Legal Issues in the Digital Age. 2024. Vol. 5. No. 4.
Вопросы права в цифровую эпоху. 2024. Т. 5. № 4.

Research article
JEL*:* K00
УДК*:* 34
DOI:10.17323/2713-2749.2024.4.73.91

Deepfakes: Search for a Model of Legal Regulation

Vladislav Olegovich Demkin

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

vodemkin@hse.ru

ORCID: 0000-0002-1079-425X, Google Scholar: OfF2SrUAAAAJ, SPIN-code RINT: 1755-2053

Abstract

Modern studies of law, political science and other humanities reveal a major public concern about deepfake technologies, with legal regulation thereof only emerging. This paper looks into the main models whereby such technologies are regulated in Russia, China, European Union, United States and United Kingdom. Effective regulation of technologies should have as its main goal the protection of personal rights through methods of private and public law while striking a balance between relevant interests of other subjects to social relations. The study employs a variety of methods: comparative method (to analyze how deepfake technologies are regulated under various legal systems); method of rising from the abstract to the concrete (to move from regulation of AI to specific ways of regulating deepfake technologies); and the formal dogmatic method (to analyze legal provisions and their place in the regulation of deepfake technologies). The study provides a list of parties to AI-related social relations whose interests should be accounted for in developing the underlying regulation. The author points out certain fundamental questions to be resolved for legal regulation of deepfake technologies to emerge in Russia, and concludes by proposing answers to the said questions and identifying the vector of regulatory development.

Keywords
deepfake; deceptive information; human rights; personal dignity and honor; artificial intelligence; digital personality.

For citation: Demkin V.O. (2024) Deepfakes: Search for a Model of Legal Regulation. *Legal Issues in the Digital Age,* vol. 5, no. 4, pp. 73–91. DOI:10.17323/2713-2749.2024.4.73.91

Background

In their current state, machine learning technologies allow to design computer models (replicas) of real people using their biological features. Thus, the widely used AI-based deepfake technologies allow to make synthetic media representing persons under certain aspects including obscene, for instance, in pornography [Pfefferkorn R., 2020: 265][1]. The problem of correct use of novel technologies to prevent deceptive information (fakes) has gained considerable public interest. In particular, according to *Google Trends*, search queries containing the word “deepfake” started to appear in February 2018, only to proliferate afterwards[2]. Society is wary of wide dissemination of technologies that allow to create deceptive information and content involving real persons but having nothing to do with the reality.

The term “deepfake” was coined from two English words: “deep learning” (to imply the use of neural network) and “fake”.

There are different criteria to classify deepfakes. First, they can be targeted or not depending on the proposed recipient of information fabricated through their use [Roberts T., 2023: 2]. By the nature of underlying content, deepfakes split into: 1) commercial (used for business development purposes); 2) original and creative (for example, in motion pictures); 3) vindictive; and 4) political [Meskys E., Kalpokiene J., Jurcys P. et al., 2020: 25].

While the term “deepfake” is technical, the national law in countries responds differently to the questions resulting from growing public interest to the problem of technological usage.

[1] See also: Horrifying new AI app swaps women into porn videos with a click. MIT Technology Review. Available at: https://www.technologyreview.com/2021/09/13/1035449/ai-deepfake-app-face-swaps-women-into-porn/ (accessed: 27.07.2024)

[2] Google search statistics. Available at: https://trends.google.com (accessed: 27.07.2024)

How the legal systems of Russia, China, EU, United States, United Kingdom regulate deepfake technologies is discussed below.

1. Legal Regulation of Deepfake Technologies

1.1 Searching for Ways to Regulate Deepfake Technologies

Russia is currently at an early stage of conceptualizing regulation of artificial intelligence including deepfake technologies. These efforts largely involve proposals to add the right to "voice" or "personality" to the list of moral rights; regulate the right to voice as exclusive right to intellectual assets, and speech synthesis — along the lines of licensing agreements; treat human voices and images as biometric personal data subject to relevant regulation when used in generative neural networks; toughen the liability for specific offences involving generative neural networks; and designate deepfake content as such.

Deepfake technology is defined by both judicial practice and legal doctrine. In Supreme Court Plenum Resolution No. 17 "On specific questions raised by courts in handling administrative offences aiming to undermine the procedure for information support of elections and referendums" of 25 June 2024, deepfakes are understood as misleading and misrepresenting images, audio and audiovisual information including created through the use of computer technologies[3].

The legal doctrine defines deepfakes as photographs, video or audio created by artificial intelligence to replicate the reality (normally by stacking the existing images and videos over source images or video clips) [Kalyatin V., 2022: 87]; [Pfefferkorn R., 2020: 248]; as AI-based technologies to produce or edit video or pictorial content in order to show something that never happened [Young N., 2019: 8]. This definition generally matches the technical one.

In October 2023, the Council for Digital Economic Development under the Federation Council of the Federal Assembly of Russia held an AI desk meeting on legal guarantees for natural persons when their speech is synthesized (generated by computer) which is also a variety of deepfake technology[4]. Following the discussions, the Federation Council decided

[3] SPS ConsultantPlus.

[4] A. Sheikin chaired AI desk meeting of the Council for Digital Economic Development under the FC. Available at: URL: http://council.gov.ru/events/main_themes/148788/ (accessed: 27.07.2024)

to draft amendments to the Civil Code of Russia for human voice to be treated as intangible goods like human image. Other discussions focused on the need to designate synthetic speech created through the use of deepfake technologies; and liability for making public (posted in the web) speech recordings without seeking consent of the person concerned[5].

In furtherance of the idea to allow or prohibit voice synthesizing, the National Federation of Music Industry (NFMI) proposed broader protection of “digital image”. According to the NFMI General Director N.A. Danilov, such provision would more efficiently protect performers from commercialization of personality through the use of deepfakes[6].

In absence of regulation, legal gaps are to be filled by court practice. Supreme Court Plenum Resolution No. 17 mentioned above provides using deepfakes in pre-election campaigning is a violation under Article 5.12 of the Code of Administrative Offences “Production, dissemination or publication of campaign materials in defiance of legal provisions on elections and referendums” (in particular, of paragraph 1.1, Article 56, Federal Law “On the Principal Guarantees of Russian Citizens’ Right to Vote and Take Part in Referendums”)[7].

The legal doctrine is also devising ways to regulate deepfake technologies.

For example, V.O. Kalyatin considers deepfake content from the perspective of intellectual property law by exploring the questions of attribution, use of intellectual property assets, pictures and images of natural persons, confidential information, as well as the assignment of exclusive right to source materials and resulting deepfakes [Kalyatin V., 2023: 17]. M.B. Dobrobaba argues for the development of tools that allow to identify and address deepfakes; for designation of deepfake content in social media; for tougher liability for violating third party rights through the use of deepfake technologies [Dobrobaba M., 2022: 117]. A.V. Minbaleev proposes to adopt basic AI federal law to be complemented by specific regulations to address specific processes or technologies such as generative neural networks including deepfake technologies [Minbaleev A., 2023: 15].

[5] Federation Council proposed to designate synthesized voice. Available at: URL: https://tass.ru/obschestvo/19816417 (accessed: 03.08.2024)

[6] Coming short of one voice. Available at: URL: https://www.kommersant.ru/doc/6805009 (accessed: 01.08.2024)

[7] SPS Consultant Plus.

In drafting specific provisions to regulate deepfake technologies, one should take into account the need to strike a balance between the interests of those governed by the underlying regulation: 1) persons whose image, voice, "personality" (broadly understood as dynamic set of physiological, biological, emotional features) are used in the deepfake content; 2) those creating artificial video, audio clips and images; 3) website owners/platform administrators providing tools and technologies to produce deepfakes; 4) government represented by public authorities for protection of rights and interests of natural and legal persons. Striking a balance of interests in shaping and improving legal regulation in the area under discussion also attracts scholarly notice. Thus, there is a perceived need, on the one hand, to fix up a system of rules applicable to deepfake technologies and envisaging liability for violations while, on the other hand, avoiding barriers to technological progress as a whole or prohibiting deepfake technologies altogether. There is also a need to keep the balance between values promoted by the legal system and new technological boundaries [Vinogradov V., Kuznetsova D., 2024: 239].

Internationally, the regulation of deepfake technologies is at various stages of development.

Thus, the Chinese regulatory model can be called administrative as it purports to impose additional requirements on owners of the tools for production of deepfakes. Believing that deepfakes are fraught with major social risks, the Chinese government wants not only to designate artificially created content but also to add new elements to the list of criminal offences: dissemination of non-designated deepfake content as news.

As regards regulation of deepfake technologies, China was the first to establish strict and detailed rules for production and dissemination of deepfake content through the use of neural networks, with the Regulation on Deep Synthesis Information Web Services adopted on 25 November 2022[8] imposing obligations on owners of deepfake creative tools.

The EU regulation of deepfake technologies is AI-focused, with the Artificial Intelligence Act ("Regulation") approved in 2024[9] defining de-

[8] Regulation for Deep Learning Management of Information Web Services (in Chinese). Available at: http://www.cac.gov.cn/2022-12/11/c_1672221949354811.htm (accessed: 27.07.2024)

[9] Regulation of the European Parliament and Council of 13 June 2024 Laying Down Harmonised Rules On Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU)

velopment methods of deepfake technologies (including by establishing regulatory sandboxes); creating and defining the powers of a special AI supervisory authority to be set up, and applicable penalties. The adoption of special instrument will allow to fine tune the regulation of modern technologies while accounting for the needs of all parties involved — natural persons as owners of "special" rights in the digital world; businesses as users and beneficiaries of technologies; and the governments as regulators of the underlying relations.

The United States and the United Kingdom follow a different approach: there is currently no specific regulation of AI and deepfake technologies, with deepfake related legal problems addressed by adding up to the existing elements of crime. In legal literature it is pointed out that the American method of regulation is particular in its unwillingness to prohibit AI-generated content due to priority of individual rights including freedom of speech and expression. The U.S. sources argue that society is wary of deepfake regulations believed to infringe on the freedom of speech [Joost L., 2023: 312][10]. Other American authors come to similar conclusions. In particular, it is noted that deepfakes can be regarded as a form of self-expression protectable by the First Amendment to the U.S. Constitution. Under another argument, deepfake content occupies an intermediate position, with some instances of use to be protected by the freedom of expression and others criminalized [Blitz M., 2020: 300].

Some laws in the United States require from different organizations and public agencies to make AI-related reports and propose response. In 2019, the U.S. adopted a number of laws regulating AI-related operations of public authorities and certain state-owned firms. Such instruments oblige them to monitor the progress of deepfake technologies worldwide, assess the underlying risks, and also promote public-private partnership to conduct relevant research and counter deceptive information[11]. These instruments

2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act). Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202401689 (accessed: 22.07.2024)

[10] See for example: As Deepfakes Flourish, Countries Struggle with Response. Available at: https://www.nytimes.com/2023/01/22/business/media/deepfake-regulation-difficulty.html?action=click&module=RelatedLinks&pgtype=Article (accessed: 27.07.2024)

[11] First Federal Legislation on Deepfakes Signed into Law. Available at: https://www.wilmerhale.com/insights/client-alerts/20191223-first-federal-legislation-on-deepfakes-signed-into-law (accessed: 05.08.2024)

have no direct impact on mechanisms for protection of personal rights or regulation of deepfake technologies. In addition, pursuant to the Identifying Outputs of Generative Adversarial Networks Act ("IOGAN") of 2020, the National Research Foundation and the National Institute of Standards and Technology are required to support and promote research of the methods of generative adversarial networks[12].

There are ongoing discussions in the United States of the draft NO FAKES ACT (Nurture Originals, Foster Art, and Keep Entertainment Safe Act)[13] designed to introduce more detailed regulation of digital replicas. The draft submitted to the U.S. Congress by both Republican and Democratic senators defines a digital replica as a newly created, computer-generated, electronic representation of the image, voice, or visual likeness of an individual that: (A) is nearly indistinguishable from the actual image, voice, or visual likeness of that individual; and B) is fixed in a sound recording or audiovisual work in which that individual did not actually perform or appear. The consent to produce and use such digital replica is assumed to be a digital replication right that is material, heritable and transferrable or assignable in full or in part (along the lines of exclusive right to intellectual property). In order to dispose of this right, one has to retain a professional attorney (lawyer, solicitor, trade union). Thanks to the federal act for protection of digital identity, the right to publicity that provides similar protection to human image but is not recognized in some states can become universally enforceable.

Deepfake technologies are also regulated at the state level. For instance, the State of California has a number of laws governing deepfake content: Assembly Bill No. 602[14] banning erotic deepfake content without approval of the person represented who is free to claim damages from the content creator, and Assembly Bill No. 730[15] banning the dissemination of election-related deepfake content sixty days prior to the voting day ("a person ...,

[12] Identifying Outputs of Generative Adversarial Networks Act. Available at: https://www.congress.gov/bill/116th-congress/senate-bill/2904/text (accessed: 28.07.2024)

[13] Nurture Originals, Foster Art, and Keep Entertainment Safe Act. Available at: https://www.ilga.gov/legislation/BillStatus.asp?DocNum=5594&GAID=17&DocTypeID=HB&LegId=153975&SessionID=112&GA=103 (accessed: 28.07.2024)

[14] Assembly Bill № 602. Available at: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=20212 0220AB602 (accessed: 28.07.2024)

[15] Assembly Bill № 730. Available at: URL: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=20192 0200AB730 (accessed: 28.07.2024)

committee, or other entity shall not, within 60 days of an election ... distribute, with actual malice, materially deceptive audio or visual media of the candidate with the intent to injure the candidate's reputation or to deceive a voter into voting for or against the candidate").

The law of Texas criminalizes the creation of misleading deepfakes seeking to impact the voting outcome if they are published 30 days prior to the voting day[16]. The State of Tennessee is peculiar for civil law regulation of deepfake technologies, with the person's name, voice and likeness protected as personal rights. Thus, any action to use these assets (including to produce deepfakes) without seeking the person's (or owner's) consent are deemed illegal and result in liability[17].

1.2. Obligations of Deepfake Technology Owners

The Russian law currently imposes no specific requirements on the owners of deepfake technologies or platforms. Operations of deep learning generative models are largely based on the terms of service (to be accepted by users at registration or simply during content production). Thus, those using Vassily Kandinsky creations to produce pictures and videos under its terms of service cannot use intellectual outcomes, identifications, third party personal data and information that constitutes any secret whatsoever. That is, users of this technology are assumed to seek third party consent in order to use the underlying items in the neural network. Also, there is a prohibition to use video and pictorial outputs that violate provisions of the Russian law[18]. Yandex neural network technologies have similar terms of service[19].

Legal systems containing (and discussing the adoption of) specific provisions to regulate AI technologies — in particular, deepfakes — will impose extra obligations on service owners.

[16] An act relating to the creation of a criminal offense for fabricating a deceptive video with intent to influence the outcome of an election. Available at: https://capitol.texas.gov/tlodocs/86R/billtext/pdf/SB00751F.pdf#navpanes=0 (accessed: 05.08.2024)

[17] An act to amend Tennessee Code Annotated, Title 39, Chapter 14, Part 1 and Title 47, relative to the protection of personal rights. Available at: https://www.capitol.tn.gov/Bills/113/Bill/SB2096.pdf (accessed: 05.08.2024)

[18] Sberbank's Kandinsky Terms of Service and Acceptable Use Policy of AI Services. Available at: https://www.sberbank.com/common/img/uploaded/files/promo/kandinskiy-terms/kandinskiy-terms-of-use.pdf (accessed: 17.07.2024)

[19] Yandex Foundation Models Terms of Service. Available at: URL: https://yandex.ru/legal/cloud_terms_yandex_foundation_models/ (accessed: 17.07.2024)

The Chinese regulatory model is focused on the engagement with owners of a proposed service. Under the 2022 Regulation on Deep Synthesis Information Web Services, they are required to assume three obligations of active influence: on users (through authentication and blocking); on inputs (by checking the adequacy of rights clearance); and on outputs (ensuring transparency through designation of deepfakes). This is claimed to be the established tradition of communication network governance in China where the government increasingly relies on technological companies for observance of web regulation standards and on relevant corporate initiatives [Hine E., Floridi L., 2022: 608].

1.3. User Authentication

Under the 2022 Regulation on Deep Synthesis Information Web Services, all administrators of deepfake technologies in China are required to have authentication using mobile phone numbers or specific public identifiers. Technologies of this kind are treated as web services governed by the Cybersecurity Law of China[20] (applicable to communication network owners or managers, network service providers or other persons of similar status). This law requires to deny service to those failing to provide personal data to the service owner. The web operations history of users can be made available to law enforcement bodies. This allows to easily identify and penalize the author of a particular deepfake, thus ruling out anonymous action in the Internet. So, the Chinese regulatory model assumes that deepfakes can be used only on the basis of authentication of those who make them.

Deepfake service owners are equally required to monitor the legitimate use of people's personal data and other sensitive information, as well as to censor the clips containing black-listed words. Such regulation applies to deepfakes irrespective of the subject and purpose. For service owners, the obligation to ensure legitimate use of content (via platforms, websites, mobile apps and databases) often means a need to introduce certain amendments to the underlying terms of service (to prevent the use of illegally produced content) and to deny access to materials contested by the subject on the basis of minimally required evidence of ownership. Thus, compliance with the same rules can be also expected in case of Chinese deepfake services. The mandatory user authentication requirement will simplify the

[20] Cybersecurity Law of China. Available at: https://www.cac.gov.cn/2016-11/07/c_1119867116.htm (accessed: 07.08.2024)

identification of wrongdoers while resulting in excessive processing of sensitive personal data by deepfake generative services. With such regulation focused primarily on web service owners, this model can be called administrative. Overall, it is successfully embedded into China's web regulation system characterized by a fair measure of state control, propensity for zero anonymity and attempts to put the interests of socialist society above those of private users.

In the EU and the United States, it is out of question to adopt provisions on mandatory user authentication. The Artificial Intelligence Act does not require from deepfake service owners to ensure authentication on the basis of personal identifiers. In our view, this is because such processing of sensitive identifiers would be contrary to the EU's regulatory policies of personal data protection. Such approach does not contradict the Act's principal objective of protecting personal rights since it ensures a comprehensive approach including in stressing the importance of correct personal data processing.

1.4. Designating Deepfake Content

Many legal systems worldwide will require from deepfake makers or enablers to designate the underlying material as artificial. This is believed to disclose to third parties significant information on its nature and to warn that the content disseminated in this manner is not credible.

Deepfake technology administrators are required to designate AI-generated content. The so-called transparency requirement is designed to warn the public that the content is artificial. The Chinese government criminalized publication of news created through the use of artificial intelligence and not designated as such[21] as early as in 2020.

The EU's Artificial Intelligence Act of 2024 requires from technology owners to designate content created through the use of AI.

In the United States and some other countries it is proposed to require from deepfake makers to designate their output accordingly. Thus, the draft DEEPFAKES Accountability Act of 2023 describes a procedure for desig-

[21] China seeks to root out fake news and deepfakes with new online content rules. Available at: https://www.reuters.com/article/us-china-technology/china-seeks-to-root-out-fake-news-and-deepfakes-with-new-online-content-rules-idUSKBN1Y30VU (accessed: 27.07.2024)

nating each type of content as created through the use of AI[22]. Under the U.S. law, in accordance with the draft COPIED ACT[23] each owner of a tool (website, platform or application) for AI-generated content (including deepfakes) must enable users to designate output to signal its artificial origin. The output so designated will not be usable for business purposes and neural network training. In the event of abuse, victims are free to claim damages and termination of the content's illegal use. Thus, a decision to designate will be taken individually by each user.

The EU and U.S. law is softer on requirements to technological companies and artificial content makers as not every image should be designated (the U.S. draft laws make this altogether voluntary); not every media should be prohibited for publication by virtue of the freedom of speech; and sensitive personal data of content makers is not to be processed in all cases.

From the perspective of barriers to technological progress, the requirement to designate all content being created is not reasonable. Deepfake outputs can be used in different formats, forms and types and for different purposes including private. For example, they are often used as robot secretaries at banks and health centers where it is assumed that callers do not deal with a real person. Meanwhile, designating such robot secretaries as artificial (for example, by a conversation starter) will undermine their commercial value and user attractiveness.

Moreover, it is not feasible to require to designate absolutely all AI-generated content. In particular, voice assistants, robotic secretaries at businesses (such as banks or health centers) should be able to quickly and precisely answer user queries in line with business objectives. In fact, a reasonable user can expect that his interlocutor is actually a software. Apparently, there should be exceptions from the general rule that requires to designate AI-created content. Anyway, such decision and discussions of a possible draft law should be based on the engagement with industry representatives, that is, technological companies already in possession of similar AI technologies as such designation will contradict the requirement that appearance should be attractive to users.

[22] DEEPFAKES Accountability Act 2023 (H.R. 5586). Available at: https://www.congress.gov/bill/118th-congress/house-bill/5586/all-actions (accessed: 08.08.2024)

[23] The Content Origin Protection and Integrity from Edited and Deepfaked Media Act. Available at: https://www.commerce.senate.gov/services/files/3012CB20-193B-4FC6-8476-DDE421F3DB7A (accessed: 28.07.2024)

A special consultative body on artificial intelligence including deepfakes may be helpful in mapping content to be designated as well as in addressing other issues. The EU's Artificial Intelligence Act of 2024 requires to set up the European Artificial Intelligence Board ("EAIB") to monitor the use of AI technologies. The EAIB will be authorized to issue opinions, recommendations and other guidelines, interpret legal provisions, develop best practices, harmonize AI-related technical standards, collect relevant data from member states on implementation of the Act and performance of regulatory sandboxes.

A special body may be also created in Russia[24]. While the new EAIB will act as such in the EU, the same functions can be assumed in Russia by a special division of the Federal Service for Supervision of Communications, Information Technology and Mass Media (Roskomnadzor).

1.5. Criminal, Administrative and Civil Liability for Deepfakes

It has a sense to believe there is currently no basis to penalize deepfake technologies either under administrative or criminal law as they are just tools to commit such offences. Meanwhile, it should be remembered that in some cases their use will make offences more harmful to society: artificial pictorial and audio content is highly delusive and gives a semblance of reality; it is web spreadable, only to become viral quickly and easily; it may be hard to refute; and attributing or identifying the source of deepfakes is problematic. Thus, it makes perfect sense to consider the use of AI technologies as a circumstance aggravating administrative or criminal liability, and conventionally also civil liability, for example, by suggesting a higher amount of compensation for violation of exclusive right to intellectual outputs and means of identification.

As noted in Supreme Court Plenum Resolution No. 17 mentioned above, the use of deepfake content is a way to commit the offence described in Article 5.12 of the Code of Administrative Offences: "production, dissemination or publication of campaign materials in defiance of legal provi-

[24] Artyom Sheikin, Deputy Chairman of the Digital Economic Development Council under Federation Council announced possible establishment of a special Roskomnadzor division to monitor the creation and use of AI technologies including deepfakes, address operational issues of neural networks, handle complaints, take decisions to apply sanctions, block website access to the Internet. Available at: URL: https://senatinform.ru/news/senator_sheykin_v_rkn_mozhet_poyavitsya_otdel_po_kontrolyu_za_ispolzovaniem_ii/ (accessed: 27.07.2024)

sions on elections and referendums" (in particular, of paragraph 1.1, Article 56, Federal Law "On the Principal Guarantees of Russian Citizens' Right to Vote and Take Part in Referendums").

The following possible offences involving deepfake technologies are fraught with considerably higher social risks: production/dissemination of extremist content; slander; violation of privacy; violation of copyright and related rights; fraud, coercion to perform or abandon a transaction; public call for terrorist action; public defense or advocacy of terrorism; illegal production and sale of pornographic content or items.

Circumstances aggravating administrative liability for committed offence are prerequisites that the penalty is fair and individual, something that primarily serves to achieve the purpose of correcting the behavior of those convicted to administrative liability and preventing further offence [Sundurov F., Talan M., 2015: 175]. They allow the court to justify a penalty approximating the maximum under the Code of Administrative Offences or Criminal Code. It is also worth noting that legal literature identifies among aggravating circumstances special ones [Sundurov F., Tarhanov I., 2016: 204] as constituting qualified factors of a specific offence — such as murder motivated by blood revenge[25]. Unlike "general" aggravating circumstances, they constitute specific offences not intended for assessing social danger of others.

So, the fact of using deepfake technologies to commit an offence is perfectly qualified to become an aggravating or special aggravating circumstance due to the wrongdoer's awareness of higher social danger.

The use of such services to commit other offences can be currently taken into account in Russia as an aggravating factor or circumstance. As was explicitly noted in the aforementioned Supreme Court Plenum Resolution, the use of deepfakes in the context of political campaigning constitutes an offence under Article 5.12 of the Code of Administrative Offences, with liability equally applicable to those who made and commissioned deepfake content. The only possible sanction is administrative fine of five to twenty thousand rubles for private individuals; thirty to fifty thousand rubles for officials; and one hundred to five hundred thousand for legal entities[26].

[25] Criminal Code of Russia, Law No. 63-FZ of 13.06.1996 // SPS ConsultantPlus.

[26] Code of Administrative Offences of Russa, Law No. 195-FZ of 30.12.2001 // SPS ConsultantPlus.

The Criminal Code of China qualifies the offences likely to be invoked in the production of other people's images using deepfake technologies. Thus, Article 235 envisages sanctions for production, dissemination and even possession of obscene images, audio recordings and texts while Articles 310, 313 prohibit slander including fraudulent dissemination of deceptive rumors. The use of the Internet to commit these crimes is an aggravating circumstance[27].

In addition, victims of deepfake content can expect that their defamation claim will be satisfied. Under the Civil Code of China, anyone offending honor, dignity and reputation of others should compensate the resulting damages and stop the violation. It is noted that the burden of proof in such cases is to be assumed by the plaintiff also supposed to justify the amount of damages, something that is not quite easy [Tianren L., Yue D., 2023].

Under the law of the United Kingdom and some American states, the use of AI technologies to create exclusively pornographic deepfakes is treated only as a way of committing offence already covered by criminal law[28]. Thus, in the State of Virginia, Articles 18.2-386.1 and 18.2-386.2 prohibit to create and disseminate other people's images without their consent regardless of the technology being used. These are class 1 offences punishable by a fine of up to USD 2,500 or prison sentence of up to 12 months. The same offences committed against minors become criminal charges that envisage more severe punishment[29].

The Online Safety Bill passed in the United Kingdom in 2023 is designed to regulate web activities of natural and legal persons including digital offences.

Germany's Bundesrat published a draft law to introduce criminal liability of up to two years in prison for digital fraud (deepfake) against personal rights (including those of deceased individuals). As an exception from elements of crime, a person will not be liable if the deepfake was made in pursuit of "prevailing" legitimate interests in arts, sciences, education, cov-

[27] Criminal Code of China. Approved 14.03.1997 at the 5th session of the National People's Congress. Available at: https://law.moj.gov.tw/ENG/LawClass/LawParaDeatil.aspx?pcode=C0000001&bp=44(accessed: 27.07.2024)

[28] See, for example: Deepfakes and American Law. Available at: https://www.davispoliticalreview.com/article/deepfakes-and-american-law (accessed: 30.07.2024)

[29] Code of Virginia. Available at: https://law.lis.virginia.gov/vacode/ (accessed: 09.08.2024)

erage of news or history etc. Naturally, no liability will arise where no personal right was infringed — for example, where the production of deepfake was consented by the person in question[30].

Thus, it is revealing that the Chinese government regards the use of deepfakes as fraught with major social risks and therefore does not only require to designate any artificial content but also to add new elements of crime to those covered by criminal law — dissemination of non-designated deepfake content as news.

The EU's Artificial Intelligence Act provides for exterritorial effect with the content to be designated, a special consultative body (EAIB) established, and relevant technologies consistently developed to comply with legal provisions.

Unlike the EU, the United States and the United Kingdom have not introduced special regulation of AI technologies, with legal problems related to deepfakes being addressed by adding such novel element of crime to those already existing. It is noted in literature that the American way of regulation is particular in its unwillingness to prohibit AI-generated content due to the need to observe personal rights protected by the First Amendment to the U.S. Constitution, that is, freedom of speech and freedom of expression.

It is noteworthy that the problem of deepfake-related criminality is increasingly observed around the globe, with the use of deepfake often recognized as independent element of crime or covered by special regulation.

2. Proposals on Regulating Deepfake Technologies in Russia

Special provisions regulating deepfake technologies should be developed with a view of striking a balance between the interests of all parties to the relevant social relations. The regulatory practice in Russia should be apparently "soft". The areas of social relations where it is prohibited to make and disseminate deepfakes should be limited to those vital for society and private individuals, such as those affecting the most sensitive sides of life.

The answer to the question on the regulatory vector of deepfake technologies in Russia should proceed, at the very minimum, from the following:

[30] Entwurf eines Gesetzes zum strafrechtlichen Schutz von Persönlichkeitsrechten vor Deepfakes. Available at: https://www.bundesrat.de/SharedDocs/drucksachen/2024/0201-0300/222-24(B).html (accessed: 10.08.2024)

Possibility to process personal data of individuals captured by deepfake content using artificial intelligence (including generative neural networks);

Considering the use of deepfake technologies as an aggravating circumstance in certain types of offence;

Identifying types and cases of using deepfake content (of certain type) where it should be designated as artificially created;

Identifying the need to set up a public authority to decide, advise, recommend, collect best practices regarding the use of AI-created content. Identifying the need to establish an entity among major technological companies as part of self-regulation of private businesses to identify common policies and technological development vector for AI-created content.

Thus, the question of whether someone's image and voice amount to biometric personal data (therefore required to be processed as biometric) appears to be among the most debatable at the intersection of personal data and deepfake content. Such data can be processed only if consented by the personal data subject in writing and only via the Unified Biometric System, with other restrictions, terms and conditions equally applicable.

In this case, we believe the answer to the question to be negative because of the constitutive feature of biometric personal data described in Article 11 of the Federal Law "On Personal Data": the operator should use such data to identify the personal data subject. If the operator is understood in this case as the operator (owner) of the technology for production of deepfake content, the use of someone's image and voice for such identification is not presumable. The same holds true for those who use this infrastructure to create content since they will often create faked images and audio to represent others in a certain light. So, they already know the personality in question while the information system for processing user supplied data is not always able to compare someone's biological characteristics and personality, that is, to identify a person.

Thus, in case of deepfake content, personal data may be and often is processed but such processing is not presumable but depends on actually proved circumstances of specific case. This means that personal data processing must be triggered by the presence of at least one of the legitimate grounds established by Article 6 of Federal Law No. 152-FZ "On Personal Data" of July 27 2006 for the category of "normal" data. These include the data subject's consent and performance of the contract with the data subject. However, the actual duty to provide legitimate basis just as the risks of non-compliance are assumed by those who make deepfakes. In particular, this is reflected in the terms of service of various platforms that enable the production of deepfakes.

Conclusion

Viewed from the legal perspective, the problem of using deepfake technologies has numerous aspects since it reflects provisions of different branches of both private and public law. While some countries are proposing their way to regulate artificial intelligence (including deepfakes), others only start discussing a possible course of action.

The approaches discussed in this paper are largely focused on those who own deepfake (and other AI) technologies rather than on victims of deceptive information or those accused of propagating it. Their rights, duties and liabilities are deemed duly regulated by the existing provisions of criminal and civil law based on the established practice. Depending on circumstances, they cover slander, fraud, offence to personal dignity and honor, and sometimes the dissemination of deceptive socially important information as news.

The Russian legal system still does not have specific regulation of AI.

The author has identified a number of questions to be addressed in approving provisions (if any) to regulate deepfake content, and proposed answers including:

Possibility to process personal data of individuals captured by deepfake content using artificial intelligence (including generative neural networks): such data are not biometric personal data under the general rule and can be processed on the “general basis”;

Considering use of deepfake technologies as an aggravating circumstance in certain types of offence: such legal novelty is admissible and even desirable since deepfake content considerably aggravates the social danger of offence;

Identifying types and cases of using deepfake content (of certain type) where it should be designated as artificially created: it is proposed to establish an exhaustive list (if any). However, the duty to designate will not slow down the progress of these technologies or positively affect the prevention of deepfake-related crime;

Identifying the need to set up a public authority to decide, advise, recommend, collect best practices regarding the use of content created by artificial intelligence. Identifying the need to create an entity among major technological companies as part of self-regulation of private businesses to identify common policies and technological development vector for AI-created content: such associations of market players are believed necessary and useful for identifying the AI-related regulatory development vector, adopting guidelines and the underlying rules of procedure. A special-purpose consultative public body can be set up under the Roskomnadzor.

The problems of implementing personal rights in a new context (in particular, digital) cannot be adequately and comprehensively addressed unless the methods and means of private and public law are used in conjunction.

References

1. Blitz M.J. (2020) Deepfakes and other Non-Testimonial Falsehoods: When is Belief Manipulation (Not) First Amendment Speech. *Yale Journal of Law & Technology,* vol. 23, pp. 161–300.

2. Chesney B., Citron D. (2019) Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, vol. 107, issue 6, pp. 1753–1819. DOI: 10.15779/Z38RV0D15J.

3. Criminal Law of Russia. General part. Textbook (2016) Sundurov F.T., Tarhanov I.A. (eds.). Moscow: Statut, 864 p. (in Russ.)

4. Hine E., Floridi L. (2022) New Deepfake Regulations in China are a Tool for Social Stability, but at what Cost? *Nature Machine Intelligence,* vol. 4, pp. 608–610.

5. Joost L. (2023) Place for Illusions: Deepfake Technology and the Challenges of Regulating Unreality. *University of Florida Journal of Law and Public Policy*, vol. 33, no. 2, pp. 309–332.

6. Kalyatin V.O. (2022) Deepfake as a Legal Problem: New Threats or New Opportunities? *Zakon*=Law, no. 7, pp. 87–103. DOI: 10.37239/0869-4400-2022-19-7-87-103 (in Russ.)

7. Mason S., Seng D. (2021) Artificial Intelligence and Evidence. *Singapore Academy of Law Journal,* issue 33. pp. 241–279.

8. Meskys E., Kalpokiene J., Jurcys P. et al. (2020) Regulating Deep Fakes: Legal and Ethical Considerations. *Journal of Intellectual Property Law & Practice,* vol. 15, issue 1, pp. 24–31. DOI: 10.1093/jiplp/jpz167.

9. Pfefferkorn R. (2020) «Deepfake» in the Courtroom. *Boston University Public Interest Law Journal,* vol. 29, issue 2, pp. 245–276.

10. Rafil R. (2023) Libel, Slander and Defamation Law: the Basics. Available at: https://www.findlaw.com/injury/torts-and-personal-injuries/defamation-law-the-basics.html (accessed: 22.07.2024)

11. Roberts T. (2023) How to do Things with Deepfakes. *Synthese*, issue 43, pp. 1–18. DOI: 10.1007/s11229-023-04044-2.

12. Sundurov F.T., Talan M.V. (2015) Punishment in Criminal Law: study guide. Moscow: Statut, 296 p. (in Russ.)

13. Tianren L., Yue D. Defamation and Privacy Law in China. Available at: https://www.carter-ruck.com/law-guides/defamation-and-privacy-law-in-china/ (accessed: 22.07.2024)

14. Velasco C. (2022) Cybercrime and Artificial Intelligence. An overview of the work of international organizations on criminal justice and the international applicable instruments. *ERA Forum*, issue 23, pp. 109–126. DOI: 10.1007/s12027-022-00702-z.

15. Vinogradov V.A., Kuznetsova D.V. (2024) Foreign Experience in Legal Regulating Deepfake Technology. *Pravo*. *Zhurnal Vysshey shkoly ekonomiki*=Law. Journal of the Higher School of Economics, vol. 17, issue 2, pp. 215–240. DOI:10.17323/2072-8166.2024.2.215.240 (in Russ.)

16. Young N. (2019) Deepfake Technology: Complete Guide to Deepfakes, Politics and Social Media. North Charleston (S.C.): Independently published, 160 p.

Information about the author:

V.O. Demkin — Postgraduate Student.

The article was submitted to editorial office 26.08.2024; approved after reviewing 30.09.2024; accepted for publication 07.10.2024.