributes significantly to energy consumption and waste production, telemedicine reduces the need for patients to travel for consultations that have significant positive effects [Ravindrane R., Patel J., 2022]. However, there is a great deal at stake here, particularly in terms of the responsibility of both healthcare professionals and public authorities, a subject that would justify an article in its own right [Grynbaum L., 2011]; [Corgas-Bernard, 2014]; [Paley-Vincent C., Gombault N., 2011].

References

1. Alajouanine G. (2022) Lettre ouverte au Président de la République pour un Grand Chantier «Zéro Déserts Médicaux!» Académie francophone de télémédecine et de e-santé, 24 janvier.

2. Babinet O., Isnard Bagnis C. (2020) La e-santé en question(s). Rennes: Presses de l’EHESP, 144 p.

3. Babinet O., Isnard Bagnis C. (2021) Les déserts médicaux en question(s). Rennes: Presses de l’EHESP, p. 147.

4. Battesti Ch., Delhomme I. (2023) L’accès aux soins se dégrade dans les zones rurales. *Insee Flash Pays de la Loire,* n° 137, 14 mars.

5. Bourdaire-Mignot C. (2011) Téléconsultation: quelles exigences? Quelles pratiques ? *Revue de droit sanitaire et social,* p. 1003.

6. Carlioz P. (2016) La fausse notion de ‘’désert médical’. *Review Generale de Droit Medical,* n° 58, mars 2, p. 61.

7. Cayol A. (2020) L’adaptation du système de santé français face à la pandémie de Covid-19 par le développement des usages de la télémédecine. *Droit, Santé et Société,* no.1, p. 25.

8. Chabas Ch. (2019) Les vétérinaires ruraux, espèce en voie de raréfaction. *Le Monde*, 7 mars.

9. Cordier C. (2023) Les sages censurent une dizaine d'articles du PLFSS mais valident les objectifs de dépense. *Hospimedia*, 22 décembre.

10. Corgas-Bernard C. (2014) Responsabilité civile médicale et nouvelles pratiques numériques: l'exemple de la télémédecine. *Les Petites affiches*, n° 162-163–164, p. 27.

11. Duguet S. Déserts médicaux : la régulation de l'installation des dentistes inquiète les médecins. *Public Sénat*, 25 juillet 2023.

12. Durat (de) G. (2017) Déserts médicaux et numériques: la double peine. *Les Echos*, 23 août.

13. Durupt M., Bouchy O. et al. (2016) La télémédecine en zones rurales : représentations et expériences de médecins généralistes. *Santé Publique*, no. 4, p. 487.

14. Fichaux J. (2022) Les déserts médicaux s'étendent de plus en plus en milieu urbain. *Club santé social*, 26 juillet.

15. Gallois F., Rauly A. (2019) Le développement de la télémédecine au prisme des référentiels de politiques publiques. Une cartographie de trois nations européennes. *Politiques & management public,* no. 3, p. 275.

16. Grynbaum L. (2011) La responsabilité des acteurs de la télémédecine. *Revue de droit sanitaire et social*, p. 996.

17. Lapinte A., Legendre B. (2021) Renoncement aux soins: la faible densité médicale est un facteur aggravant pour les personnes pauvres. *Études et résultats,* n° 1200.

18. Manus J.-M. (2020) Après les déserts médicaux, les déserts vétérinaires?. *Revue Francophone des Laboratoires,* n° 520, mars, p. 19.

19. Manus J.-M. (2022) Les déserts médicaux concernent aussi les soins dentaires. *Revue Francophone des Laboratoires*, n° 539, fév., p. 17.

20. Marié R. (2024) Du renforcement de la prévention à l'amélioration de l'accès aux soins. *Droit social*, n° 3, p. 230.

21. Nayrac C. (2023) L'ordre des pharmaciens pointe une démographie fragile et se pose en force de proposition. *Hospimedia*, 11 juillet.

22. Paley-Vincent C., Gombault N. (2011) Télémédecine, réglementation et responsabilité. *Revue Responsabilité*, déc., n° 44, p. 10.

23. Payne R., Clarke A., Swann N. et al. (2023) Patient safety in remote primary care encounters: multimethod qualitative study combining Safety I and Safety II analysis. *British Medical Journal of Quality* & *Safety*, 28 November. DOI: doi: 10.1136/bmjqs-2023-016674.

24. Polton D., Chaput H., Portela M. (2021) Remédier aux pénuries de médecins dans certaines zones géographiques — Les leçons de la littérature internationale. *Les dossiers de la DREES,* n° 89.

25. Prudhomme C. (2023) Téléconsultations: dire stop aux cabines de télémédecine. L'Humanité, 18 décembre.

26. Ravindrane R., Patel J. (2022) The environmental impacts of telemedicine in place of face-to-face patient care: a systematic review. *Journal of Medical Internet Research,* vol. 9, no. 1, pp. 28–33.

27. Renaudie O. (2013) Télémedecine et téléservice public. *Revue française d'administration publique,* no. 2, p. 381.

28. Sauer F. (2011) Europe et télésanté. *Revue de droit sanitaire et social*, p. 1029.

29. Schmidt N. (2017) Entretien — Les déserts médicaux se créent aussi là où on ne les attend pas. Entretien avec Emmanuel Vigneron, géographe. *Observatoire des inégalités,* 18 août.

30. Sebai J., El Manzani Y. (2023) Adoption de la télémédecine par les professionnels de santé publics français pendant la pandémie de COVID-19. *Management and Prospective*, no. 2, p. 13.

31. Vallancien G. (2012) Il faut en finir avec le faux problème des déserts médicaux. *Les Echos*, 1er et 2 juin.

32. Véran O. (2013) Des bacs à sable aux déserts médicaux : construction sociale d'un problème public. *Les Tribunes de la santé, no.* 2, n° 39, p. 77, 79.

33. Vioujas V. (2015) La télémédecine: entre expérimentations réussies et généralisation au ralenti. *Revue de droit sanitaire et social,* p. 681.

Information about the author:

G. Rousset — Senior Lecturer, Director, Centre for Research in Law and Management of Health Services.

The article was submitted to editorial office 08.07.2024; approved after reviewing 11.07.2024; accepted for publication11.07.2024.

Legal Issues in the Digital Age. 2024. Vol. 5. No. 2.
Вопросы права в цифровую эпоху. 2024. Том 5. № 2.

IT, Industries, Law: Media

Research article
УДК: 342
DOI:10.17323/2713-2749.2024.2.97.112

Regulating Digital Era: A Comparative Analysis of Policy Perspectives on Media Entertainment

Reeta Sony A.L.[1], Shruti Chopra[2]

[1, 2] Centre for the Study of Law and Governance, Jawaharlal Nehru University, New Delhi 110067, India,

[1] reetasonyjnu@gmail.com

[2] shruti.jmi@gmail.com, https://orcid.org/0000-0002-3642-3326

Abstract

The rapid proliferation of digital media platforms has democratized content creation and distribution, enabling a vast spectrum of voices to be heard. It has brought about a significant shift in media entertainment landscapes worldwide, with India being a prominent case study due to its vast and diverse content consumption patterns. The massive content on the Internet has also raised concerns regarding misinformation, copyright infringement, and cultural sensitivity. Therefore, in the context of media entertainment, the regulation of the digital era presents a vast complex array of challenges for policymakers. Thereby, it analyzes the regulatory challenges and policy perspectives; addressing how India is navigating the complexities introduced by digital technologies. The study outlines India's current regulatory framework including legislative measures. Apart from this, the paper contains exploration of challenges of balancing free speech with societal norms in a country characterized by its cultural pluralism. The authors of the article argue that rational regulation is able to help to prevent the spread of misinformation, protect national security, and ensure privacy. It can play a pivotal role in promoting national integration by fostering unity, preventing communal tensions, and ensuring equal representation. To achieve the objectives, the paper analyzes three case studies — Swami Ramdev v. Facebook, Inc.*

(2019), the Tandav Controversy (2021), and the Tiktok ban for privacy and security concerns (2020). In the later section, the authors analyze and compare the regulatory framework of different states including India, the United States, European countries, Australia, and China. In the end, the paper summarizes the need for changes in the regulatory framework and also recommends policy measures that may be implemented to safeguard the consumers' interest, preserve cultural values, and ensure the integrity of content.

Keywords

freedom of speech and expression; media consolidation; over-the-top services; consumption; digital media; regulatory framework.

For citation: Reeta Sony A.L., Shruti Chopra (2024) Regulating Digital Era: A Comparative Analysis of Policy Perspectives on Media Entertainment. *Legal Issues in the Digital Age,* vol. 5, no. 2, pp. 97–112. DOI:10.17323/2713-2749.2024.2.97.112

Introduction

The history of media has evolved significantly from one-way dissemination of information to a dynamic two-way communication channel, and with the emergence of the Internet, it has undergone radical transformation. The shift from the traditional form of media (radio or television) to newest media has remarkably transformed the milieu of media entertainment globally. The Internet is incorporated into the lives of humankind similar to the radio and television before. Broadly defining, the media can refer to tools, platforms, and channels that can be used to create, produce, and share knowledge among society. However, digital media may be defined as the digitized content (text, graphics, audio, and video) that can be transmitted over Internet or computer networks. Digital media is considered as part of the convergence between interactive media, online networks, and existing media forms [Flanagan A.J., Metzger M.J., 2008]. In today's world, digital media has given access to society not only to consume content on different platforms but also to produce and disseminate the content extensively. It has opened up new avenues for content development, consumption, and delivery. The Internet users as "produsers", that is, those who are both users and producers of digital media, and coined the term "produsage" to describe this blend of production and usage in digital media environments [Bruns A., 2007]. And, in a country like India, with its vast population, this transformation to new forms of media is even more pronounced because of its rich cultural tapestry.

Having known that Internet subscriber base in India has overstepped 900 million comprising wireless subscriptions amounting to approximately 1143.93 million persons involved is transforming present-day India into of the largest online markets globally.[1] Due to the growing Internet penetration and the proliferation of smartphones, the growth in media consumption patterns has also seen a rapid increase. In India, digital platforms are the predominant medium for entertainment, news, and social interaction. Therefore, the unprecedented growth of digital media has presented unique challenges and opportunities for regulatory frameworks. The increase in digital media consumption patterns, encompassing streaming services, social media, watching television, listening to music and so on and so forth are be analyzed in Figure 1 given below.

It may be said that digital platforms have not only transformed consumer behavior but also posed significant regulatory challenges. Issues such as data breaches, misinformation, digital monopolies, and the need for content moderation have become increasingly prominent, necessitating a robust regulatory response [Pickard V., 2019]. The Government of India has responded by implementing a range of policies aimed at overseeing these digital landscapes. These policies seek to balance the dual imperatives of promoting technological innovation and ensuring consumer protection, data privacy, and national security.

In the stated context, the study presented examines the state of media regulation in India during the digital era, with an emphasis on the interaction between legislative frameworks and technological improvements. It looks at the efficiency of the laws that are currently in place, the challenges posed by digital media, and the potential paths forward to ensure that the growth of digital platforms contribute positively to societal, cultural, and economic dimensions. By analyzing policies, legal frameworks, and industry practices, this study aims to provide a comprehensive overview of India's approach to regulating the digital media sector.

Figure 1 illustrates a survey on media consumption behaviors, displaying both the percentage of respondents who reported increased consump-

[1] Statista 2024. 29 April. Change in media consumption in India 2022 by activity. Available at: https://www.statista.com/statistics/128061/india-change-in-media-consumption-by-activity (accessed: 16.05.2024)

tion over the past twelve months and those who intend to increase their consumption in the next twelve months. The figure depicts the dominance of digital media in India. Music and video streaming services like Spotify, Pandora, Netflix, and Amazon Prime are experiencing a surge. Around half of the respondents have increased their consumption of these services, and similar numbers intend to continue doing so. This trend underscores the shift away from traditional media formats towards on-demand digital content. In a nutshell, it may be concluded that the data suggests a strong consumer pivot towards on-demand, customizable media consumption facilitated by digital platforms. Traditional forms, while still holding significant sections of the market, show less dynamic growth, pointing to a potential area of concern for industries reliant on these formats.

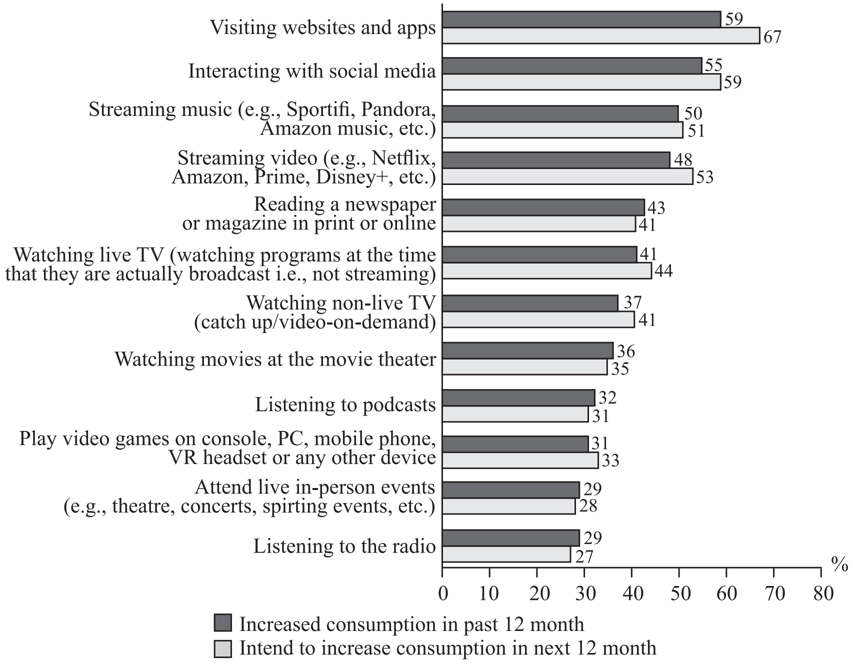


Fig. 1. Digital Media Consumption Patterns 2022 in India by Activity.

Source: Statista.com[2]

On the one hand, the digital revolution in India paves the way for remarkable opportunities in media and entertainment; on the other, it intro-

[2] Statista 2024. 29 April. Change in media consumption in India 2022 by activity. Available at: https://www.statista.com/statistics/128061/india-change-in-media-consumption-by-activity (accessed: 16.05.2024)

duces a spectrum of challenges that demand robust regulatory responses. The establishment of regulatory frameworks in the digital age is critically dependent on the complex array of opportunities and problems presented by India's media and entertainment industry's digital transformation. With the speed technology is developing, the Indian government is putting efforts to address critical issues like data privacy, content regulation, and the digital divide while leveraging the potential of digital media for innovation and economic progress. This section explores these challenges and opportunities, providing insights into how effective regulation can also increase the positive effects of digital media on society.

The rapid advancement in technology renders the need for regulating the digital era in India's media and entertainment sector. Regulations frequently find it difficult to keep up with the rapid evolution of digital technology, which may induce gaps in monitoring and enforcement. This is particularly relevant in fields like machine learning and artificial intelligence, where emerging technologies can surpass current laws. Ensuring data security and privacy in an environment where personal data is constantly shared puts forth a major challenge. Even though India is making progress with new regulations such as the Digital Personal Data Protection Act of 2023[3], the implementation and enforcement of these rules are still fragmented and inadequate, leaving customers open to fraud and security breaches.

1. Review of Literature

The digital era has significantly transformed the media entertainment landscape, necessitating a robust regulatory framework to address new challenges and opportunities. The need for regulations on digital media entertainment is extensively documented by various scholars reflecting the complexities and challenges of this rapidly evolving landscape. Some of the issues that have drawn attention and emphasized the importance of regulatory framework include: data privacy, content moderation, and the influence of digital platforms on public discourse.

The exponential growth of digital platforms and their massive usage by people around the world has heightened concerns about data privacy and the protection of personal information. The General Data Protection

[3] India.Ministry of Electronics and Information Technology. Digital Personal Data Protection Act, 2023. Available at: https://www.meity.gov.in/writereaddata/files/Digital%20Personal%20Data%20Protection%20Act%202023.pdf (accessed: 10.05.2024).

Regulation (GDPR) in the European Union serves as a landmark example, setting stringent standards for data collection, processing, and storage. The GDPR is crucial for ensuring that digital platforms adhere to high standards of user privacy and data security [Voigt P., von dem Bussche A., 2017].

On the other hand, content regulation and moderation have also become pivotal in the digital age, as online platforms gain influence over public discourse and societal norms. Three regulatory approaches are normally observed — Self Regulation and Platform Policies, Government Regulation and Legal Frameworks, and Hybrid Models. Many digital platforms rely on self-regulation, establishing internal policies and algorithms to manage content. The platforms like Facebook[4] and YouTube develop community standards and use automated systems to identify and remove harmful content [Gillespie T., 2018]. These mechanisms are essential for managing vast amounts of user-generated content but also raise concerns about transparency and accountability. Some countries adopt varying levels of government intervention in content regulation. The European Union's Audiovisual Media Services Directive (AVMSD) provides a harmonized legal framework that ensures content moderation across member states, protecting minors and preventing hate speech. Similarly, the German Network Enforcement Act (NetzDG) mandates social media platforms to swiftly remove illegal content, imposing fines for non-compliance [Tworek H., Leerssen P., 2019]. The Hybrid model combines self-regulation with government oversight. Australia's Enhancing Online Safety Act exemplifies this approach, where the private regulators have the powers to require platforms to remove content ensuring content standards while providing a government-led mechanism for addressing severe cases of online harm [Flew T., Martin F.R., 2022]. Content moderation also leads to various challenges like maintaining a free speech balance with the need to prevent harm. It is argued that platforms must curb harmful content, overly stringent regulations can stifle free expression [Garton Ash T., 2016]. This tension is evident in the varied global responses to content moderation, where cultural and political contexts significantly influence regulatory frameworks. The second challenge is algorithmic moderation and bias. Automated content moderation systems are prone to biases and errors [Noble S.U., 2018]. The algorithmic biases can disproportionately target marginalized communities, exacerbating existing social inequalities. This underscores the need for

[4] Meta Platforms Inc. is recognized in Russia as an extremist organization and banned. The social networks Facebook and Instagram belonging to it are banned in Russia.

transparent and accountable moderation practices that consider the ethical implications of algorithmic decision-making. The third and most prevalent challenge is the global platforms and the local norms. The global nature of digital platforms poses a challenge for content regulation, as local norms and laws vary widely. The platforms navigate these complexities, often leading to inconsistent enforcement of content standards [Suzor N., 2019]. This variability can undermine the rationality of content moderation policies and erode user trust.

The monopolization of media ownership by a few powerful moguls also poses significant challenges to market competition, content diversity, and cultural influence [Baker C.E., 2001]. Some of the most prominent media conglomerates, own major assets across various segments of the entertainment industry. The key players include Disney, Comcast, and Warner Bros. Discovery. Disney's acquisition of 21st Century Fox in 2019 has significantly expanded its media empire, including film studios, television networks, and streaming services like Hulu and Disney+ [Vogel H. L., 2020]. This acquisition has positioned Disney as a dominant player in the media and entertainment sector, influencing both content production and distribution [Lotz A.D., 2021]. Similarly, Comcast's extensive holdings include television networks, film studios, and the streaming service Peacock. Its vertical integration, combining content creation and distribution, exemplifies the monopolistic tendencies in the media industry [Napoli P.M., 2001]. Discovery manages a portfolio that includes HBO Max, Discovery+, and numerous television networks and film studios.

Such consolidation practices have significant implications on market control and competition, indeed. It further reduces competition and increases the influence of a few large players in the media market by creating high barriers to entry for smaller players [McChesney R.W., 2015]. This control allows media moguls to dictate terms in content licensing, advertising rates, and consumer pricing, often leading to higher costs for consumers and reduced market dynamism (Bagdikian B.H., 2004]; [Noam E.M., 2015]. It may also have a significant impact on content diversity. The consolidation by media moguls often leads to homogenization of content, where diverse and independent voices are marginalized. The focus on profit maximization drives these conglomerates to produce content that appeals to the broadest audience, often at the expense of niche or culturally specific content. This homogenization undermines the diversity of viewpoints and cultural representation in media [McChesney R.W., 2008]; [Doyle G., 2002]. Another important implication is cultural influence, the extensive reach and influence of media moguls enable them to shape public discourse and cultural norms. By controlling major news outlets, film studios, and television networks, these conglomerates can influence public opinion and political agendas. This concentration of power raises concerns about media pluralism and the role of media in a democratic society [Curran J., 2011]; [Zuboff S., 2019].

The literature on regulating digital media entertainment reveals significant variations across different regions, highlighting diverse approaches and challenges. The European Union's robust framework, including the Digital Services Act (DSA) and the General Data Protection Regulation (GDPR), emphasizes user protection and accountability of digital platforms. In contrast, the United States maintain a more liberal stance, prioritizing freedom of expression with sector-specific regulations like the California Consumer Privacy Act (CCPA) addressing privacy concerns. China's stringent regulatory environment is characterized by heavy state control through laws such as the Personal Information Protection Law (PIPL) and strict content censorship managed by the Cyberspace Administration of China [Priya V.B.A., 2023]. India and Australia have selected a balanced approach, implementing moderate to high regulation levels with acts like the Digital India Act 2023 and the Online Safety Act 2021 to ensure user privacy and content accountability [Narayanan R., 2024]. These varying regulatory frameworks reflect the complexities and evolving nature of digital media governance globally. A comparative analysis has been made in the later section of the study.

2. Need for Digital Media Regulation

Digital media has become the primary source of information these days. Therefore, the need for digital media regulation is crucial for several reasons, particularly in balancing the growth of digital media platforms aligning with societal norms, privacy, and security. It may be said that digital media can significantly influence public opinion, societal norms, and cultural values. Thus, the non-regulated media content may spread hatred and violence and become a source of misinformation, affecting societal norms and values. On the contrary, the regulations may facilitate the constructive discourse and exchange of cultures and values. Another concern that is seeking importance in the context of the need for digital media regulation is privacy and protection from unauthorized access and misuse. The Internet is a vast pool of information and users on the Internet enter their personal information on various websites. Thus, securing the personal information of the users becomes a paramount concern while addressing privacy issues. The regulations may enforce companies to strictly adhere to privacy laws by implementing stringent data protection measures. Moreover, misleading advertisements, fraud, and other deceptive practices often harm customers. Thus, considering the need to protect the customers and users on the digital platform, it becomes essential to introduce the practice of responsible innovations that should respect user rights, privacy, and ethical considerations. At the same time, the regulatory regime should also maintain the balance between the freedom of speech and expression and harmful content on digital platforms.

Societal Norms	Need: Public Safety, Societal Harmony Risk: Hate Speech, Violence, Misinformation
Privacy Protection	Need: Stringent Data Protection Measures Risk: Misuse of Personal Data Information
Security Measures	Need: Protection of Digital Infrastructure and Sensitive Information Risk: Hacking, Phishing, Malicious Activities
Responsible Innovation	Need: Innovation that Respects User Rights, Privacy, and Ethical Considerations Risk: Unethical Activies
Consumer Protection	Need: Protect Consumers by Ensuring that Digital Services & Products Meet Certain Quality and Safety Standards Risk: Misleading Advertisements, Fraud, and Other Deceptive Practices
A Balance Between the Free Speech and Harmful Content	Need: Balance the Right to Free Speech & to Prevent the Dissemination of Harmful or Illegal Content Risk: Conduits of Harmful Activities

Fig. 2. Need for Regulation on Digital Media[5]

The successful regulatory framework can not only mitigate risks but also enhance the benefits of digital media for society. By striking the right balance between regulation and freedom, India can harness the power of digital media for the greater good of society and national cohesion. It has been

[5] Compiled by the Author using data from Flew and Martin (2022) for societal norms, Belli and Zingales (2019) for data protection, Gillespie (2018) for hate speech and harmony; Anand and Brass (2021) for responsible innovation; and author's own analysis for the balance between free speech and harmful content.

very well argued that media plays an important role in preventing the spread of inflammatory content and misinformation, which can fuel communal tensions and violence. Furthermore, regulating digital platforms may ensure the protection of diverse cultural expressions and equitable representation of various communities, fostering a sense of inclusion and unity. The role of regulation in promoting educational and informative content, it can also enhance digital literacy and civic engagement, strengthening national identity and solidarity. The strategic implementation of digital media regulation is essential for safeguarding national security and fostering an integrated, cohesive society.

3. Cases

The study presented discusses three cases broadly classified under the three major issues: regulatory issues regarding content regulation in India; technological issues regarding national security; cross-border regulatory issues regarding content moderation. Based on this classification, the Tandav Controversy of 2021 highlights the regulatory concerns that are needed to be addressed in terms of content regulation in India, the TikTok Ban in India (2020) highlights the technological issues concerning the threat to the national security, and Swami Ramdev v. Facebook Inc.* (2019). These cases triggered the need to rethink the current regulatory framework in India.

3.1. The Tandav Controversy

The Tandav controversy emerged in January 2021, following the release of the political drama series on Amazon Prime Video. The show, directed by Ali Abbas Zafar and starring Saif Ali Khan, Dimple Kapadia, and Mohammed Zeeshan Ayyub, was criticized for allegedly hurting Hindu religious sentiments. The controversy centered around two specific scenes: one involving a college play where Ayyub's character, Shiva, depicted the Hindu god Mahadeva in a manner deemed offensive by some viewers, and another where caste-related dialogue was perceived as derogatory[6]. The backlash led to multiple FIRs being filed across several states, accusing the show's creators of promoting enmity between different groups on religious grounds and insulting religious beliefs. The uproar prompted widespread

[6] India Today. 2021. January 22. The Tandav controversy. Available at: https://www.indiatoday.innewsmo/video/tantav-controversy-what-sparked-it-and-where-it-is-at-1761855-2021-o1-28 (accessed: 16.10.2023)

calls for a boycott of the series, protests, and legal actions. The show's creators issued an unconditional apology, asserting that there was no intent to offend any community or religious sentiments. They also made changes to the controversial scenes following consultations with India's Ministry of Information and Broadcasting. Amazon Prime Video also issued a public apology, expressing regret for any hurt caused and emphasizing its commitment to respecting the diverse cultural and religious sentiments of its audience. Despite these measures, the controversy highlighted the sensitive nature of religious and cultural depictions in Indian media and the significant influence of social media in mobilizing public opinion and political reactions.

3.2. The Swami Ramdev vs. Facebook, Inc.*

The case is a significant example of the complexities and challenges in digital media regulation. In this case of 2019[7], Swami Ramdev sought an injunction against these platforms to remove globally defamatory content that summarized a book banned in India for defamation. The Delhi High Court has ruled that social media platforms must take down the defamatory content globally, not just within India, if it was uploaded from Indian IP addresses. The ruling was based on the interpretation of the Information Technology Act, 2000, specifically Section 79(3)(b), and the Information Technological Rules, 2011, that mandates intermediaries to remove content once they have actual knowledge of its illegality through a court order. The Court has rejected the platforms' arguments for geo-blocking, citing the ease with which such blocks could be bypassed and the need for comprehensive removal to uphold the law's intent [Mendiratta R., Barata J., 2019].

The decision has sparked considerable debate about the jurisdictional reach of Indian courts and the balance between national regulatory needs and global free speech norms. Critics argue that such global injunctions could lead to conflicts of laws and excessive censorship, as different countries have varied standards for what constitutes defamatory or illegal content. This case underscores the tension between enforcing local laws on global digital platforms and maintaining the open and free nature of the Internet.

[7] India Today. 2019. Available at: https://www.indiatoday.in/india/story/delhi-hc-facebook-google-twitter-ramdev-1612313-2019-10-23 (accessed: 25.01.2014)

3.3. TikTok Ban for Privacy and Security Concerns

India’s ban on TikTok in June 2020 was primarily driven by concerns over national security and data privacy, following heightened tensions between India and China. The Indian government cited Section 69A of the Information Technology Act 2000; it allows the government to block access to content that threatens the sovereignty and integrity of India, defense of India, and public order. Along with TikTok, 58 other Chinese apps were also banned during this period. The decision was influenced by the fear that data collected by TikTok could be accessed by the Chinese government, given that Byte Dance, TikTok’s parent company, is based in China. TikTok collects extensive user data, including geolocation, browsing histories, and behavioral patterns, raising concerns about potential espionage and data misuse.[8]

4. A Comparative Analysis

The comparative analysis containing in Fig. 3 below presents a detailed overview of digital media and entertainment regulation across five regions: India, the United States, the European Union, China, and Australia. In India, the regulatory framework is marked as moderate to high, driven by the Digital India Act 2023, prioritizing citizen interests, with privacy protected under the Digital Personal Data Protection Act 2023 and content regulated by the Central Board of Film Certification. The United States features a low regulatory framework focused on freedom, with privacy governed by sector-specific laws like the California Consumer Privacy Act (CCPA) and general laws against illegal content, relying on platform policies for moderation.

The European Union approves a balanced approach with moderate to high regulatory frameworks such as the Digital Services Act (DSA), Digital Markets Act (DMA), and the Media Freedom Act. Privacy and data protection are robust under the General Data Protection Regulation (GDPR), and content is regulated by the Audiovisual Media Services Directive (AVMSD). China enforces a high regulatory framework dominated by state interests via the Cyberspace Administration of China, stringent privacy laws under the Personal Information Protection Law (PIPL) and Data Security Law (DSL), and strict censorship and content control measures.

[8] Euronews. 2024. Which countries have banned TikTok? Cybersecurity, data privacy, espionage fears. *Euronews*, March 24. Available at: https://www.euronews.com/next/2024/03/14/which-countries-have-banned-tiktok-cybersecurity-data-privacy-espionage-fears (accessed: 10.04.2024)

Countries	Regulatory Framework	Privacy ans Data Protection	Content Regulation
India	Citizen's interest prevails Digital India Act 2023	The Digital Personal Data Protection Act, 2023	Digital India Act 2023 Central Board of Film Certification
United States	Regulation centred on principle of freedom	Sector-specific privacy laws and state-level regulations like the California Consuner Privacy Act (CCPA)	General laws against illegal content, with reliance on platforms' own policies for content moderation
European Union	The Digital Services Act (DSA) and Digital Markets Act (DMA) Media Freedom Act	Central Data Protection Regulation (CDPR)	Audiovisual Media Services Directive (AVMSD)
China	State Interest Prevails Cyberspace Administration of China	The Personal Information Protection Law (PIPL) & the Data Security Law (DSL)	Restrictive, with stringent censorship and content measures
Australia	Citizen's interest prevails Online Safety Act 2021	The Privacy Act 1988	Australian Communications and Media Authority (ACMA) The News Media and Digital Platforms Mandatory Bargaining Code

▩ Low □ Moderate ▢ Moderate to High ■ High

Fig. 3. A Comparative Analysis of Regulatory Framework[9]

Australia, similar to India, maintains a moderate to high regulatory stance with the Online Safety Act passed in 2021 and prioritizing citizen interests. Privacy is regulated by the Privacy Act of 1988 and content is overseen by the Australian Communications and Media Authority (ACMA) along with the News Media and Digital Platforms Mandatory Bargaining Code. This comparative analysis depicts the diverse regulatory regimes, varying from minimal regulation in the United States to stringent controls

[9] Compiled by Author from MeitY for The Digital Personal Data Protection Act (2023), Digital India Act (2023) and Central Board of Film Certification; Federal Communications Commission for USA regulatory principles, California legislative information for CCPA, Electronic Frontier Foundation for General laws against illegal content; European Commission for DSA, DMA, European Protection Data Board for GDPR; NPC Observer for PIPL and DSL, China Media Project for Restrictive Content Control Measures; eSafety Commissioner for Online Safety Act 2021, Office of the Australian Information Commissioner for The Privacy Act 1998, Australian Competition and Consumer Commission (ACCC) for News Media and Digital Platforms Mandatory Bargaining Code.

in China, highlighting the different priorities and approaches to managing digital media and entertainment globally.

Conclusion

There is no doubt that the digital era demands robust and flexible regulatory frameworks to address the multifaceted challenges of media entertainment. These frameworks must protect user rights, promote ethical content practices, and support a dynamic and inclusive digital ecosystem. Continued collaboration among policymakers, digital platforms, and civil society is essential to achieve these goals and ensure that the media serves the public interest. The study highlights various challenges including the monopolization of media ownership by powerful conglomerates like Disney, Comcast, and Warner Bros. Discovery poses significant challenges to market competition, content diversity, and cultural representation. This concentration of ownership leads to the homogenization of content, reducing the plurality of voices in media and impacting public discourse and democratic processes. The other significant challenges include privacy, data protection, consumer protection, and also maintaining the balance between the freedom of speech and expression and the prevention of harmful content. Effective regulation in the digital age requires a nuanced approach that balances free speech with harm prevention, ensures transparency and accountability in content moderation, and accommodates diverse cultural norms and legal frameworks. Regulatory bodies must enforce antitrust laws and promote policies that support independent media outlets and public service broadcasting to maintain a pluralistic media environment. In the later section, a comparative analysis of media entertainment regulation across different regions reveals the diverse approaches and challenges faced by policymakers in the digital era. The European Union's robust regulatory framework, exemplified by the General Data Protection Regulation and the Digital Services Act, underscores a strong commitment to user privacy and platform accountability, setting high standards for global data practices.). In contrast, the United States prioritizes freedom of expression, relying on sector-specific regulations like the California Consumer Privacy Act to address privacy issues while maintaining a liberal market environment. Similarly, China's stringent regulatory environment, characterized by the Personal Information Protection Law and extensive content censorship, prioritizes state control and political stability, often at the expense of individual freedoms. Meanwhile, India and Australia strike a balance between regulation and innovation, with frameworks like the Digital India Act 2023 and the Online Safety Act 2021 aimed at protecting user interests and ensuring content accountability.

References

1. Anand N., Brass I. (2021) Responsible innovation for digital identity systems. Data & Policy, 3, e35. doi:10.1017/dap.2021.35

2. Bagdikian B.H. (2004) *The New Media Monopoly.* Boston: Beacon Press, 299 p.

3. Baker C.E. (2009) *Media, Markets, and Democracy.* Cambridge: University Press, 216 p.

4. Belli L., Zingales N. (2019) Platform values: Conflicting rights, artificial intelligence and tax avoidance by digital platforms. Internet Governance Forum. Available at: https://www.intgovforum.org/system/files/filedepot/57/special_issue_platform_values_igf_consolidated_.pdf (accessed: 20.04.2022)

5. Bruns A. (2007) Produsage: Towards a broader framework for user-led content creation. In: Creativity & Cognition: Proceedings of the 6th ACM SIGCHI Conference on Creativity & Cognition, pp. 99–106. Available at: https://doi.org/10.1145/1254960.1254975 (accessed: 16.11.2023)

6. Curran J. (2011) *The Media and Democracy.* New York–Oxford: Routledge, 255 p.

7. Doyle G. (2002) *Media Ownership: The Economics and Politics of Convergence and Concentration in the UK and European Media.* London: SAGE Publications, 192 p.

8. Flanagan A.J., Metzger M.J. (2008) *Digital Media, Youth, and Credibility.* Cambridge (Mass.): MIT Press, 202 p.

9. Flew T., Martin F.R. (eds.) (2022) *Digital Platform Regulation.* London: Palgrave, 263 p.

10. Garton Ash T. (2016) *Free Speech: Ten Principles for a Connected World.* New Haven (Conn.)–London: Yale University Press, 491 p.

11. Gillespie T. (2018) *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media.* New Haven (Conn.)–London: Yale University Press, 297 p.

12. Lotz A.D. (2021) *Media Disrupted: Surviving Pirates, Cannibals, and Streaming Wars.* Cambridge (Mass.): MIT Press, 200 p.

13. McChesney R.W. (2008) *Free Press: The Political Economy of Media.* New York: Monthly Review Press Publishers, 589 p.

14. McChesney R.W. (2015) *Rich Media, Poor Democracy: Communication Politics in Dubious Times.* New York: New Press, 496 p.

15. Mendiratta R., Barata J. (2019) Ramdev v. Facebook, Inc.* Available at: https://wilmap.stanford.edu/entries/ramdev-v-facebook-inc-2019 (accessed: 16.10.2022)

16. Metzger M.J., Flanagan A.J. (2008) Digital Media and Youth: Unparalleled Opportunity and Unprecedented Responsibility. *Journal of Communications*, vol. 58, no. 3, pp. 333–341.

17. Napoli P.M. (2001) *Foundations of Communications Policy: Principles and Process in the Regulation of Electronic Media.* New York–Springfield: Hampton Press, 344 p.

18. Narayanan R. (2024) Legal compliance for OTT platforms in India: Understanding the regulatory landscape. Available at: https://www.khuranaandkhurana.com/2024/04/03/legal-compliance-for-ott-platforms-in-india-understanding-the-regulatory-landscape/ (accessed: 19.06.2024)

19. Noam E.M. (2016) Who Owns the World's Media? Media Concentration and Ownership around the World. Oxford: University Press, 440 p.

20. Noble S.U. (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism.* New York: New York University Press, 248 p.

21. Pickard V. (2019) *Democracy without Journalism? Confronting the Misinformation Society*. Oxford: University Press, 264 p.

22. Priya V.B.A. (2023) Overview of over-the-top (OTT) platforms in India: analysis of IT Rules 2021, judicial battles, balancing act of artistic freedom of speech and expression, and challenges for regulation in India. Available at: https://legalresearchandanalysis.com/overview-of-over-the-top-ott-platforms-in-india-analysis-of-it-rules-2021-judicial-battles-the-balancing-act-of-artistic-freedom-of-speech-and-expression-and-challenges-for-regulation-in-india/ (accessed: 21.04.2024)

23. Suzor N. (2019) *Lawless: The Secret Rules that Govern our Digital Lives.* Cambridge: University Press, 218 p.

24. Tworek H., Leersseen P. (2019) An Analysis of Germany's NetzDG Law. Available at: annenbergpublicpolicycenter.org. (accessed: 24.05.2023)

25. Vogel H.L. (2020) Entertainment Industry Economics: A Guide for Financial Analysis. Cambridge: University Press, 680 p.

26. Voigt P., von dem Bussche A. (2017) The EU General Data Protection Regulation: a practical guide. Berlin: Springer International Publishing, 122 p.

27. Zuboff S. (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* New York: Public Affairs Publishing, 704 p.

Information about the authors:

Reeta Sony A.L. — PhD, Associate Professor.
Shruti Chopra — PhD, Research Associate.

The article was submitted to editorial office 03.06.2024; approved after reviewing 28.06.2024; accepted for publication 28.06.2024.

Legal Issues in the Digital Age. 2024. Vol. 5. No. 2.
Вопросы права в цифровую эпоху. 2024. Том 5. № 2.

E- Government

Research article
УДК: 342
DOI:10.17323/2713-2749.2024.2.113.126

Technologies Versus Justice: Challenges of AI Regulation in the Judicial System

Elena Vladimirovna Burdina[1], Viktor Nikolaevich Kornev[2]

[1, 2] Russian State University of Justice, 69 Novocheremushkinskaya Str., Moscow 117418, Russia,

[1] elenburdina@yandex.ru

[2] kornev51@yandex.ru

Abstract

The article examines issues of using artificial intelligence in such a sensitive area of human activity as justice. The authors refer to numerous facts on attempts to create a kind of "smart court" in various countries. At the same time, these attempts run up against circumstances that indicate the need to establish legal restrictions on the use of artificial intelligence in the administration of justice. Moreover, according to the authors' reasoned conviction, there are areas in which the robot judge turns out to be powerless to replace human intelligence. Based on the philosophical and legal approach to assessing such a phenomenon as digitalization and the phenomenology of legal judgment, the authors conclude that the adoption of a court decision that meets the requirements of the principle of justice is something beyond the reach of artificial intelligence. Such a decision can only be made by a human judge, but not by a robot. AI systems in the judicial system should support rather than supersede judges.

Keywords

artificial intelligence; robot; judge, judicial decision; "smart court"; phenomenology; justice.

For citation: Burdina E.V., Kornev V.N. (2024) Technologies Versus Justice: Challenges of AI Regulation in the Judicial System. *Legal Issues in the Digital Age,* vol. 5, no. 2, pp. 113–126. DOI:10.17323/2713-2749.2024.2.113.126

Background

Dramatic social changes caused by the fourth industrial revolution and its principal offspring — artificial intelligence (AI) — are challenging the judicial system, with society and judges faced with problems never seen before. By their sheer impact on the value basis of the judicial system, these challenges need to be promptly addressed by theory and systemic regulation.

In many areas of industrial production and public services the digital technologies including AI are regarded as a factor of development and a modern method (benefit) for reducing production costs, improving labor productivity and management performance, providing for new usability, and ensuring better living standards and individual comfort. Future-focused expressions like “smart home”, “smart plant” or “smart city” reflect the current trend to make AI systems part of the economic and social texture and to create economically viable models [Filipova I.A., 2021: 92–105].

In the wake of this rhetoric, the doctrinal literature and case law increasingly employ the word combination “smart court” that assumes the use of automation, digital data communication/processing systems and AI across the board including legal procedures, case management and administration.

While countries are now only at the early stage of AI introduction, this technology increasingly permeates the judicial system with no resistance on the part of judges, only too eager to test new capabilities for addressing professional tasks.

Meanwhile, AI is fraught with evident threats (named *digital risks*), something that pushes researchers and practitioners to look for answers to the question of AI feasibility in the judicial field in general and legal decision-making in particular, as well as of the forms and methods to regulate its usage.

While the opportunities of using AI are welcomed rather than questioned by judges themselves, there is no shared view on the meaningful use of this technology to render justice. Also, there is a bitter controversy around the extent and legitimacy of AI use in legal decision-making.

1. "Smart Court": AI Judicial Uses in Russia and Elsewhere

The question of possible use of machine algorithms in court is not new either for international or domestic science.

The noted American mathematician Norbert Wiener, one of the founders of cybernetics, first posed the question of using cybernetics to deal with legal issues back in 1958 [Wiener N., 1958: 117].

A similar question was discussed by professor S. Levi (France) in his presentation "Cybernetics and Law" at the Second International Congress on Cybernetics (Belgium, 1958). The speaker, in particular, argued that cybernetics should be used both to create and use laws since lawyers "have to deal with increasingly difficult situations resulting from complex organization and fast pace of living of the modern society" [V.A. Ilyin et al., 1961: 368].

The same question was formulated more specifically by L.E. Allen in his report for International Conference on Machine Languages in Cleveland (United States, 1959) on machine discovery and verification of grammarlogical ambiguities in pleadings.

It was stated already at that time: even the most advanced machine would never become a substitute for human creativity, with cybernetics exploring only quantitative aspects of management processes. Cybernetic devices are just auxiliary technologies for addressing the legal problems of enforcement and management.

In his article "Cybernetics and Law", D.A. Kerimov, legal section chairman of the Research Council on Cybernetics in the USSR Academy of Sciences, noted in 1962, that "any ideas to fully replace human creative intelligence with machines are to be strongly condemned". He was outraged that "there are lawyers who are serious about feasibility and rationality of developing a cybernetic device to replace the judge!" [Kerimov D.A., 1962: 102, 103].

While technologies have advanced considerably by now, the main question put in the simplest but essentially valid form — can artificial intelligence replace the judge? — is yet to be addressed.

The answers to these questions are produced by way of experimenting and building up innovative experience of using digital technologies (including AI) in court.

Despite the intrinsic conservatism of procedural form, the judicial power cannot remain outside digital communications emerging at executive agencies and businesses (as prompted, in particular, by interagency digital communication with many adjacent bodies).

AI tools demonstrate an enormous potential for courts — such as data processing, audiovisual identification, search and analysis of legal documents. AI provides social advantages for exercise of the right to judicial protection as it facilitates access to justice by offering a claim drafting wizard as well as advice on simple and frequently asked questions.

The opportunities for using AI in legal proceedings are extensively explored under different legal systems. Internationally, these technologies are tested to address various tasks including to examine and resolve disputes and to deliver final judgments.

Thus, in China "...major issues brought about by the era of digital technologies and cybernetics, era of artificial intelligence and dissemination of blockchain" are dealt with at the government level.

The Supreme People's Court of China Resolution "On Regulation and Application of Artificial Intelligence in the Judicial Field" (2022) purports to introduce an improved AI system at courts for comprehensive support of justice and lower burden on judges. It is envisaged by 2030 to put in place an applied and theoretical system for AI use in the judicial field and to develop relevant standard rules consistent with generally accepted standards and principles of justice. The Supreme People's Court resolution identifies AI in-depth integration with litigation and enforcement, court services and administration, as well as modernization of the judicial system and services across the board as strategic areas of development. [1]

The introduction of AI into China's judicial system has already provided sizeable economic and financial gains by allowing to reduce the workload of judges by more than one third and save 1.7 billion hours of working time and over 300 billion yuan (45 billion US dollars) in the period from 2019 to 2021.

Chinese digital services cover all stages of case examination and resolution from pre-trial settlement to enforcement of judgments, including case file management and archiving processes. The judicial system makes active use of Big Data technologies, intelligent data processing for speech recognition, case analysis, file error correction, similar case search, case document drafting assistance.

[1] Available at: https://ru.chinajusticeobserver.com/law/x/the-supreme-people-s-court-the-opinions-on-regulating-and-strengthening-the-applications-of-artificial-intelligence-in-the-judicial-field-20221208 (accessed: 26.01.2024)

Each judge's desk is connected to the Smart Court SoS digital system. As reported by the Supreme People's Court, this system daily analyzes and draws conclusions on approximately 100,000 cases nationwide to monitor the progress of each case and prevent abusive or corrupt practices[2].

In India, virtual courts examine same-type claims for violation of traffic rules based on AI-aided algorithmic proceedings[3].

AI is actively used to provide access to justice.

In Germany, AI systems support the processing of mass claims (in particular, those to road vehicle manufacturers with regard to sales). Essentially of the same type, such claims differ in minor details: motor type, price, mileage, etc. AI is used to process data and draft the final certificate.

In Portugal, the Justice Ministry is in process of developing a virtual assistant based on GPT system to facilitate people's access to information.

Trial courts in Singapore are testing generative AI to process claims for divorce and some other civil cases[4].

In Russia, large-scale introduction of AI is hinged on Online Justice super-service to be made operational not later than 1 January 2025, with services to include weak AI technologies to be used in proceedings including for automatic drafting of judgments based on analysis of claims and case files, decoding audio minutes, searching/analyzing legal precedents, and performing administrative routine (record keeping and archiving).

V. Momotov argues that weak AI can be used to examine civil and administrative cases for collection without recourse, primarily in summary proceedings, as "decision-making is largely technical and not related to analysis of legal relations between the parties"[5].

[2] China's court AI reaches every corner of justice system, advising judges and streamlining punishment // South China Morning Post. 13.07.2022. Available at: https://www.scmp.com/news/china/science/article/3185140/chinas-court-ai-reaches-every-corner-justice-system-advising (accessed: 26.01.2024)

[3] The Courts and COVID-19: Adopting Solutions for Judicial Efficiency. 04.06.2020. Available at: https://ecommitteesci.gov.in/the-courts-and-covid-19-adopting-solutions-for-judicial-efficiency/ (accessed: 22.01.2024)

[4] Singapore courts to test generative AI. Available at: URL: https://tass.ru/ekonomika/18851511?ysclid=lrv9b7s79v584002374 (accessed: 26.01.2024).

[5] Presentation "Smart courts and the future of judicial power" by V. Momotov, Chairman of the Judicial Council at the XVIII Conference of Supreme Court Chairpersons of SCO Member States in Delhi, 11 March 2023. Available at: URL: http://ssrf.ru/news/lienta-novostiei/50081?ysclid=lrv9t2lweb234395325 (accessed: 29.01.2024)

The funded knowledge of AI use in the judicial system shows that its introduction solves three main objectives: reducing the workload of judges and court staff across both procedural and administrative (essentially auxiliary) segments; ensuring faster proceedings; satisfying people's needs in cheaper, more accessible and convenient forms of access to justice.

Thus, the early experience of introducing AI into countries' justice systems shows economic, social and administrative gains, with conveniences and advantages offered by this technology to encourage further expansion not only into judiciary communities but also government and society.

2. AI in the Justice System: Challenges of Institutionalization

The current period is prioritizing the search for reliable regulatory system to fence off adverse implications of AI use, and for sources likely to be acceptable for multi-tier social regulation.

Apparently, AI institutionalization challenges related to new "digital risks" for the judicial system, need to be addressed via the law.

Adoption of regulations should be accepted as an ideal regulatory method since the law itself is the supreme regulator [Maltsev G.V., 2016: 770].

However, instant regulation of the problem like in the age of stability and all-over codification, as mentally (and habitually) expected by the legal profession is not feasible and even practically impossible since it is hard to formalize *as due* the procedure for AI systems tested at courts over short periods, often as test samples, or yet to be developed.

As such, the problem of necessary regulators can be addressed at the first stage of AI introduction via not only legal but also non-legal social regulators, primarily ethical corporate standards would later provide a robust social basis for legal regulators.

Overall, the AI regulatory system appears more sustainable and effective given the diversity of social regulators combining legal and ethical regulation as reinforcement.

The study of doctrinal literature provides similar views with regard to the search for an adequate regulatory system.

V. Sinyukov argues with good reason that such a system, given the intervention of technical regulators, cannot rely on highly abstract provisions emerging through a long evolutionary process but instead should be "highly empirical and concrete" [Sinyukov V.N., 2021: 26].

Social regulation should be bidirectional, with public interests essentially opposing each other to contain and encourage AI development [Djeffal C., 2019: 255–284]. On the one hand, it should ensure protection of the society and individuals from negative implications of digital technologies and to make them safe while, one the other hand, to encourage innovative AI development for judicial purposes.

The Recommendation on the Ethics of Artificial Intelligence (hereinafter Recommendation)[6] passed at the UNESCO Conference of 23 November 2021 attended by 193 member states is believed to the first global source of AI regulation in the international practice. Its starting point is that control arrangements should be based on values and principles not to be violated through the use of technologies.

The Recommendation provides the aims, values and principles of AI use, as well as guidance for all areas where AI is introduced.

The main values underlying all policy measures and regulations relevant for AI are respect and protection of human rights and fundamental freedoms and human dignity. No human being or community should be harmed or subordinated, whether physically, economically, socially, politically, culturally or mentally during any phase of AI system lifecycle.

Throughout the lifecycle of AI systems the quality of life of human beings should be enhanced (clause 14 of the Recommendation). Values to be supported by AI include: environmental and ecosystem flourishing (clauses 17–18), promotion of diversity and inclusiveness (clauses 19–21), and living in peaceful, just and interconnected societies (clauses 22–24).

The principles of AI ethical regulation include: proportionality and do no harm, safety and security, fairness and non-discrimination, sustainability, right to privacy and data protection, human determination, transparency and explainability, responsibility and accountability, awareness and literacy, multi-stakeholder and adaptive governance and collaboration.

China demonstrates a sustainable procedural strategy with regard to AI regulation in the judicial system, with the Supreme People's Court Resolution "On Regulating and Promoting the Use of AI in the Field of Justice"

[6] Recommendation on the Ethics of Artificial Intelligence. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000381137 (accessed: 29.01.2024)

passed in 2022 containing five principles that identify the parameters of AI technologies used in courts.

The general principles include those of security and legitimacy, integrity, fairness, auxiliary role in decision-making, transparency and credibility, compliance with public order and good customs.

The principle of security and legitimacy essentially prohibits to use AI technology and products to the detriment of national security and legitimate interests of individuals and organizations.

The principle of fairness and integrity requires to follow the fundamental principles of justice and ensure fair trial and equal opportunities to stakeholders.

AI's auxiliary role in proceedings is a critically important rule (principle) since it prohibits AI to deliver judgments instead of the judge.

Under the principle of transparency and credibility, all AI algorithms are subject to control, assessment and registration by the relevant authorities. Such algorithms should be verifiable to make the procedure and outcomes of AI use predictable and credible.

The meaning of the fifth principle is that the use of judiciary AI systems should not undermine public order and good customs.

We believe that the sustainable procedural approach to AI regulation in the justice system contains the outlines of the applicable legal regime and provides the framework to institutionalize this phenomenon in the judicial field. The next step is to formulate special standards that will establish the legal regime for AI across different types of proceedings to examine different categories of cases from the perspective of common procedural principles and judicial practice.

3. Using AI for Decision-Making: Red Lines

The legal literature provides the views on obvious advantages of AI compared to human intelligence, with some authors considering the matter of replacing judges with robots — at least, in e-courts — as closed [Fursov D.A., 2021: 46–53].

In our view, the problem of using AI to deliver judgments or intermediate orders depends on the general legal theory and legal philosophy at the same time. The critically important issues are, firstly, those of legitimate sources of judicial power and, secondly, those of the nature of judgments to be made in rendering justice.

The academic discussion of the problem prompts the following question: "Does the delegation of decision-making authority from a legitimately appointed judge to artificial intelligence (machine) match the nature of judiciary power?"

There is apparently no profound study of how AI systems have impacted judiciary institutions shaped by millennia of human history. The current cursory effects to reduce the workload and accelerate proceedings cannot serve as a criterion for their unlimited use.

The lagged effects of transition from "man-man" to "man-machine-man" or "man-machine" patterns in the communicative model of justice threaten not only to undermine the outcomes of justice but equally to cripple the actor — the judge as the embodiment of judicial power — with a profound debasement.

The doctrine shows an increasing number of authors who adopt the view that using AI to deliver judgments is contrary to the idea of the rule of law [Djeffal C., 2022: 33–44] and fair trial.

The fundamental importance of AI acceptability for judicial decision-making calls for a number of interrelated ideas to set the limits of what is acceptable from the perspective of legal philosophy and theory and other fields of knowledge.

The ongoing processes are hinged on the solution to the dilemma of what comes first: artificial intelligence based on mathematically computable algorithms or human mind capable of perceiving and understanding the facts of life including for rendering justice. That is, the principal question is whether AI (robot) can replace human judge in legal proceedings.

The advocates of using AI in the judicial system believe they can thus significantly simplify, accelerate and facilitate case examination at court. Therefore, they focus on technological aspects of undisputable benefit in the digital age, only to bypass the main question of correlation between computation-based AI and conscious thinking proper of physical activity of human mind endowed with intelligence.

There are at least four viewpoints in this regard: every thinking is computation; in particular, the sense of knowledgeable cognition is in fact the outcome of corresponding computation; cognition is a characteristic manifestation of physical activity of human mind [Penrose R., 2005: 35]; though any physical activity can be simulated through a set of computations, numerical simulation cannot be the effective cause of cognition; cognition results from the relevant physical activity of human mind but this physical activity cannot be adequately simulated by computational means; cognition cannot be explained in any physical, mathematical or scientific terms whatsoever [Johnson–Laird P.N., 1983: 252].

The view of the philosopher John Searl in support of the third point is especially interesting in light of this discussion [Searl J.R., 1992].

Positively, justice is not about technologies. It is a process we implement to make a judgment based on our interpretation of legal principles and provisions as well as actual circumstances of the case, and no constructed syllogism or subsumption — mechanically matching the actual circumstances with a legal provision or rule of behavior — will help AI to sort it out.

A legal decision is an act of judgment containing new knowledge that may be true or false. Deciding whether an assertion is true or false requires cognition characteristic only of human mind and unavailable to its electronic simulation.

AI is thus simulated intelligence that, unlike genuine intelligence of the judge, does not require to understand or perceive the legal principles and provisions to be applied.

Interpretation as intellectual and volitional activity has always been and will be the legal profession's main purpose of activity because of interpretative nature of law that adds up to its other properties. Since regulations and other forms of law create rights and obligations involving often different forms of liability, only interpretation can serve the purpose of their "right" understanding.

Understanding is part and parcel of genuine intellect: in interpreting a regulation, the judge perceives its meaning and legislator's intention, that is, the will and purpose pursued by the legislator in adopting a certain regulation.

Human consciousness is characterized by such intellectual phenomena as thinking, volition and judgment that are proper of the judicial decision-making. Since, these phenomena are not shared by AI, it is not truly conscious. Consciousness, cognition and understanding are the abilities that no computing system fully has or can ever learn, just like it is incapable of aesthetic perception and judgment of what is ethical, beautiful or good as these things require cognition.

While AI can simulate these abilities, it will require additional controlling impact on the part of external, sensitive and conscious being — man [Penrose R., 609]. We thus agree with researchers who assert that only man can render justice [Kleandrov M.I., 2018: 15–25].

Law is a highly complex phenomenon that manifests itself at different levels of human existence and each time in a different quality [Kaufman A., 2019: 18–29].

It is explored by philosophy of law, theory of law and sociology of law, each at its own viewing angle. This fact is indicative of the integrative nature of law that allows to regulate relevant part of human behavior and express the interests that make up the foundation of legal provisions and principles [Yershov V.V., 2019: 17]. Without it justice and rule of law relying on a set of abstract principles and rules will not only run up against human existence but pose some sort of a threat to it as an instrument of formalized digital government.

Any public (including state) institution is underpinned by the idea of justice. A robot is incapable of just, that is, fair decision-making. It is human legacy because only man can understand what is fair and what is not [Zorkin V.D., 2017: 2]. Where a public institution is efficient and formal but not fair, its legitimacy cannot be sustained. Such institution should be either reformed or abolished.

According to D. Rawls, truth and fairness accept no compromise [Rawls D., 1995: 19, 20], not even when a departure from fairness is compensated by economic or social benefits. Fairness is a major institutional attribute.

Fairness does not only essentially define justice as a special government activity to examine and resolve cases but also a mechanism supposed to result in a just, that is, fair outcome. In a sense, the judicial system and proceedings are the two aspects (functional and administrative) of legal and formal embodiment of the concept of fairness created by generations of humankind through successive reforms and transformations. Fairness of the judicial system and the concept of the rule of law implemented by court are based on the axiom of judicial power exercised by man — the judge appointed or elected by formal procedure. Any departure from this axiom is a violation of legal succession essentially amounting to revolution in the legal sense.

All revolutions have a cost in terms of public good and seem to be a weapon with devastating social, cultural and politico-legal consequences. A revolutionary method of delegating the decision-making authority to AI falls short of the task to protect rights, only to result in a damage to fundamental values of judicial power well beyond any economic benefit.

A full and uncontrolled delegation of the decision-making authority from a human judge to AI is incompatible with the nature of judicial power.

With technologies opposing the fundamental values of justice as fair trial, there is a need to regulate the extent and forms of control over the use of artificial intelligence as well as formulate relevant prohibitions.

A control mechanism for acceptable use of AI to deliver judgments should have at its core, in our view, the axiom of irreplaceable human judge. Based on this concept, it should include the following components: a) identifying process stages, case categories, types of judgments involving AI; b) prohibiting automated judgments, that is, those made without human control; c) right of the judge to decide whether to involve AI for assistance; d) principle that AI-generated decision is auxiliary; e) principle of personal responsibility of the judge for decisions being made; and f) identifying the risks likely to result from large-scale judicial use of AI across the board.

The incontestable and unconditional premise that the authority to render justice can never be delegated to artificial intelligence should be specifically enshrined in law. AI should support rather than supersede judges[7].

Adoption of this standard (principle) will allow to end up theoretical and practical (often fruitless) discussions of the problem and to focus on practical implementation of technologies in the judicial system without risking to undermine its fundamental values.

Conclusion

No “smart court” ideology involving large-scale use of digital technologies and AI can be promoted as a path for the judicial system to follow, unless fundamental principles of legal theory and judiciary power are strengthened and developed.

[7] CCJE Opinion No. 26. 2023. Moving forward: the use of assistive technology in the judiciary. Available at: https://rm.coe.int/ccje-opinion-no-26-2023-final/1680adade7 (accessed: 20.01.2024)

The nature of judicial power is incompatible with full and uncontrolled delegation of the decision-making authority to artificial intelligence. Using AI to make judgments is contrary to the rule of law and fair trial: AI systems can function only as an auxiliary technology.

AI in the judicial system is at the stage of inception characterized, along with innovations, by moderate conservatism, building up of empiric experience, development of models acceptable for justice and administrative and procedural rules to involve AI without undermining the humanitarian, human nature of fair and legitimate judgment.

As part of the experimental stage, it appears useful to provide for a "pilot court" regime as a way to reduce the risks from implementation of digital innovations at court and to assess the prospects of the relevant organizational forms. This will create a space to develop, validate and introduce digital technologies into procedural, record-keeping and administrative activities of courts.

With technologies opposing the fundamental values of justice as fair trial, it is necessary to institutionalize AI and create a multi-tier regulatory system that will fence off any adverse implications of AI, establish the extent and forms of control over its use and formulate relevant prohibitions.

The best regulatory strategy for AI in the justice system is the procedural guarantee approach that will define the principles of its use AI for the judicial system should support rather than supersede judges. A control mechanism for acceptable use of AI to deliver judgments should be based on the axiom of irreplaceable human judge and include the following: a) identifying process stages, case categories and types of judgments involving AI; b) prohibiting automated judgments, that is, those made without human control; c) right of the judge to decide whether to involve AI for assistance; d) principle that AI-generated decision is auxiliary; e) principle of personal responsibility of the judge for decisions being made; f) identifying the risks likely to result from large-scale judicial use of AI across the board.

References

1. Cleandrov M.I. (2018) Reflections on the Topic: Could there be a Robot as a Judge? *Rossiyskaya justitcia*=Russian Justice, no. 6, pp. 15–25 (in Russ.)

2. Djeffal C. (2022) Wie sollen wir künstliche Intelligenz regeln? *Public Law Research*, vol. 3, no. 12, pp. 33–44.

3. Djeffal C. (2019) AI, Democracy, and the Law. The Democratization of Artificial Intelligence: Net Politics in the Era of Learning Algorithms. In: *Digitale Gesell-*

schaft, vol. 25, pp. 255–284. Available at**:** https://ssrn.com/abstract=3535735 (accessed: 16.10.2023)

4. Ershov V.V. (2019) Law in the Context of the Metamodern Paradigm. *Pravosudie*=Justice, no. 2, pp. 15–34 (in Russ.)

5. Filipova I.A. (2021) Labor Law in Transition to a Digital Society: ongoing Changes and Contours of the Future. *Zhurnal rossiyskogo prava*=Journal of Russian Law, vol. 25, no. 3, pp. 92–105 (in Russ.)

6. Fursov D.A. (2021) Robotization of Legal Activities. *Rossiyskoye pravosudie*=Russian justice, no. 4, pp. 46–53 (in Russ.)

7. Ilyin V.A. (ed.) (1961) Philosophical Issues of Cybernetics. A collection of essays. Moscow: Izdatelstvo sotcialo-ekonomicheskoi literatury, 392 p. (in Russ.)

8. Johnson–Laird P.N. (1983) How Could Consciousness Arise from Computations of the Brain? In: C. Blakemore, S. Greenfield (ed.) Mindwaves: thoughts on intelligence, identity and consciousness. Oxford: Blackwell, 550 p.

9. Kant I. (2010) *Critique of the Pure Reason*. Moscow: Mysl, 736 p. (in Russ.)

10. Kaufman A. (2019) Philosophy of law, theory of law, legal dogma. *Gosudarstvo i pravo*=State and Law, no. 5, pp.18–29 (in Russ.)

11. Kerimov D.A. (1962) Cybernetics and Law. *Sovetskoye gosudarstvo i pravo*=Soviet State and Law, no. 11, pp. 98–104 (in Russ.)

12. Larenz K. (1979) *Methodenlehre der Rechtswissenschaft*. 4 Aufl. Berlin: Springer–Verlag, 525 S.

13. Maltsev G.V. (2016) *Social Foundations of Law.* Moscow: Norma, 800 p. (in Russ.)

14. Penrose R. (2005) *Shadows of the Mind. In Search of Science of Consciousness.* Moscow: Institute of Computer Studies Press, 688 p. (in Russ.)

15. Rawls D*. (*1995*) The Theory of Justice*. Novosibirsk: University Press, 534 p. (in Russ.)

16. Searl J.R. (1992) *The Rediscovery of the Mind.* Cambridge (Mass.): University Press, 220 p.

17. Sinyukov V.N. (2021) Law and Order Theory of V.V. Borisov in the Modern Russian Context: Significance and Prospects. *Zhurnal rossiyskogo prava*=Journal of Russian Law, vol. 25, no. 3, pp. 17–32 (in Russ.)

18. Wiener N. (2019) *The Cybernetics and the Society.* Moscow: AST Publishers, 200 p. (in Russ.)

19. Zorkin V.D. (2017) The Essence of Law. *Zhurnal constitutcionnogo pravosudia*=Journal of Constitutional Justice, no. 5, pp. 1–11 (in Russ.)

Information about the authors:

E.V. Burdina — Doctor of Sciences (Law), Associate Professor,
V. N. Kornev — Doctor of Sciences (Law), Professor.

The article was submitted to editorial office 07.03.2023, approved after reviewing 28.07.2023, accepted for publication 15.09.2023.

Legal Issues in the Digital Age. 2024. Vol. 5. No. 2.
Вопросы права в цифровую эпоху. 2024. Том 5. № 2.

Research article
УДК: 342
DOI:10.17323/2713-2749.2024.2.127.147

Digitizing Law-Making at Federal Executive Agencies

Anton Aleksandrovich Doschatov

Ministry of Economic Development of Russian Federation, 10 Presnenskaya Embankment, Moscow 123112, Russia,

a.doshchatov@molprav66.ru

Abstract

With digital technologies gaining ground in all spheres of life across the world, the digitization is becoming especially relevant, only to require in-depth technical and legal analysis. Current digital development of public administration in Russia calls for a change of approach to drafting and фадоption of regulations. Today's technologies are expected not only to make sure that regulation is timely and complete, but that it is as efficient as possible in view of an enormous and ever growing number of tasks, as well as non-controversial and comfortable for all those involved in the law-making process. Annually the Justice Ministry and its territorial offices receive about 1 million draft regulations for anti-corruption and legal review and state registration at the relevant level. It is exactly for this reason that this paper purports to conduct a comprehensive analysis of digitization processes affecting law-making at federal executive agencies using the example of the state information system "National shared environment for collaboration between the parties to the law-making process" (SIS Normotvorchestvo) and to identify operational problems. The study aims to explore the current operational status of the system, analyze issues and constraints related to digitization of law-making, identify potential advantages and benefits to be gained via digitization, and to discuss further prospects. The methodological basis of the research includes formal legal method of inquiry as well as logical method allowing to present findings and draw conclusions; methods of analysis and synthesis; comparative legal method.

Keywords

digitization; law-making; state information system; automation; digital change; parties to the law-making process.

For citation: Doschatov A. A. (2024) Digitizing Law-Making at Federal Executive Agencies. *Legal Issues in the Digital Age,* vol. 5, no. 2, pp. 127–147. DOI:10.17323/2713-2749.2024.2.113.126

Background

The international community's transition to the global technocratic concept of development since the late 20th— early 21th centuries has caused an explosive growth of IT industries and set the stage for accessibility of information technologies across various spheres of daily life and at different levels of social networking. In the same period, information technologies became instrumental and provided resources for globalization and integration processes taking place in the world economic system particularly in the political, military, industrial, socioeconomic, spiritual, research and technological spheres.

The year 2023 has witnessed in Russia a transition of all federal executive agencies towards electronic exchange of non-public documents (NPD).[1] While in 2020 the Government Office did not exchange such documents electronically, almost 80 federal executive agencies and public authorities are now involved in such exchange. This transition to interagency e-document exchange has caused the share of e-documents to grow 2.4 times, from 35% to 84%. By the early 2025 it is also expected to integrate the state information system "National shared environment for collaboration between the parties to the law-making process" (SIS Normotvorchestvo; hereinafter the new SIS) into the Government Office e-document exchange system to put in place an e-platform for drafting federal constitutional, federal laws and regulations, as well as facilitate interagency exchange and control of execution of high executives' instructions, and improve the quality of draft documents and add-on applications. All these efforts are expected to improve public administration functions through better quality of work by the information system's users.

1. Law-Making at Federal Executive Agencies: Urgency of Digitization

In terms of its impact on civil society, digitization is comparable to the fifth industrial revolution. Explosive transformation and expansion of so-

[1] On approving the Provision on circulation of non-public documents at federal executive agencies, authorized office for atomic energy and authorized office for space activities. The Government of the Russian Federation Resolution No. 1233 of 11 March 1994 // Collected Laws of Russia, 2005, No. 30. Part II. Article 3165.

cial relations into virtual environment inevitably affect regulation pushing legal provisions to adapt to a new social development model (that is, become flexible).

Digitization of the economic system and of administrative decision-making, in particular, through the introduction of digital technologies into operations of federal executive agencies, means for any modern government a transition to totally different paradigm for stronger national competitiveness, higher living standards of population and society, and more robust GDP and GNP growth [Stenkin D.S., 2023: 64].

The digital age has witnessed a qualitative change to the way the government makes and implements administrative decisions by actively using innovative means to receive, process and transmit information in machine-readable forms. Meanwhile, proper quality of public administrative decisions in the absence of legal instruments adequate to the digitization processes of public administration cannot be maintained without due regulation[2].

Digitization of law-making at federal executive agencies is more limited in extent and largely takes place through operational expansion of already well-established IT technologies. References to provisions of the federal project "Digital Governance" of the Russia's Digital Economy National Program leave no doubt public control (supervision) and provision of public and municipal services are the priority areas for digitization of federal executive agencies[3] [Kabytov P.P., 2020: 115].

As a type of public governance, law-making is aimed at creating, improving, amending or revoking legal provisions.

A consolidated description of the methods to apply digital technologies to law-making was given for the first time by the Institute of Legislation and Comparative Law under the Federal Government in the draft Federal Law "On Regulations" (2012) in view of digitization of social life across the board and the need to introduce modern technologies to the law-making process. The 2019 draft provides for more detailed regulation: apart from using information technologies for access to already published legal instru-

[2] Information Society Development Strategy for 2017–2030. Presidential Decree No. 203 of 09 May 2017 // Collected Laws of Russia, 2017. No. 20. Article 2901.

[3] Passport of the national project "Russia's Digital Economy National Program" approved by the Presidium of the Council for Strategic Development and National Projects under the President of Russia, protocol No. 7 of 04 June 2019.

ments; its Chapter 16 "Using information technologies" contains provisions on automated law-making systems.

With regard to digitization of administrative decision-making, it is necessary to make a distinction between informatization and digitization. The use of computers, gadgets, applications and Internet is informatization of public institutions. In contrast, digitization is hinged on accessible and mobile Internet services and AI (Strategy for development of artificial intelligence in Russia approved by Presidential Decree No. 490 of 10 October 2019 in force since 2019 as a strategic planning document),[4] self-learning machines, principally new software and computer technologies (distributed ledger technology, etc.) [Lipen S.V., 2019: 25].

Abundance of documents in force of various status is fraught with legal risks for digitization of public administration due to a need for continuity between earlier and subsequently adopted regulations.

In March 2024, following Presidential Instructions No. 2242 of 31 December 2020, No. 1383 of 05 August 2021 and No. 1553 of 01 September 2022, the Federal Government approved the strategic focus of digital change in public administration as largely aiming to ensure sustainable and safe information exchange between public authorities, society and businesses.

Digital change in public administration has the following priorities: automation and simplification of government operations in terms of interagency cooperation and organization of standard processes; shared information environment for intra- and interagency e-collaboration between federal executive agencies and public authorities in constituent territories.

The main problem is a lack of shared domestic tool for exchange of documents and information under the applicable law and a lack of possibility to exchange legally meaningful e-documents within the framework of public administration.

In the course of implementing strategically focused projects it is expected to introduce AI technologies for automation of standard processes to save time spent on addressing routine tasks and searching for sound decisions.[5]

[4] On the Development of Artificial Intelligence in Russia (annex to the 2030 National Strategy for Development of Artificial Intelligence). Presidential Decree No. 490 of 10 October 2019 // Collected Laws of Russia, 2019, No. 41. Article 5700.

[5] Approving the Strategic Focus of Digital Change in Public Administration. The Government of the Russian Federation Instruction No. 637-r of 16 March 2024 // Official web portal of legal information, 20.03.2024.

2. Regulating Use of Public Information Systems

Under Article 14 of Federal Law No. 149-FZ "On Information, Information Technologies and Data Protection" of 27 July 2066, state information systems are established for the purpose specified in this federal law. The Federal Government approves the requirements to the procedure for creating, developing, commissioning and de-commissioning state information systems. The information contained in such systems ranks as public resources, with federal executive authorities to make sure that information posted to these systems is reliable and up to date.[6]

While the national law does not define a "state information system", the author believes it to be a set of interrelated software and hardware designed to collect, store, process and provide information at public agencies and development institutions. State information systems ensure centralized data management for better operational quality of federal executive agencies and faster administrative decision-making.

Within a state information system there may be a large number of information systems responsible for different aspects of public administration such as accounting and control (standard cloud solution for automatic control operations:[7] "Governance" automated system[8]); analysis and forecasting (shared interagency information and statistical system);[9] e-document exchange (non-public e-document exchange between federal executive agencies) and public finance administration (integrated SIS "E-Budget").[10]

Developing approaches to digital change in public administration will require to elucidate the terms "digital space" and "digital twin". The Dictionary of Terms and Concepts of Digital Change defines digital space as

[6] Rossiyskaya Gazeta, 03 August 2006.

[7] On State Information System "Standard Cloud Solution for automatic control (supervision) operations". The Government of the Russian Federation Resolution No. 482 of 21 April 2018 // Collected Laws of Russia, 2018, No. 18. Article 2633.

[8] On the Automated State Information System "Governance". The Government of the Russian Federation Resolution No. 1088 of 25 December 2009 // Collected Laws of Russia, 2010, No. 1. Article 101.

[9] On the Shared Information and Statistical System. The Government of the Russian Federation Resolution No. 367 of 26 May 2010 // Collected Laws of Russia, 2010, No. 22. Article 2779.

[10] On the Integrated State Information System for Public Finance Administration "E-Budget". The Government of the Russian Federation Resolution No. 658 of 30 June 2015 // Collected Laws of Russia, 2015, No. 28. Article 4228.

“space for integration of digital processes, communication tools, information resources and a combination of digital infrastructures based on regulatory provisions and mechanisms for their organization, administration and use”. “Digital twin” is a virtual digital model (prototype) of real physical objects or processes simulating internal processes, technical parameters and behavior of a real object under the effect of noise and environment” [Demidov A.Yu., Lukashov A.I., 2021: 31].

Federal Government Resolution No. 1646 of 2020 defines digital change as a combination of actions by public authorities to improve public administration, public service provision and performance of public functions through the use of electronic data and introduction of IT technologies into relevant operations.[11]

The Supreme Eurasian Economic Council also defined the concept of digital change in its Decision No. 1 of 2017 “On the main areas of implementing the EEU’s digital agenda until 2025”. The Eurasian Economic Commission believes digital change to be a manifestation of quality transformational changes embodied not only in individual digital changes but also in a principal transformation of economic structure to move value creation centers where digital resources are put in place.[12]

Public good increasingly includes today the outcomes of digital transformation of public administration strategically focused, firstly, at increasing real incomes and purchasing power of the population, secondly, improving investment attractiveness of the country and, thirdly, ensuring national security [Nazarenko T.S., 2023: 150].

M. Zherebtsov describes in his article different approaches to digital change. In the first place, he identifies the structural approach supported by the e-government’s infrastructural approach with static implementation plans extended over many years. The second approach is dynamic and based on flexible and iterative plans including the administrative process reform and promoting civil society’s involvement in public administration [Zherebtsov M., 2019: 583].

[11] On ways to ensure efficiency of measures for using IT technologies in operations of federal executive agencies and governance bodies of state extrabudgetary funds (annexed to the Provision on agency digital change programs). The Government of the Russian Federation Resolution No. 1646 of 10 October 2020 // Collected Laws of Russia, 2020, No. 42 (part III). Article 6612.

[12] Supreme Eurasian Economic Council Decision No. 2 of 11 October 2017. Available at: http://www.eaeunion.org/ (accessed: 12. 05. 2022)

The achievement of the said objectives translates into higher quality and more systemic public functions such as primarily the following:

public regulation of the national (including sectoral and regional) economic system;
development and implementation of public policies in various sectors and constituent territories;
public and municipal service provision;
control and supervision;
managing public and municipal property.

Introducing digital technologies into operations of public authorities, updating and, where inadequate or incomplete, amending the relevant regulatory framework as may be necessary merits special attention [Yastrebov V.B. et al., 2021: 142].

The objectives to be addressed today in public and municipal administration call for a need to transform the approaches to drafting and adoption of regulations by federal executive authorities. New digital technologies are required to make sure that regulation is not only adequate and timely but also non-controversial and capable of handling considerably more tasks of increasing complexity, with convenient services available to all those involved in the law-making process.

N. Popova argues in her paper that avoidance of overlapping and ambiguity of governance processes to make all e-documents legally relevant (generally binding) is an urgent requirement of digitization of public administration. Digital solutions to be created and developed require an innovation such as developing a shared state digital platform relying on a shared dataset synchronized with regard to public authorities on the basis of one-stop-shop principle for rapid and efficient administrative decision-making [Popova N.F., 2020: 50].

Digitizing law-making means improving this process through implementation and use of information technologies. This is the purpose of the shared national system for development and adoption of regulatory decisions that the Ministry of Economic Development has been creating in Russia since 2018. These efforts purport to create a shared digital space for regulatory drafting at federal executive agencies.

That cooperative work relies on collaboration technologies and intelligent tools for core administrative processes involved in law-making. Digitization allows major cost savings in regulatory drafting while offering a number of opportunities. As such, digitization brings about paperless communication between the parties to the drafting process; online joint editing of drafts; implementation control of regulatory decisions; transparency of the drafting process.

A system for non-public document exchange and execution control including through the use of cloud services purports to create a shared information environment and to avoid paperwork in record management based on automatic execution control and optimized drafting processes.

This system is designed to automate the Government Office processes and can be used as a standard solution in automating work processes at public authorities.[13] It is supposed to operate on the basis of the GosTech shared national platform which is an ecosystem for development and operation of state information systems and which includes shared hardware/ software and methodology.

Under sub-paragraph "e", paragraph 11, Section II of Government Resolution No. 1646 of 10 October 2020, the Ministry of Economic Development has approved the 2021–2023 departmental program of digital change (Ministerial order No. 876 of 30 December 2020) to introduce digital technologies in public administration for higher quality of prioritized socially important in-demand public (municipal) services across the country.[14]

The experiment to develop, migrate to and build up state information systems (including the SIS Normotvorchestvo) on the GosTech digital platform is envisaged by Government Resolution No. 1674 of 12 October 2020.[15] This digital platform was created for a number of reasons as 826 federal and 3,303 regional state information systems in operation by 2022 were normally designed in accordance with requests of federal executive agencies.

[13] On approving the Provision on the information system for intra- and interagency document exchange and execution control including through the use of cloud services. The Government of the Russian Federation Resolution No. 198 of 17 February 2022 // Collected Laws of Russia, 2022, No. 8. Article 1193.

[14] Available at: URL: https://cloud.consultant.ru/cloud/cgi/online.cgi?req=doc&base=LAW&n=393064&cacheid=2AD2C8CACA77D6DD946F5DFED1E6F45C&mode=splus&rnd=hmBvJw#lrU3E1UMvwdXaKuT1 (last accessed on 10 12 2021)

[15] On the experiment to develop, migrate to and build up state information systems and components thereof on the GosTech shared nationwide digital platform. The Government of the Russian Federation Resolution No. 1674 of 12 October 2020 // Collected Laws of Russia, 2020, No. 42 (part III), 19 October. Article 6637.

With each state information system offering largely standard functionalities (up to 80%), the federal authorities are "reinventing the wheel" instead on concentrating on provision of fast and convenient services to individuals and businesses. For this reason, it was decided to integrate all information systems into a single digital cloud (platform).

Different information systems are used at the federal and regional levels for digitized public administration. They are integrated into data systems of organizations involved in performance of public functions and provision of public and municipal services. At the regional level, it is exemplified by the Sverdlovsk Oblast information system for monitoring socioeconomic development (Sverdlovsk Oblast Government Resolution No. 977-PP of 27 December 2022). It is mainly designed to create a shared database of regional socioeconomic development indicators (linking them between themselves), ensure digital change and improve the quality of regional governance.[16]

This state information system is currently designed to manage regional public programs following a new governance system approved for regional programs by the Ministry of Economic Development and the Ministry of Finance in February 2023.

The adoption and implementation of administrative decisions through the use of state information systems requires that these systems are not only reliable in terms of technology but also sustainably operational in the legal environment. The issue of ownership to state information systems used for digital public administration should be resolved in legal terms. Legal certainty in this issue should be viewed as another prerequisite of efficiency and legitimacy of public decision-making.

The Federal Government operational guidelines for the period until 2024 (No. 8028p, para 13, approved by the Government on 29 September 2018) say the following: "The current regulatory environment falls short of the task to make the regulation of social relations more flexible and adaptable to the ever changing technological context. Despite systemic efforts to improve the business climate, the law still has numerous gaps and administrative barriers for development of businesses focused on information technologies and datasets". Overall, there is a need to develop a mechanism for managing regulatory changes in the digital economy to timely adapt normative regulation to the tasks of digital development.

[16] Official web portal of the Sverdlovsk Oblast, 2022, No. 37471, 29 December.

Fro developing Russia’s digital economy based on the use of data systems (with the Data Economy national project to be implemented starting from 2025), it is necessary, by introducing digital technologies and digital platforms, to save time and administrative costs involved in provision of public and municipal services (including with regard to law-making) and operation of agencies that make up the government system.

By 2024, the Federal Government has accomplished a number of priority tasks: introducing digital technologies and platform solutions in public administration and streamlining (standardizing) public and municipal service provision.

3. The Federal Concept of Machine-Readable Law

With messages and ideas of machine-readable law currently in the making, there are technologies that can convert legal provisions into machine-readable formats and regulations into a computer code thus opening up new opportunities for collaboration in the legal environment not only between man and computer but between computer systems themselves.

The Machine-Readable Law Development Concept[17] was drafted in autumn of 2022 pursuant to paragraph 1.23 of the passport of the national program “Normative Regulation of the Digital Environment” (developing a set of proposals to encourage cooperation between e-document exchange operators).[18]

The Concept became Russia’s first official strategic planning document in the area of machine-readable law and a major step forward to introduce the underlying technologies in law-making. Thus, it codifies the notions of technologies of machine-readable law and identifies their main development vectors. Machine-readable law incorporates legal provisions described in programming languages and text markups suitable for use in information technologies.

Moreover, machine-readable law also includes the tools for application of such provisions: data systems and software. Thanks to these technologies,

[17] Concept of machine-readable law approved by the Governmental Commission for digital development and use of IT to improve living standards and business environment. Minutes No. 31 of 15 September 2021.

[18] Passport of the federal project “Normative Regulation of the Digital Environment” approved by the Presidium of the Governmental Commission for digital development and use of IT for better living standards and business environment. Minutes No. 9 of 28 May 2019.

provisions will be translated into a computer code. The Concept of machine-readable law is itself focused at introducing the SIS Normotvorchestvo for seamless drafting, coordination and approval of regulations to enable federal executive agencies to work collectively on draft regulations in the Live mode.

Under the Federal Concept of Machine-Readable Law, the Normotvorchestvo that has equivalent systems in federal constituent territories will ensure seamless drafting, coordination and approval of regulations by federal executive agencies and enable different public authorities to work jointly on regulatory drafts.

Thus, it was used as a mold for the Document Approval information system of the Moscow City Government (Resolution No. 1239-PP of 25 September 2019) designed primarily to automate regulatory drafting, coordination and approval[19] and operated by the IT Department of the Moscow Government.

In 2022, the national legal system witnessed for the first time the emergence of a digital regulation: Federal Education Supervision Service order No. 1112 of 03 November 2022 created it through the use of functionalities of the Digital Regulation Constructor whose design is supervised by the Ministry of Digital Development and Mass Communications with methodological support by the Ministry of Economic Development.

The Legaltech segment is a promising development area and surely a baseline criteria for shaping a robust digital control loop in law-making (including for drafting bylaws). Registration of the first digital regulation is an example of important step in the right direction towards digitization of the national legal system.

Public authorities and development institutions need to continue developing such Legaltech instruments in close cooperation with the industry to put in place non-controversial regulatory framework in order to simplify the national law both at the federal and regional level.

4. Developing and Introducing the SIS Normotvorchestvo at Federal Executive Agencies

With approximately 19.000 federal, 1 million regional and 10 million local government regulations annually published in Russia, drafting aver-

[19] Moscow City Government Resolution No. 1239-PP of 25 September 2019 // Bulletin of the Moscow City Mayor and Government, 2019, No. 55.

agely takes 328 days and involves between 5 and 7 federal executive agencies. A draft document will pass about 12 approval loops, with processing of one integration taking up to 30 days.

A. Ivanov argues in his paper that there is currently no common legal approach to digitize bylaw drafting, with technologies advancing at much faster pace than the underlying practices. As such, digitization of law-making at executive authorities will require comprehensive systemic approach, consistent non-controversial enforcement practices and underlying mechanisms for legal regulation and application of machine-readable provisions [Ivanov A.A., 2018: 37].

The need to develop a state information system for bylaw drafting is explained by the following problems: belated drafting, lack of relevant draft versions, work with different draft versions and also impossibility of automatic enforcement (machine-readability).

The Ministry of Economic Development has been making efforts since 2019 to introduce the new SIS: national shared environment for collaboration between the parties to the law-making process" at federal executive agencies jointly with the Government Office.[20]

The state information system is developed under sub-paragraph "g", paragraph 21, and sub-paragraph "a", para 55, of the 2017–2030 Information Society Development Strategy for Russia (free, sustainable and safe communication between individuals and organizations, public authorities, local government bodies) and also paragraph 11 of the 2030 national objective "Digital Change" ("digital maturity" of the key economic sectors).

This digital project is related to implementation of the "Digital Economy for Russia" national program and directly aims at addressing relevant tasks and achieving strategic objectives. The currently developed draft Federal Government Resolution "On the state information system "National shared environment for collaboration between the parties to the law-making process" has passed all necessary stages of public discussions at regulation.gov.ru, but is still to be submitted to the Government Office.

The project of digital change is designed to enable public authorities, specific regulatory agencies involved in drafting work to draft regulations

[20] On the state information system "National shared environment for collaboration between the parties to the law-making process" (SIS Normotvorchestvo). Draft of the Federal Government Resolution.

in electronic form. A rationale for this project in public administration is the outcome to be achieved as envisaged by another federal project "Digital Public Administration" of the National Digital Economy Program: "Digital technologies have been introduced in the area of public administration and provision of prioritized socially important in-demand public (municipal) services".

Implementation of the project and completion of interventions envisaged by the road map will ensure the following: performance by public authorities of their policy development/implementation and/or regulatory functions in electronic form; lower time costs for those involved in drafting work through the use of mechanisms for collaborative document processing and semantic text analysis that enable automatic proposals of alternative wordings and circular changes; shared digital information environment for the parties involved in drafting allowing to use the information on pending regulations for governmental decision-making.

The procedure for operation of this state information system is established by Ministry of Economic Development Order No. 400 "On the federal state information system "National shared environment for collaboration between all parties to the law-making process involved in drafting regulatory decisions" of 09 July 2019.[21] Once implemented, this idea will introduce, in particular, collaborative drafting in order to systematize the law-making process at federal executive agencies in a shared space under the one-stop-shop principle.

This information system will be organized and operationalized under a federal executive agency order developed for internal operations of departments within the Ministry of Economic Development. Another order of the Ministry also approved the functional "backbone" consisting of the following: managing instructions, document package, document approval, editing and review, storage, notification, managing the document package approval route and document package register.

The main difference between the new SIS and present-day public portal for regulatory drafting regulation.gov.ru is that the former is non-public and designed only for federal executive staff and expertise entities such as the Expert Council under the Federal Government and the Civic Chamber of Russia.

[21] Ministry of Economic Development Order No. 400 of 09 July 2019 // SPS Consultant Plus.

The system regulation.gov.ru was set up under Federal Government Resolution No. 851 “On the procedure for disclosure of information on regulatory drafting and outcomes of public discussion of regulatory drafts by federal executive agencies” of 25 August 2012[22] passed in furtherance of Presidential Decree (“Decree of May”) No. 601 “On the guidelines for improvement of the public administration system” of 07 May 2012.[23]

In 2023, the new SIS became operational at 37 departments of the Ministry of Economic Development, with pilot connectivity available to a number of federal executive agencies and integration with regulation.gov.ru being achieved.

In 2024, the system is expected to become fully operational and integrated with the state e-document exchange system, with the instructions register and database of draft standard documents, forms and letterheads to be put in place, the development intellectualized and migration to the GosTech platform prepared.

In 2025, it is expected to make the system mandatory for drafting work across federal executive agencies, achieve its integration with pravo.gov.ru and State Duma law-making system and develop a mobile app for online downstream tracking of draft regulations (until adopted and signed by heads of the federal executive agencies with enhanced digital signature).

The technological processes implemented in the new SIS with relation to the lifecycle of regulatory drafts will be mandatory for federal executive agencies and recommended to other public bodies and organizations (not specified in the draft Resolution).

The specific functionalities of the system will ensure information interaction and integration with outside systems and resources of federal executive agencies and Government Office including those involved in planning, managing and controlling the execution of regulatory drafting instructions issued by the Office and Chairman of the Government.

Also, the system’s functionalities will ensure integration with those run by the public authorities in constituent territories (regional executive and legislative bodies), thus forming a shared register of regional regulations, as well as integration with local governments’ systems to put in place a shared register of municipal regulations in the web portal for legal information.

[22] Rossiyskaya Gazeta, 2012, No. 200, 31 August.

[23] Rossiyskaya Gazeta, 2012, No. 102, 09 May.

The requirements to have a shared format of draft regulations for text markup complying with universal specifications for electronic issuance, machine processing, posting, storage and dissemination of regulatory texts, as well as with shared formats are to be approved by the Federal Guard Service under Presidential decree No. 90 of 03 March 2022.[24] Once approved, the shared format of the pending regulatory draft will need to be synchronized with the shared format of the regulatory draft. Further development of regulatory draft markups will ensure automatic enforcement allowing to form a register of normative requirements in accordance with preset parameters.

Digitization of regulatory drafting in the new SIS contains three cyclic components.

The first component is drafting (underlying instruction in the shared register integrated with the state e-document exchange system, automatic planning — draft routing with control points, regulatory drafting with a built-in word processor based on legal drafting rules, electronic draft markup, alternative wordings).

The second component is approval (while in-house approval takes place in a shared file, outside approval requires the system to generate documents containing all comments/amendments, draft refinements and settlement of differences in the form of tables and minutes of conciliation meetings).

The third component is submission and signature (transfer of a shared relevant document package with complete information on the approval history and persons in charge (draft versions, tables of comments and differences) to the state e-document exchange system and automatic archiving of regulatory drafts).

As for a positive effect, this digitization process will save 30% of time spent on drafting, approval and adoption of regulations while providing federal executive agencies concerned with online shared access to relevant regulatory drafts and editing tools.

Further development of this digital project assumes automatically generated draft amendments, checking whether the provisions of effective regulations need to be amended, identifying semantic contradictions with other regulations, checking the links within and between regulations for correctness, expanded analysis of the given subject field (for example, an-

[24] Collected Laws of Russia, 2022, No. 10. Art. 1470.

ti-corruption and legal due diligence), recommendations to users on legal drafting rules to be followed, searching for regulations governing similar legal relationships, builder of regulatory drafts.

In January 2024, the Deputy Government Chairman (head of the Government Office) chaired a meeting on integrating the new SIS with operations of federal executive agencies since 2025. As a follow-up, an official instruction was issued to set up a project office bringing together representatives of federal executive agencies to propose methodological and practical recommendations for interaction with this state information system.

Each state information system for law-making has its upsides and downsides. The upsides include monitoring key indicators at the level of Deputy Chairman, instruction routing and execution control, draft approval and signing, fast document exchange, online text processing, online commenting and amending, shared markup format, IT infrastructure free of foreign software.

In 2024, the key performance indicators for the state information system are: about 1,500 users, 5 federal executive agencies involved in regulatory drafting and approval, 14 federal executive agencies involved in regulatory approval, 7 types of regulations, 20 draft regulations submitted to the Government in structured electronic format.

5. Digitizing Law-Making at Executive Agencies: International Experience

The rationale behind the choice of countries was that the United States is often believed to be among the most advanced countries in terms of computer technologies in general and artificial intelligence in particular while Kazakhstan and Belarus are Russia’s neighbors, members of the Eurasian Economic Union and partners crucial for foreign policy, with their experience likely to be useful for the development of digital law-making in Russia.

The most relevant public initiative in the United States is related to the use of AI technologies to identify and remove outdated and redundant provisions of federal regulations. This initiative has yielded positive results. Following the first successful experiment at the Department of Health and Social Services in 2019, it was decided to run an automatic analysis of law to identify redundant and archaic provisions also at the Departments of Labor, Transport and Agriculture. This AI-based technology finally detected hundreds of errors and outdated provisions, such as the one requiring the sender to deliver documents by fax. This and other AI-based technologies are part of the national AI development strategy in the United States.

While this strategy is reflected in different regulations, the principal guiding document is the 756-page long Final Report of the National Security Commission on Artificial Intelligence.[25] According to the non-public National Security Presidential memorandum (NSPM) of 11 February 2019 ("Protecting the United States advantage in artificial intelligence and related critical technologies"), the U.S. strategy purports to protect AI technologies critical for economic interests and national security from hostile state and non-state actors [Kamolov S.G. et al., 2023: 92].

As for EEU members' experience of regulating the development of bylaws and amendments thereof, it is worth noting Law of Belarus No. 130-3 "On Regulations" passed 17 July 2018, where Article 2, para 15 specifies that the text of regulation under control is its version in force as of specific date, drafted on the basis of original text and regulatory amendments and posted to the reference database of legal information of the Republic of Belarus.[26]

Since 01 July 2022, public authorities (agencies) of Belarus perform regulatory drafting via the Normotvorchestvo automated information system for support of the law-making process under Presidential Decree No. 415 "On improving the speed and quality of law-making" of 17 November 2020.[27]

This information system was developed in furtherance of an instruction by the President of Belarus to digitize law-making and implement information technologies on a wide scale at all stages of regulatory drafting and adoption. However, there are exceptions: requirements of the Decree do not apply to regulatory drafts containing state secrets or non-public information.

This state information system will allow Belarus to put in place a complete and transparent drafting cycle for all kinds of regulations, dramatically reduce the amount of correspondence between public agencies, ensure traceability, work with different versions and collaborative draft editing by public agencies, simplify, digitize and, consequently, expedite the law-making process, as well as to start implementing AI components within the

[25] Available at: https://digital.library.unt.edu/ark:/67531/metadc1851188/ (accessed: 22.11.2023)

[26] National web portal of the Republic of Belarus, 2018, No. 2/2568, 31 July.

[27] Ibid. 2020, 1/9332, 19 October.

system. In the future, the system under development will have the capability to adopt machine-readable regulations assumed to be enforced by AI.

In the Republic of Kazakhstan, the Government has passed Resolution No. 827 “On approving the “Digital Kazakhstan” state program” of 12 December 2017[28] to create the Zandylyk automated information system as a structural component of the said state program developed by the General Prosecutor’s Office jointly with the Supreme Court of Kazakhstan. The system can check regulations drafted (or adopted) by a prosecutor or judge for compliance with formal requirements of Kazakhstan’s criminal law and law of criminal procedure, and collect judicial statistics across regions in the Live mode.

Conclusion

The current evolution in law should be regarded as an intermediate stage, with digital technologies fitting into an established legal system to extend and expand the effect of traditional legal instruments rather than operating as new independent legal realities. The doctrine should further theoretically develop and accumulate the experience of synergy between legal tools and computer hardware/software.

The introduction of modern digital technologies in public administration offers an enormous potential at the price of possible risks. The use of “cross-cutting” digital technologies in public administration will result in an efficient governance system. They will expedite interagency collaboration, increase the extent of protection of state information systems, reduce the number of civil servants, ensure availability and high quality of public and municipal services, and accelerate critical decision-making both at the federal and regional level.

Meanwhile, the processes of digitization are fraught with major challenges both for the government and society. The main risks of introducing digital technologies in public administration are legal gaps regarding their use in different governance areas [Zubarev S.M., 2020: 29, 39].

The SIS Normotvorchestvo is critical for law-making efficiency. However, the system is facing multiple problems that prevent it from becoming fully digital.

[28] On approving the Digital Kazakhstan state program. Government of Kazakhstan Resolution No. 827 of 12.12.2017. Available at: URL: https://primeminister.kz/assets/media/gosudarstvennaya-programma-tsifrovoy-kazakhstan-rus.pdf (accessed: 22.11.2023)

The implementation downsides of such digital projects are also worth noting. The main obstacle to making the new SIS a fully digital law-making tool for federal executive agencies are technical problems such as outdated systems and compatibility issues. Organizational issues including inadequate coordination between federal executive agencies and their subdivisions (departments, directorates, offices) and a lack of clear agency-level digitization strategies and plans also obstruct the digitization of law-making.

Legal problems including a lack of mature regulatory framework for digitizing law-making — particularly of a Government-adopted regulation — makes it difficult to integrate the information system into operations of federal executive agencies; a lack of agency-level regulations also creates barriers for the system's use at ministries and departments. Another difficulty is inadequate mechanisms for intellectual property protection which should be improved when developing and introducing the said state information system.

There are also cultural issues like low awareness and reluctance of civil servants to accept new technologies, fear of dismissal and change to long-established processes. These are barriers for digitizing law-making at federal executive agencies. At the Ministry of Economic Development the process of document approval in the form of orders takes place in the Normotvorchestvo platform, only to cause concern among the staff due to technical defects of the system.

The processes of adoption and implementation of administrative decisions will bring about legal risks which normally have adverse implications and are harmful for law-protected interests of individuals, society and public corporations.

Nevertheless, the digitization of law-making holds the promise of considerable benefits for federal executive agencies including higher efficiency, major time and cost savings of drafting and implementing regulations, better access to information and broader involvement of civil society [Amelin R.V., Channov S.E., 2023: 250].

For seamless law-making at federal executive agencies, it is feasible to use AI technologies from two neural networks. Normative provisions and legal terminology (words and their combinations) as well as the forecasted outcome should be fed to the first network. The second neural network will learn from the first by processing and analyzing the database information and proposing plausible versions [Gvozdetsky D. S., 2019: 23].

For successful digitization of the SIS Normotvorchestvo it is useful to develop a clear strategy and plan (roadmap for the system's implementation at federal executive agencies — in simple terms, approval of the agency-level program of digital change), and remove the above barriers including technical, organizational, legal and cultural aspects.

The process of governmental decision-making in the context of digitization is inevitably subject to change that assumes not only a special digital infrastructure to be created but also relevant regulation, with the digital infrastructure objectively to become the basis for administrative decision-making. Moreover, the benefit from the governance process as a whole and individual administrative decisions in particular will depend on the development level of digital infrastructure.

Collaboration between all parties to digital change — federal, regional, municipal authorities, business community, science and education, civil society organizations — assumes mutually beneficial cooperation at the regional and interregional levels for exchange of experience in adopting new knowledge, introducing breakthrough digital technologies and applying relevant decisions for multiple positive effect [Samorukov A.A., 2022: 12].

The studies of and practical efforts towards digitization of law-making at federal executive agencies will allow to develop and refine the SIS Normotvorchestvo and ensure more transparent and open involvement of public authorities and civil society in law-making.

References

1. Amelin R.V., Channov S.E. (2023) *How digital technologies have revolutionized law.* Moscow: NORMA, 280 p. (in Russ.)

2. Demidov A.Yu., Lukashov A.I. (2021) Specific approaches to digital change in public administration. *Gosudarstvennaya Sluzhba*=State Service, no. 1, pp. 28–34 (in Russ.)

3. Gvozdetsky D.S. (2019) Digitizing law-making at government agencies: theoretical aspect. *Pravovoye gosudarstvo: teoriya i praktika*=Rule of Law: Theory and Practice, no. 4, pp. 20–24 (in Russ.)

4. Ivanov A.A. (2018) On the depth of computerization in law. *Zakon*=Law, no. 5, pp. 35–41 (in Russ.)

5. Kamolov S.G. et al. (2023) Predominant of national AI development strategies in Russia, Germany and the United States. *Voprosy gosudarstvennogo i munitcipalnogo upravleniya*=Issues of Public Administration, no. 2, pp. 85–105 (in Russ.)

6. Kabytov P.P., Starodubova O.E. (2020) The impact of digitization on realizing executive power. *Zhurnal rossiyskogo prava*=Journal of Russian Law, no. 11, pp. 113–126 (in Russ.)

7. Legal Support of the Implementation of National Socioeconomic Development Projects in Russia (2021) V.B. Yastrebov (ed.). Moscow: Prospekt, 208 p. (in Russ.)

8. Lipen S.V. (2019) CIS regulatory drafting laws as indicators of informatization and digitization processes in the legal system. *Aktualnie problemy rossiyskogo prava*=Current Issues of the Russian Law, no. 8, pp. 22–33 (in Russ.)

9. Nazarenko T.S. (2023) Digital change in public administration as strategic public good. *Strategirovanie: teoriya i praktika*=Strategizing: Theory and Practice, no. 2, pp. 140–157 (in Russ.)

10. Popova N.F. (2020) The need in digitization of public administration in Russia. *Administrativnoye pravo i protcess*=Administrative Law and Process, no. 2, pp. 48–53 (in Russ.)

11. Samorukov A.A. (2022) Digital change in public administration. *Vestnik Povolzhskogo instituta upravleniya*=Bulletin of the Volga Institute of Administration, no. 1, pp. 4–13 (in Russ.)

12. Sorokopud M.C. (2024) Digital change in public administration in Russia: reasons and prospects. *Vestnik Belgorodskogo yuridicheskogo instituta MVD imeni I.D. Putilina*=Bulletin of the Belgorod Juridical Putilin Institute of the Interior Ministry, no. 1, pp. 18–24 (in Russ.)

13. Stenkin D.S., Chuchadeev D.A. (2023) Introducing digital technologies at executive agencies (exemplified by control and supervisory agencies). *Rossiyskoye pravo: obrazovanie, praktika, nauka*=Russian Law: Education, Practice, Research, no. 4, pp. 64–75 (in Russ.)

14. Zherebtsov M. (2019) Taking stock of Russian e-government. *Europe-Asia Studies*, no. 4, pp. 579–607.

15. Zubarev S.M. (2023) *Efficiency of public administrative decisions in the context of digitization*. Moscow: Prospekt, 184 p. (in Russ.)

16. Zubarev S.M. (2020) Legal risks of digitization of public administration. *Aktualnie problemy rossiyskogo prava*=Current Issues in the Russian Law, no. 6, pp. 23–32 (in Russ.)

Information about the author:

A.A. Doschatov — Officer, the Ministry of Economic Development of Russia

The article was submitted to editorial office 27.05.2024, approved after reviewing 17.06.2024, accepted for publication 17.06.2024.

Legal Issues in the Digital Age. 2024. Vol. 5. No. 2.
Вопросы права в цифровую эпоху. 2024. Том 5. № 2.

Review
УДК: 342, 347
DOI:10.17323/2713-2749.2024.2.148.169

Digital Platforms in the Focus of National Law

Ludmila Konstantinovna Tereschenko[1], Olesya Evgenievna Starodubova[2], Nikita Alekseevich Nazarov[3]

[1, 2, 3] Institute of Legislation and Comparative Law under the Government of the Russian Federation, 34 Bolshaya Cheremushinskaya Str., Moscow 117218, Russia,

[1] Adm2@izak.ru

[2] olesyastarodubova@gmail.com

[3] naznikitaal@gmail.com

Abstract

The paper contains an overview of the research workshop *Digital platforms: new environment for collaboration* at the Institute of Legislation and Comparative Law under the Federal Government (ILCL) on 23 April 2024, and findings of the expert survey *Digital platforms in the focus of national law*. Presentations by participants were systematized by the authors of the overview around the most relevant subjects related to digital platform operations: general issues of digital platform regulation; digital platforms' impact on human rights; digital platforms in public administration; digital platforms in private law, criminal law and specific branches. The workshop was attended by researchers representing the Institute of Legislation and Comparative Law, Kutafin Moscow State Law University, National Research University–Higher School of Economics, Moscow State University, Russian Academy of National Economy and Public Administration under the President of Russia, Plekhanov Russian University of Economics, Orenburg State University, Siberian Federal University, Gubkin Russian State University of Oil and Gas, Saint Petersburg University of the Interior Ministry, MTS Joint-Stock Company, etc. Legal themes under discussion included the legal nature of digital platforms; digital platforms' impact on the economy, public administration, human rights and the underlying changes to their amount and content; associated risks; digital platforms and anti-trust regulation; networking effect of digital platforms; labor relations and platform employment; labor right protection and digital platforms; dangers of discrimination related to digital platforms.

Keywords

digital platforms; platform economy; data economy; platform employment; crowdfunding platforms; recommendation technologies; personal data; artificial intelligence; data security; discrimination; trusted technologies.

For citation: Tereschenko L. K., Starodubova O.V., Nazarov N.A. (2024) Digital Platforms in the Focus of National Law. *Legal Issues in the Digital Age,* vol. 5, no. 2, pp. 148-169. DOI:10.17323/2713-2749.2024.2.148.169

The research workshop *Digital platforms: new environment for collaboration* was held at the Institute of Legislation and Comparative Law (ILCL) on 23 April 2024**.** Below is the review of the Research Workshop "Digital Platforms: New Environment for Collaboration" and Findings of Expert Survey

1. General Issues of Digital Platform Regulation

In opening the session, the workshop moderator L.K. Tereschenko, Senior Researcher at the ILCL, Doctor of Sciences (Law), Associate Professor, Honored Lawyer of Russia, Russian Academy of Sciences expert, has noted that digital change has brought about numerous new things and phenomena that did not exist before, just as associated relations. It would be safe to say these include digital platforms created and operating in both private and public domains. In the private domain, digital platforms tend to be viewed as a business model for online connectivity between sellers and buyers to exchange products, services and information. With digital platforms at the core, the market structure is changing. Digital platforms are transforming the way markets operate by exercising new forms of clout on the market, competition and human rights. While downplaying the role of law, digital platforms take the regulatory initiative, only to replace law in a number of cases for specific agents, with both consumers and sellers in a weak position vis-à-vis platform owners.

A.A. Efremov, Senior Researcher, ILCL laboratory for IT regulation and data protection, Doctor of Sciences (Law), Associate Professor, discussed the legal nature and prospects of platform law.

The speaker has underlined digital platforms were a key vector of the Data Economy national project. He has noted that the term "digital platforms" has made its way to the national legislation, with the relevant bylaws developed in absence of the generally acknowledged approach. Digital platforms are collaborative tools for agents of public relations permeating all spheres of life: economic, cultural, public administration, with expansionary trend affecting both the regulatory mechanism and its components and the implementation of human rights.

The speaker has identified the following approaches to the definition of platform law:

comprehensive inter-system regulation supported by international law;
specific local regulation applicable to specific platforms (ecosystems).

He has outlined the development prospects of platform law:

digital platforms as a tool of geopolitical and economic struggle: prohibitions and competition of extraterritorial jurisdictions;
standardization of requirements to digital platforms at the legislative level within specific countries, harmonization of regulatory approaches within the framework of international organizations;
promoting public regulation, especially as part of anti-trust, consumer protection, personal data and labor laws.

A. Minbaleev, Head, Chair of information law and digital technologies at Kutafin Moscow State Law University, Doctor of Sciences (Law), Associate Professor, RAS expert, discussed in his presentation the legal nature of digital platforms.

He referred to the example of China, a leading economy in terms of IT development, with the Chinese government not so much regulating the relations themselves as defining the operating rules for key digital platforms that implement these relations. The government will appoint the main operators in the given field, establish the underlying operational requirements and accomplish deregulation by delegating some authority to define policies in a number of aspects including meta-universe, personal data, artificial intelligence, trusted technologies etc.

Russia has adapted much the same practice, with specific issues resolved by major digital platforms followed by subsequent integration. This way of regulating information relations is the legacy of the fact that major digital platforms actually develop and revolutionize technologies and technology-related relations.

Restricting digital platform operation is another vector of public regulation visible in anti-trust policies, human rights and personal data protection and e-trade.

The speaker has suggested a number of ways to conceptualize digital platforms in legal terms: object-based approach: digital platform is a complex set of information relations bringing together several information objects such as ITC network, websites, data systems, information technologies, information, data. Comprehensive regulation: provisions governing all information objects are eq ually applicable to regulation of digital platforms.

Legal fiction based models:

A) agent-based approach: digital platform is regarded as person at law and a party to legal relations — information, civil, labor, etc. — and can have a set of rights and duties. This model is questionable, its advocates comparing it to the concept of e-person.

B) digital platform as information environment for collaboration between the said agents acting on the basis of certain resources with involvement of various social media and other resources. It is possible to clearly identify the range of agents and objects of information environment: agents — software developers, business agents integrated into the digital platform, users, service providers ensuring operation of digital platforms. This approach is legally convenient as it allows to single out the entire range of agents and objects and to regulate the underlying relations on this basis.

It is crucial to define digital platforms in legal terms. The relations involved in their operation should be regulated on the basis of concepts and objects existing in the legislation.

P. Kabytov, Acting Head, ILCL laboratory for IT regulation and data protection, Candidate of Sciences (Law) , discussed the specific status of digital platform operators.

The speaker has noted that digital platform operators possess specific rights or powers that are quasi-public in terms of impact on users, whether they offer goods and services or post content via the platform or consume these goods, services and content. In particular, digital platform operators impose mandatory rules on users (of indefinite range), exercise coercive power, resolve disputes between platform users (sellers and buyers).

Due to deviations in platform operator behavior including those resulting in violation of rights and legitimate interests of users, there is a need to introduce requirements to specific parts of platform rules to be checked for compliance with legislation, as well as to set the basic principles of the procedure established by operators to challenge coercive action and dispute resolution policies.

A. Antopolsky, Associate Professor at the Plekhanov Russian University of Economics, Candidate of Sciences (Law), discussed in his presentation the question of conceptual framework for digital platforms. He has noted that a legal definition should be based on clear and, importantly, usable (operable) criteria allowing to distinguish digital platforms from other information systems. Meanwhile, most definitions used in official documents fail to meet these criteria.

The speaker also emphasized that the risks related to digital platforms in the public domain include overcentralized governance processes. These risks have not been adequately addressed. While in a compact, decentralized system, defects constantly faced by ordinary users (private individuals and lower-level employees) could be easily identified and removed, they will often remain hidden for system operators and developers in a more complex centralized system.

2. Digital Platforms' Impact on Human Rights

V. Naumov, Senior Researcher, Information Law and International Data Security Desk, Institute of State and Law of the Russian Academy of Sciences, Head, Intellectual property and ITC section and Managing Partner, Nextons Saint Petersburg office, Doctor of Sciences (Law), focused on the issue of exercising the right to refuse digital platform technologies.

The rapid pace of technological change radically transforms social relations resulting in a new dimension of digital divide between generations, only to pose a challenge to mankind maintained by digital platform owners. The loss of energy that once emanated from human communication affects the foundations of human relations. Global propaganda of digital life and digital services results in existential threats in the context of geopolitical risks and influences (as users originally relied on Western digital platforms, there are issues of migration to analogous domestic platforms).

As always, the legal system is a laggard, with legal and technical terms out of grip with the reality. The legal academic community does not take part in multi-disciplinary discussions. The use of digital platforms largely follows in the wake of fashion while the regulatory plans for digital change outlined in strategic planning documents fail to be implemented in full.

The speaker reported the findings of a survey related to the right of choice of technologies conducted among daily users of digital technologies [Fedotov M.A., Naumov V.B. et al., 2024: 8–28]. The issue of refusal of digital technologies is becoming critical. The current priority of technological communication with the government without involvement of human operators — from the integrated portal of public and municipal services (functions) to the GosTech integrated nationwide digital portal — is causing serious concern even among users with a high level of computer literacy and good knowledge of digital technologies.

As a matter of conclusion, V. Naumov has identified the areas where the right to refuse digital technologies can be implemented:

amending Federal Law No. 149-FZ "On Information, Information Technologies and Data Protection" of 27 July 2006 (the most organic way);
amending Federal Law No. 152-FZ "On Personal Data" of 27 July 2006;
amending Article 10, Federal Law No. 135-FZ "On Protection of Competition" of 26 July 2006 (Article 10 "Prohibition to abuse a predominant position"). Where man interacts with technologies, he is objectively a weaker party (though not in the economic sense), and there is discrimination.

M. Bundin, Associate Professor, Chair of administrative and financial law at Lobachevsky National Research University of Nizhny Novgorod, Candidate of Sciences (Law), discussed the issue of personal data protection with regard to digital platforms.

The theme of platform regulation is closely related to that of platform-based personal data processing, with the transparency of the underlying data processing algorithms and the competition of legal grounds for data use being among the most fiercely debated issues. The Roskomnadzor repeatedly recalled the need to draw a distinction between the legal grounds for processing personal data of different legal nature — the terms of service and personal data owner's consent to process the data. The terms of service is a type of private law contract between the service owner and the user amendable under civil law whereas personal data processing consent is a public law instrument that allows the owner of information to define and/or change the legal regime applicable to information (personal data). The final goal of the Roskomnadzor is to introduce constraints on consents to process personal data that providers tend to impose indiscriminately as the terms of service.

However, it is worth recalling that online services are often free, only to be later monetized by service owners through possible use of users' personal data for other purposes. Any restrictions on such "secondary" use will therefore backfire on users as platform owners will be unable to offer services for free.

It is high time to discuss in detail and elaborate on the issue of delineating legal grounds of the terms of service and consent to personal data processing with regard to digital platforms and online services, especially those offered for free.

E. Diskin, Candidate of Sciences (Law), Researcher at the National Research University–Higher School of Economics, argued that censorship at digital platforms is a form of discrimination.

E. Savchenko, Researcher at the ILCL department of social law, analyzed the impact of digital platforms on human rights. When discussing the impact of digital platforms on human rights, one has to remember that a digital platform is above all an information system; however, in view of the current progress of information technologies, there is a need to specify this definition given in the Law "On Information, Information Technologies and Data Protection". Digital platforms — for instance, in the cultural domain — change the format of sharing cultural values through the so-called "digital rights", one of which is the right of access to digital platforms in the cultural domain, something that, as some researchers believe, can be viewed as access to the Internet. However one can have access to the Internet but be deprived of information, for example, on digital platforms in the cultural domain created by executive authorities, public, commercial and non-profit organizations for concerted action to implement people's constitutional right of access to cultural heritage and participation in cultural life of the country. For this reason, the speaker believes, the access to digital platforms is part and parcel of the right of access to information in the Internet.

Convergence of digital platforms and human rights significantly transforms the content of labor relationships as observed in the following presentations.

T. Korshunova, Candidate of Sciences (Law), Senior Researcher at the ILCL law and social security department, discussed the main trends of extending employment and social guarantees to digital platform workers, in particular, engaged in delivery and taxi services in some countries (Italy, Norway, Germany), and made a presentation of newly-published guide *Judicial Practices and Development of Labor and Social Security Law* [Korshunova T.Yu. et al., 2024: 3–248].

S. Kamenskaya, Candidate of Sciences (Law), Senior Researcher at the ILCL law and social security department, noted the increasing importance of social protection of those working at digital platforms without the status of workers in the classical (traditional) sense and identified the issue of voluntary adhesion of self-employed and other individuals with non-typical forms of employment to the social insurance system.

M. Stepanov, Candidate of Sciences (Law), Associate Professor, Senior Researcher at the ILCL `department of legal theory and multi-disciplinary studies, discussed digital platforms in the context of protecting individual labor rights.

The speaker noted that digital labor platforms vividly exemplify the impact of digital technologies on the processes of recruiting, organizing and managing the participation of staff in production operations. Because of problematic regulation of platform employment, Russia still does not have specific law in this domain. At the same time, there is an urgency to regulate these relations to protect labor rights of individuals. Meanwhile, it should be borne in mind that regulating platform employment on the basis of existing labor law provisions can be damaging to the development of this economic segment.

3. Digital Platforms in Public Administration

O. Stepanov, Senior Researcher at the ILCL center for judicial law, Doctor of Sciences (Law), Professor, analyzed the prospects and risks of implementing the Government as Platform concept.

The development of the Government as Platform Concept is closely related to operating parameters of the Universal Biometry System (UBS), with a platform solution for the UBS expected to be developed on the voluntary basis. Meanwhile, the standard biometry potential is currently rather restricted. As a result of attacks on personal data storage and identity thefts, the UBS is extremely slow to develop. Moreover, the doctrinal discussions often suggest that personal data theft is used not only to get credit in a fake name but also by terrorists in an attempt to legalize the origin of criminal funds via “unduly charitable donation” that could be made via a stolen digital identity with a full set of digital profile attributes. Here we deal with spoofing made possible by the technical opportunity to mask one set of data with another via substitution and falsification of the ordinary sample.

The UBS development prospects will be considerably brighter if the system is positioned in one package with the universal identification and authentication system (UIAS) as a federal register of digital economic agents rather than a system for managing biometric data and a remote authentication platform.

A. Kalmykova, Candidate of Sciences (Law), Senior Researcher at the ILCL department of administrative law and process, described the experience of regulating the use of digital platforms in the supervisory and licensing domain in EEU member states.

As a result of reform, supervisory and licensing operations, with exception of certain aspects, have become fully electronic. In terms of its functionalities, the control and supervision portal is actually a digital platform enabling supervisory authorities to communicate between themselves and with business agents, as well as allowing communication between those subject to supervision. In particular, the multi-functional portal allows to monitor supervisory operations. While the term "digital platform" is not applied to the service, it is defined as a combination of information systems related through common algorithms and allowing agents to communicate between themselves. As a matter of conclusion, A. Kalmykova has noted that a legal fiction is not applicable to digital platforms in the public domain, with the latter to be viewed exclusively as an object of regulation. This approach is also shared by EEU member states.

V. Lagaeva, Postgraduate Student, Chair of information law and digital technologies, Kutafin Moscow State Law University, discussed in her presentation the details of legal regulation of digital platforms in the area of public control (supervision).

As was argued by D. Gvozdetsky, Senior Lecturer, Chair of state law and criminal law at the Plekhanov Russian University of Economics, digital platforms significantly simplify information communication between government and individuals in routine operations of the federal executive agencies. Moreover, these software products, along with other innovative solutions, are also reflected in the National Economy of Russia projects including those used in agency-level law-making primarily at the stage of developing legal solutions at the federal and lower levels.

Development of digital platforms is outlined in a number of public programs and concept papers (for example, the draft concept of the shared national environment for collaboration between all parties to the law-making process in drafting regulatory solutions developed by the Ministry of Economic Development and supported by other federal executive agencies).

In analysis of the dynamics of law-making solutions in the context of introducing innovations into the law-making cycle, the speaker also noted that the issue of more sophisticated software products based on the algorithmic mechanism for drafting standard law-making solutions is expected to be discussed in the near future (5-10 years) as part of implementing state programs (concepts) at relevant venues of the federal executive agencies.

4. Digital Platforms in Private Law

S. Chekhovskaya, Candidate of Sciences (Law), Senior Researcher at the ILCL center for private law, noted the role of digital platforms as a networking method for market participants.

The digital component is fully integrated into the modern market structure. There is a need to study the functional role of digital platforms in trade turnover as market networking organizers. As a way of economic networking, digital platforms operate so that market agents communicate through access to a specifically created IT system as a combination of integrated digital services for collaboration between all stakeholders under the rules set by the operator. The procedures envisaged by the rules are fixed and implemented by the underlying algorithm. To use the digital platform, market participants thus need to comply with both technical connectivity requirements and the rules of conduct.

The information environment for collaboration between market participants associated with technological infrastructure implements the principal advantage of the digital platform as a model based on user data collection, something that allows to maximize the value of multiple user cooperation and higher amount of user data available. These aspects affect the choice of legal means for digital collaboration between market agents: it is critically important to address legal issues of access to the platform, security/confidentiality of digital economic operations, use of special contractual patterns etc.

E. Obolonkova, Candidate of Sciences (Law), Senior Researcher at the ILCL center for private law, stressed the importance of digital platforms for attracting residents of the territories with special terms of doing business, and offering a shared service to all such territories.

The federal law makes it possible to create in Russia a range of territories with special terms of economic development because of the country's vast expanses and varying geographic and economic conditions in regions. With no major difference in either qualifications required for a resident status in such territories or available preferences, the doctrine allows to pass a general law defining the principles of their operation. While this initiative is not yet implemented, it will be of practical benefit to create a shared digital platform for such territories, something that will allow potential investors to select a territory with optimal terms of doing business on the basis of required parameters and to file electronic documents for acquiring the resident status. In the current context, this will boost investment activities and reduce financial costs for both investors and public authorities.

M. Tsirina, Candidate of Sciences (Law), Senior Researcher at the ILCL, presented an analysis of digital change for alternative resolution of disputes.

The list of LegalTech solutions include the technologies for more efficient administration of justice which is currently among the most demanded domains for innovative technological tools. These cover different platforms and applications for facilitating and optimizing the administration of justice, as well as technologies for alternative dispute settlement that are similar in many respects to those for better administration of justice.

The progressive introduction of information technologies at mediation courts (including international business arbitration tribunals) has been encouraged by the development of automated management of legal proceedings. Today public courts in a number of industrial economies, such as the United States, Canada, part of the EU states, South Korea, Japan, Indonesia, exhibit trends for optimizing dispute resolution procedures including where the parties use legitimate innovative web-based judicial technologies whose progressive and inevitable development is significantly affecting international and national arbitration practices also based on rather active use of alternative mechanisms in the form of ADR and ODR remote e-technologies.

The principal difference of online dispute resolution (ODR) from classical conciliation and arbitration lies in the use of e-venues for online examination of disputes (so-called technological online dispute resolution platforms that comprise computer software (including to draft, send, receive, store, exchange or otherwise process a message, ensure security of the relevant data and operation of a network of sellers and buyers involved in exchange of goods), databases, websites, domain names, systems). Online dispute resolution provides the parties with an opportunity to control the procedure and engage, apart from the arbitrator, a mediator (neutral party acting as the technological platform administrator) to technically assist with dispute resolution. This process assumes that dispute resolution (including initial registration, neutral appointment, oral hearings and discussions) largely takes place online with possible involvement, apart from the third party, of the "fourth party", a special application (artificial intelligence) that will create for "disputing parties a range of opportunities along the lines of the third party's role in the conflict". While the fourth party can act from time to time as neutral mediator by automating the negotiations in the course of dispute resolution, it will often play the role of a neutral third party to assists in the search of settlement options".

Online dispute resolution is a promising mechanism with prospects of future development (as regards providing the parties with variable terms of transition to online dispute resolution stages: online negotiations via both ODR platforms and face-to-face meetings or online broadcasts; access to the system for targeted "big data" processing; ensuring protected access to "electronic deliberations rooms"; using algorithms for automatic online resolution of standard disputes etc.).

5. Digital Platforms in Criminal Law

O. Zaitsev, Senior Researcher at the ILCL Center for criminal law and criminal procedure, Doctor of Sciences (Law), Professor, presented an analysis of the impact of new digital technologies on rights of the parties involved in criminal proceedings.

The criminal procedure is facing a phased transition from paper to electronic documents to be created on digital platforms in the law enforcement system. The priority of trusted data over paper files will allow to abandon paper altogether, with all processes transferred to the digital paper-free format. This transition should be stipulated by non-interference of third parties with criminal proceedings; protection of rights and personal safety of the parties in the event of personal data leakage etc.

E. Yamasheva, Researcher at the ILCL Center for criminal law and criminal procedure, discussed several aspects of digitizing penitentiary system in Russia.

The digital change is one of the main vectors identified in the 2030 Penal (Penitentiary) System Development Concept. Under the Concept it is envisaged to create and develop data collection and processing systems, with AI to be used for secure decision-making (including video content analysis to forecast the behavior of convicts and penal system staff). The Federal Penitentiary Service has made a decision to digitize 380 correctional facilities, with a facial recognition system to be introduced for video monitoring at the FPS offices and facilities.

The personal identification technology is already used today at checkpoints of correctional facilities to enhance security. In the future, intelligent data analysis for processing information in the penal system will improve the safety of convicts and staff through stronger information support of facilities and offices, and better forecasting and planning of work with the accused and convicts including to stop crime.

However, in the penal system AI will require normative regulation of both operational aspects and protection of rights, liberties and legitimate interests of individuals since its uncontrolled use could be harmful in many respects, only to result in disclosure of personal data, discrimination and more severe implications. With legislative amendments and comprehensive legal support of AI referred to in the National Artificial Intelligence Development Strategy, there is also a need to improve the penal law.

6. Digital Platforms in Specific Spheres

M. Drozdova, Candidate of Sciences (Law), Associate Professor at the Saint Petersburg State Transport University, examined aspects of regulation of digital logistical platforms.

I. Bashlakov-Nikolaev, Candidate of Sciences (Law), Associate Professor at the Russian Academy of National Economy and Public Administration, discussed the aspects of anti-trust regulation of digital platforms and its possible solutions.

I. Tselovalnikova, Candidate of Sciences (Law), Associate Professor at the Russian State University of Justice, noted the peculiarities of investment platforms in the digital environment.

The workshop was followed by the expert survey *Digital Platforms in the Focus of National Law* to identify a consensus among highly skilled specialists on the most controversial and crucial regulatory issues related to digital platforms.

Survey methodology

Almost one half of 60 respondents specializing in this sphere (48.3%) had an academic degree or status, the main age groups being 36–50 years (38.3%); 26–35 (23.3%); 51–70 (20%); under 25 years (18.4%). The respondents were from the following domains: science and education (54.2%); students (both master degree and postgraduate) (18.6%); public (municipal) servants (10.2%); business (10.2%); lawyers and other legal practitioners (6.8%).

The survey was carried out in two formats: onsite and online (by completing either a paper form at the event or online Yandex Form[1]). The respondents were proposed 11 questions[2] related to regulation of digital platform and 3 questions on personal status. Those responding onsite could leave their comments (see Fig. 1—11).

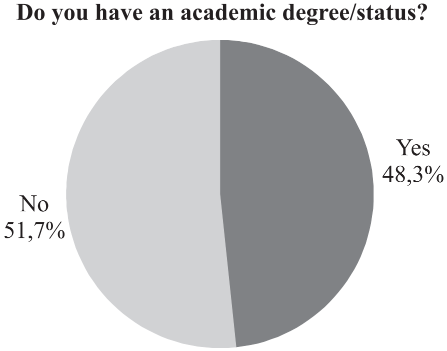


Fig. 1

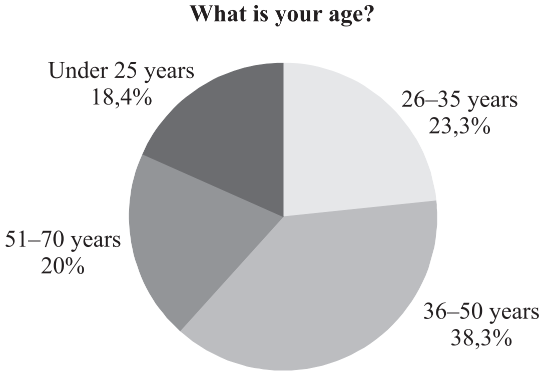


Fig. 2

[1] All questions assumed the choice of only one option. The respondents were allowed to complete the form only once and vote online until 1 May 2024.

[2] There was one question which, if answered positively, was followed by two more questions.

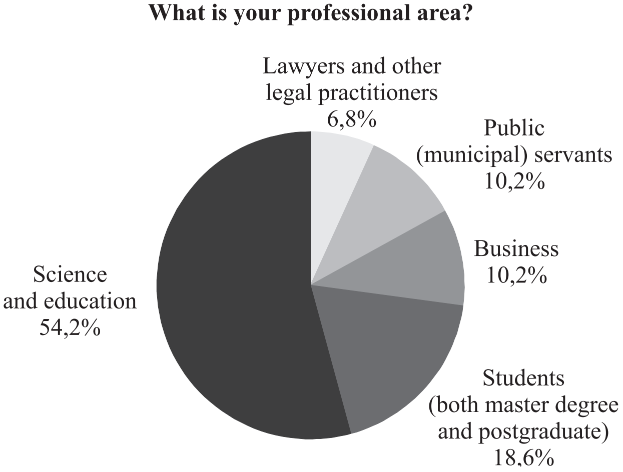


Fig. 3

Survey findings

A majority of respondents (65%) answered negatively to the first question **"Are digital platforms, information platforms, information systems and digital ecosystems identical concepts?".**

It was stated in comments to negative answers that some concepts are wider than others. Thus, a digital ecosystem may include a number of digital platforms. Moreover, it was stated in comments that these concepts differ in terms of content and purpose.

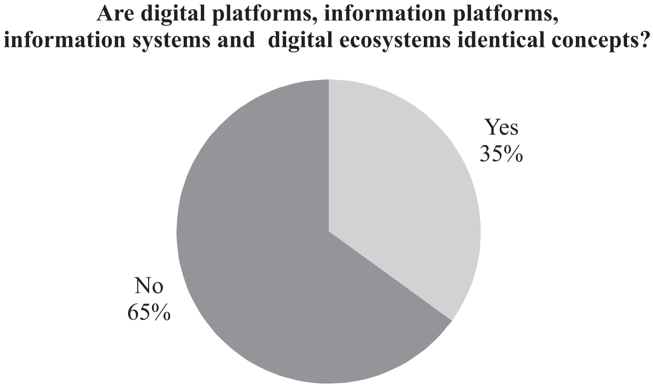


Fig. 4

The second question **"Do digital platforms need specific regulation?"** on the rationale of regulation has yielded a vast majority of positive answers

(83.3%), with the respondents noting that special regulation is required only for legal relations concerning: 1) protection of the weaker party; 2) technical regulation; 3) anti-trust issues; and 4) extent of digital platforms' use. As such, specific provisions can standardize regulation of digital platforms by way of excluding or constraining agency-specific aspects.

Do digital platforms need specific regulation?

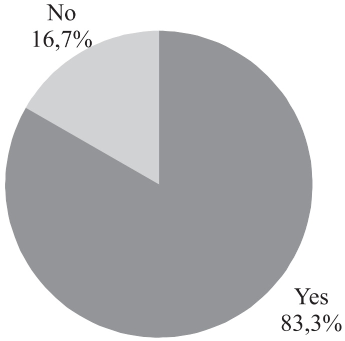


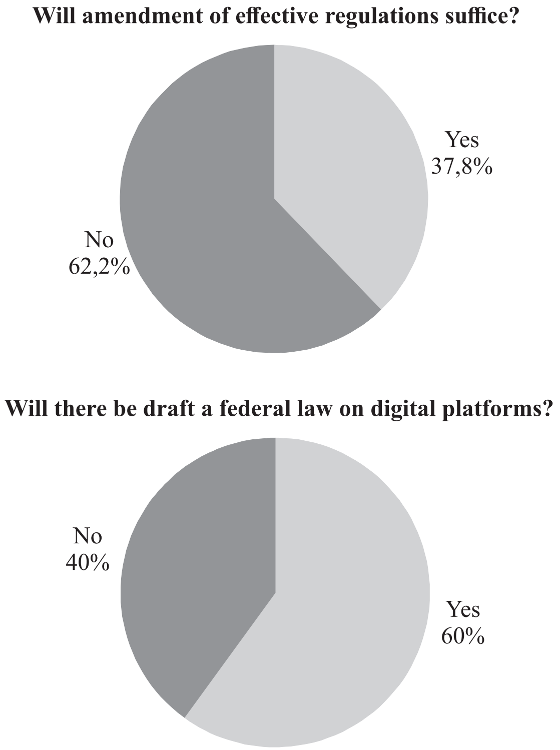
Fig. 5

The next two optional questions were designed to specify the second, with respondents asked to choose the nature of regulatory change: **"Will amendment of effective regulations suffice or is there a need to draft a federal law on digital platforms?"**.

As the survey showed, a majority of respondents were in favor of the second regulatory option. They noted in comments that the would-be federal law on digital platforms will allow to regulate these activities accurately and comprehensively but will require to amend the bulk of legal instruments for coherence with the effective regulation. More detailed regulation of specific groups of digital platforms is to be equally addressed by bylaws.

Some respondents argued that the would-be law should cover the issues of service provision and underlying dispute resolution, censorship and prohibition of access to specific digital platforms. However, it was argued in some comments that a federal law on digital platforms was premature.

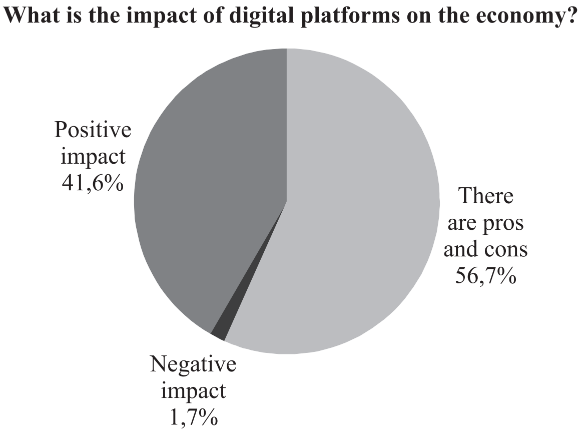
The third question was **"What is the impact of digital platforms on the economy?"**, with a majority of respondents (56.7%) believing there were both pros and cons while 41.6% noted a positive economic impact of digital platforms. Some respondents, while noting a generally positive impact on the economy, argued for more strict government control. As observed in comments, the economic upsides were: 1) easier collaboration between users; 2) broader area for collaboration; 3) stronger demand and supply of goods and services. The downsides were: 1) possibility of hacking the user infrastructure; 2) unequal user treatment, discrimination; and 3) violation of the institution of public agreement.



Will amendment of effective regulations suffice?

Will there be draft a federal law on digital platforms?

Fig. 6



What is the impact of digital platforms on the economy?

Fig. 7

The fourth question was more specific: **"What is the impact of digital platforms on public administration?"**. With a majority of respondents (55.7%) noting both pros and cons, 41.7% answered that digital platforms had a positive economic impact. The upsides of digital platform impact on public administration as observed in comments included: 1) lower maintenance and development costs of state information systems; 2) operational openness of public authorities; 3) lower bureaucracy. The downsides included data security and data leakage risks.

What is the impact of digital platforms on public administration?

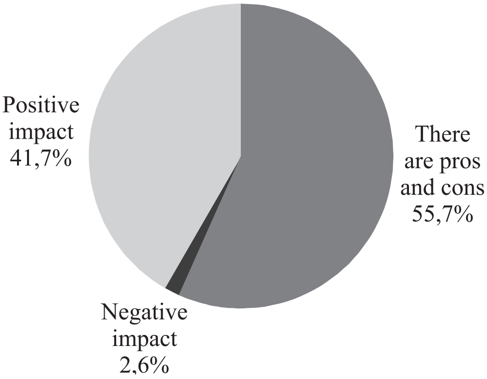


Fig. 8

The fifth question was: **"What is the impact of digital platforms on human rights?"**. While a solid majority of respondents (80%) noted pros and cons, only 16,7% believed the impact to be positive. Respondents noted in comments possible violations of human and civil rights and interests, especially since it was actually impossible to put a stop to personal data processing. Meanwhile, recommendation technologies at digital platforms based on personal data processing had a positive rather than negative impact. Adequate regulation of these technologies is therefore more preferable than banning them altogether.

An overwhelming majority of respondents (80%) answered positively to the sixth question **"Do children need more protection when using digital platforms?"**. Moreover, they noted in comments that stronger parental control and higher protection within the system were needed.

A majority of respondents (90%) answered positively to the seventh question **"Do human rights (including personal data) need more protection at digital platforms?"**.

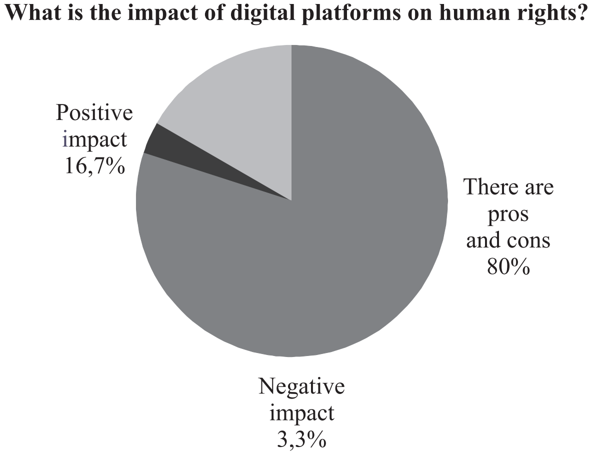


Fig. 9

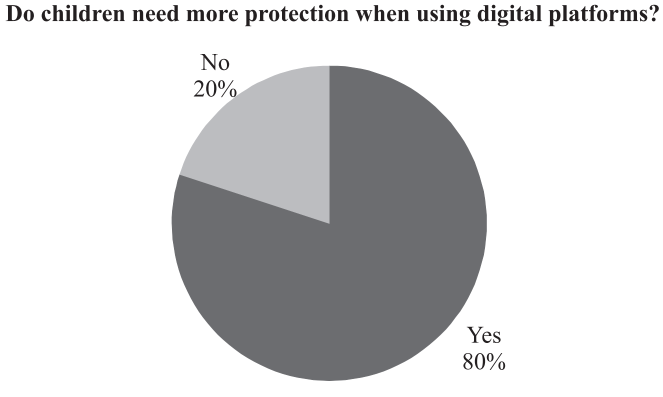


Fig. 10

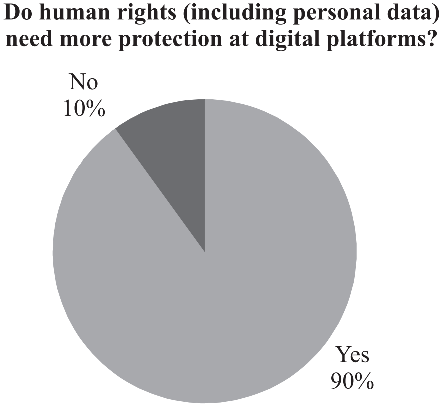


Fig. 11

A majority of respondents (60%) answered negatively to the eighth question **"Did you face any form of discrimination when using digital platforms?"**

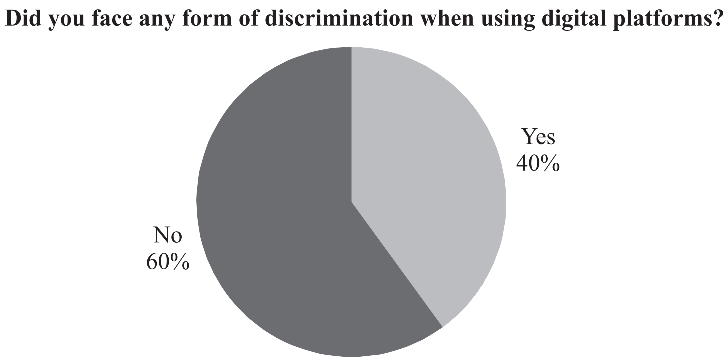


Fig. 12

An absolute majority of respondents (66,7%), however, answered positively to the next question **"Do digital platforms need to be subject to more anti-discrimination measures?"** Respondents believe that discrimination is non-transparent, implicit and shady since, for example, there is no feedback; true reasons of service denial and dynamic pricing mechanisms are unknown etc.

In their comments, respondents specified the following additional measures to amend the law: 1) a special authority to consider digital platform related disputes; 2) specifying requirements to recommendation services including to disallow the use of specific personal data; and 3) allowing to collect sensitive personal data only if consented by the person in question.

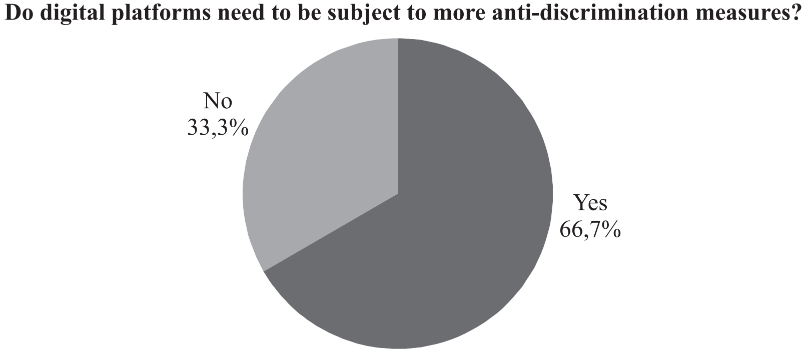


Fig. 13

The tenth question was: **“Who should be legally liable for harm resulting from operation of digital platforms?”**. It has raised the worst controversy as offline respondents[3] were allowed to choose only one option in answering other questions while in this case several reply options were possible[4]. While the answers split into two large groups without sizeable difference between them, a relative majority of experts (48.1%) believe that only the person providing services on the operator’s behalf should be legally liable.

Moreover, the comments did not reveal any common approach to the grounds for legal liability. Some believe that legal liability always result from the caused harm; others, only depending on the degree of proven guilt; still others, that the economic sector and contractual relations with the contractor also had a role to play.

In addition, respondents noted in their comments that the developer can only be liable by way of recourse under the contract with the digital platform operator.

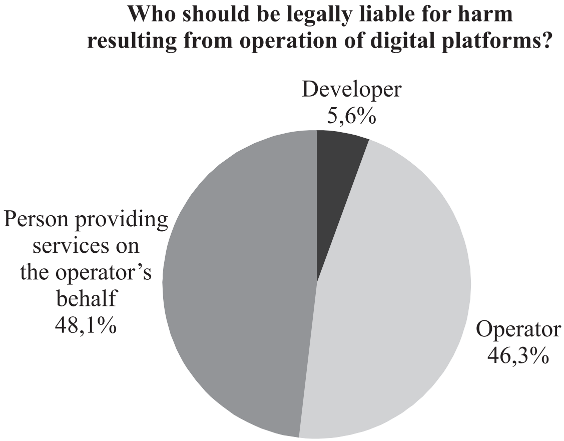


Fig. 14

A majority of respondents (66.7%) answered positively to the last (eleventh) question **“Do labor law provisions need to be amended under the impact of digital platform operations?”**

[3] Those voting online could not give more than one answer for technical reasons.

[4] The respondents who gave two answers (approximately 10% of all those surveyed) were not counted in the total sample. A vast majority of them would choose two persons: person providing services on the operator’s behalf and the operator (owner) himself.

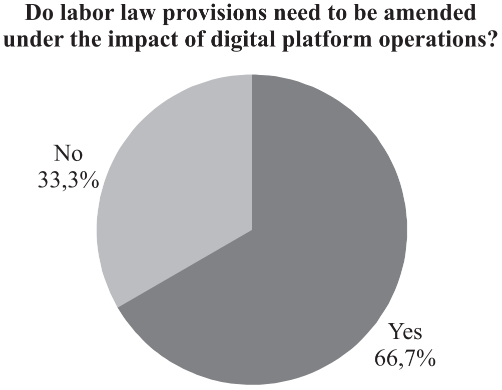


Fig. 15

References

1. Fedotov M.A., Naumov V.B. et al. (2024) The Right to Refuse Digital Technologies: Outcomes of an Expert Survey. *Trudy po intellektualnoy sobstvennosti*= Works on Intellectual Property, vol. 48, no. 1, pp. 8–28 (in Russ.)

2. Korshunova T. Yu. et al. (2024) Judicial Practices and Development of Labor and Social Security Law: a guide. Moscow: Kontrakt, 248 p. (in Russ.)

Information about the authors:

L.K. Tereschenko — Doctor of Sciences (Law), Chief Researcher, Honored Jurist of Russia.
O.E. Starodubova — Researcher.
N.A. Nazarov –Junior Researcher, Postgraduate Student.

The article was submitted to editorial office 23.05.2024; approved after reviewing 23.06.2024; accepted for publication 28.06.2024.

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.