Л.Н. Гумилев атындағы Еуразия ұлттық университетінің ХАБАРШЫСЫ.
ISSN: 2616-6844. eISSN: 2663-1318

ҚҰҚЫҚ СЕРИЯСЫ/ LAW SERIES/ СЕРИЯ ПРАВО

Халықаралық құқық, халықаралық жеке құқық / International law, Private International Law/ Международное право, международное частное право

IRSTI 10.77.51
Scientific article

https://doi.org/10.32523/2616-6844-2025-152-3-212-228

International and domestic legal frameworks on online fraud and deepfake technologies: a comparative criminal law analysis

K.M. Beaver*

College of Criminology and Criminal Justice, Florida State University, USA, Florida

(e-mail: kevinmichaelbeaver@gmail.com)

Abstract: The relevance of this study stems from the rapid development of digital technologies and the introduction of generative artificial intelligence, which has given rise to new forms of online fraud, particularly those involving deepfakes. These technologies, capable of producing audio and video content that is virtually indistinguishable from authentic sources, significantly enhance the manipulative potential of fraudulent schemes and complicate their detection. Existing international and national legal mechanisms remain insufficiently adapted to these emerging challenges.

The article aims to provide a comparative analysis of international and domestic legal frameworks addressing online fraud involving deepfake technologies, to identify legal gaps, and to suggest directions for their resolution. The object of the research is deepfake technologies and social engineering in the context of fraud, while the subject concerns criminal law mechanisms for their prevention and prosecution.

The methodological framework combines system and comparative analysis, content analysis of legal instruments and academic literature, case studies, as well as classification and generalization.

The findings indicate that international instruments, such as the Budapest Convention, lack explicit provisions addressing deepfakes, thereby reducing legal certainty and requiring expansive interpretation. At the domestic level, substantial divergence is observed: while the United States, the United Kingdom, and the European Union are developing specialized norms, post-Soviet jurisdictions predominantly rely on general provisions on fraud and forgery. Judicial practice reveals difficulties in the qualification of offenses, detection of synthetic content, forensic examination, and ensuring consistency in enforcement.

The study concludes that effective responses require the elaboration of specialized legal norms, the harmonization of international approaches, and the strengthening of institutional cooperation. The practical significance lies in providing recommendations for the modernization of national legal systems and the enhancement of international efforts to counter transnational threats of online fraud facilitated by deepfake technologies.

Keywords: deepfake technologies, online fraud, criminal law, legal frameworks, artificial intelligence.

Received: 16.09.2025. Accepted: 24.09.2025. Available online: 30.09.2025

***corresponding author**

Introduction

The rapid expansion of digital technologies is reshaping economic and social processes while simultaneously altering the structure of criminal activity. Traditional strategies for addressing cyber threats appear increasingly insufficient given the emergence of generative artificial intelligence (AI), which is capable of producing synthetic audio and video content that is nearly indistinguishable from authentic material. Under these conditions, criminal law is confronted with qualitatively novel forms of illicit behavior for which appropriate legal mechanisms are often lacking.

Evidence of rising threats to digital security is documented both in international statistics and in empirical reports of financial losses. According to Europol [1], cybercrime in Europe continues to grow, with AI-based technologies playing an increasingly significant role. Comparable patterns are reported by the FBI Internet Crime Complaint Center [2], which indicated that in 2021 alone, reported losses from internet fraud exceeded $6.9 billion, with a large proportion involving advanced social engineering techniques. Such findings indicate that cybercrime is transitioning from relatively simple schemes to more technologically complex forms of deception that require specialized investigative and legal responses.

One critical element of this transformation involves the proliferation of synthetic media, most commonly referred to as deepfakes. These applications rely on deep neural networks and machine learning methods to generate audio and video content with high levels of realism. Initially used in entertainment and digital modeling, deepfake technologies have rapidly acquired criminogenic potential, serving as tools for online fraud. Recent evidence suggests that deepfakes act as catalysts for increasingly sophisticated criminal strategies. Examples include voice imitation of corporate executives, falsified video calls, and manipulated images used to create fraudulent social media accounts. These practices enhance traditional forms of deception while making detection more difficult for both victims and law enforcement. Scholars have emphasized that synthetic media undermine trust in visual and auditory sources of information, thereby creating risks not only for individuals and organizations but also for broader social stability [3].

The accessibility of deepfake-generation software has further exacerbated the problem. The production of such content needed specialized skills and powerful computers in the past, but multiple free applications now make it accessible to criminals. The new technology enables extensive distribution of digital scams, which demands a review of existing legal systems for their effectiveness.

The online fraud enabled by Deepfake technology operates across multiple countries because digital spaces eliminate geographical restrictions. Offenders use digital environments to operate from various locations while hosting servers in foreign territories and using anonymization tools that make it difficult to identify them. Research evidence shows that cybercrime remains invisible because of its international nature. For instance, the Interpol Global Crime Trend Report 2022 [4] noted that more than 60% of surveyed countries identified financial fraud, phishing, online scams, and ransomware as high or very high threats. Yet references to deepfake-related fraud remain limited, and the absence of clear legal definitions intensifies uncertainty. At present, many jurisdictions rely on provisions originally designed for conventional fraud, privacy violations, or misappropriation of personal data, which do not adequately capture the unique risks associated with synthetic media [5].

The lack of harmonized approaches produces jurisdictional gaps frequently exploited by criminal groups. Differences in legal regimes complicate extradition, recognition of electronic evidence, and the implementation of joint investigations. As a result, effective responses to deepfake-enabled fraud cannot be achieved solely at the national level but require institutionalized international cooperation. The global character of the problem further underscores the necessity of examining legal frameworks at both universal and regional levels and evaluating the extent to which existing mechanisms can support coordinated responses to new digital threats [6].

Contemporary criminal law provisions on fraud and forgery were developed in the context of traditional forms of crime and thus do not adequately account for the digital environment. In practice, deepfake-related conduct is often prosecuted under general provisions addressing fraud, unlawful use of personal data, or violations of image rights. The current legal frameworks do not adequately protect society from synthetic media risks because these technologies can compromise identification systems and judicial processes, and online communication. The international legal framework does not contain specific rules about deepfake technology. The Budapest Convention on Cybercrime (2001) and its 2021 Protocol focus on traditional computer crimes instead of addressing modern digital threats. The current legal framework creates uncertainty about liability because practitioners must use broad interpretations, which results in inconsistent judicial decisions.

The current legal framework demonstrates an urgent requirement for a complete theoretical framework that addresses how deepfake technologies change fraud practices. A scholarly approach demands the development of precise definitions and new criteria for qualification and clear definitions of legal responsibility. From a practical perspective, it involves providing recommendations for policymakers and practitioners aimed at adapting criminal law to emerging digital threats.

The development of such approaches carries dual significance. Nationally, it contributes to the modernization of domestic legal systems and increases resilience against new forms of crime. Internationally, it establishes a foundation for harmonization of legal practices, which is essential given the cross-border nature of deepfake-enabled fraud. Consequently, research in this area may serve as the basis for legislative initiatives and institutional reforms designed to balance the protection of legal order with the safeguarding of fundamental human rights.

Methodology

The research material comprised academic publications, legal instruments, statistical reports, and documented case studies, encompassing both qualitative and quantitative data. The methodological framework included system analysis, which made it possible to examine the issue as an interaction of technological, legal, and social factors; comparative analysis, which facilitated the assessment of international approaches to regulation; content analysis of academic, legal, and empirical sources; and case studies, which provided detailed insights into specific instances of deepfake use in fraudulent schemes and their legal qualification. In addition, methods of classification and typology were applied to systematize various forms of deepfake-related fraud, while synthesis and generalization were employed to develop comprehensive conclusions and recommendations. The combined use of these methods ensured a multidimensional perspective on the transformation of fraud under the influence of deepfake technologies, allowing for the comparison of international practices, the systematization of fraudulent models, and the elaboration of well-founded conclusions.

Findings/Discussion

The present analysis focuses on two interrelated phenomena that have become central to contemporary criminal law scholarship. The first concerns online fraud, which is defined in legal doctrine as the unlawful acquisition of property or other benefits through deception or abuse of trust conducted in the digital environment using information and communication technologies. For example, the Council of Europe's Convention on Cybercrime (Budapest Convention, 2001) includes provisions on computer-related fraud, defined as the intentional causing of property loss to another through the input, alteration, deletion, or suppression of computer data with fraudulent intent to secure economic gain [7].

Empirical evidence accumulated over recent decades indicates the wide variety of such offenses, ranging from classical phishing and credential theft to more complex forms of social engineering and manipulation of digital identifiers.

The second phenomenon involves deepfake technology, which consists of generative artificial intelligence algorithms capable of synthesizing audio and video content that closely replicates authentic sources. These applications have become feasible due to advances in deep neural networks and deep learning methods, allowing for high levels of fidelity in imitating human appearance and voice. Chesney and Citron provide a detailed analysis of how deepfake technologies enable the production of media that appear to depict individuals performing or saying actions that never occurred and how such outputs may be resistant to detection [8].

Although initially developed for entertainment and digital modeling purposes, these technologies have rapidly acquired criminogenic potential. Their intersection with online fraud is reflected in the way deepfakes serve as instruments for qualitatively enhancing fraudulent practices. Whereas earlier forms of online fraud relied primarily on textual or simple visual deception, synthetic media now permit real-time identity imitation, falsification of evidence, and manipulation of victim trust at a substantially higher level.

This development generates a category of criminal conduct that extends beyond classical understandings of online fraud and raises new issues for legal classification and international cooperation in combating cybercrime. The emergence and rapid diffusion of generative artificial intelligence technologies have altered both the form and scale of fraudulent activity in digital contexts. The capacity to convincingly replicate visual and auditory characteristics of specific individuals has increased the manipulative potential of fraud while complicating detection by law enforcement. The integration of traditional social engineering with synthetic media illustrates a qualitative shift from relatively rudimentary schemes to more sophisticated models of criminal behavior.

The most common types of deepfake technologies include video generation, audio synthesis, and static image creation. Video deepfakes involve the substitution of an individual's visual identity in dynamic form. Through neural network algorithms, videos are produced in which one person’s face is superimposed on another's or an individual’s behavior is fully simulated. The high degree of realism enables their use in falsifying video conferences, remote identification processes, and other contexts in which visual presence is a key component of trust.

Audio deepfakes are based on voice cloning and the reproduction of speech while retaining distinctive vocal characteristics. Contemporary generative models can generate audio files and streaming speech that are virtually indistinguishable from the original. The application of this technology appears most often in fraud schemes, which require spoken communication for tasks including financial transaction verification and corporate directive issuance and personal interactions.

The static nature of image-based deepfakes includes photographs that people use to create fake identification documents and social media profiles, and advertising content. The lack of dynamic features in these images does not reduce their criminal potential because they enable identity theft and digital service deception through biometric identification systems. The different types of synthetic media serve as a foundation for various fraudulent activities which will receive detailed analysis in the upcoming sections. A prominent example involves the integration of deepfake into mechanisms of social engineering. By imitating the voice of a corporate executive or other trusted figure, offenders can initiate telephone calls or audio messages that induce recipients to authorize financial transfers or disclose sensitive information. This strategy aligns with the category of “CEO fraud,” which has been widely documented in corporate environments.

Video deepfakes are also employed in impersonation during remote negotiations and online identification procedures. The use of an artificial interlocutor in a video conference, or a generated likeness of a client during banking verification, complicates authentication. In combination with phishing, these methods produce highly convincing deception scenarios in which victims receive visually “verified” evidence of the counterpart’s authenticity.

The production of fake documents and fraudulent online accounts continues to use static deepfakes as a tool. The creation of fake portraits that do not belong to actual people enables the development of synthetic identities, which criminals use to get credit and create bank accounts and conduct illegal financial activities. The combination of traditional identity theft methods with deepfake technology makes it harder to detect and prevent such crimes.

Deepfakes present a major security risk because they can be used to create synthetic content for mass communication attacks and cyberattacks. The use of synthetic content in emails and social media, and messaging applications enables attackers to run phishing scams against numerous users at once. Video and audio content that mimics real people in messages proves more effective at deception because it outperforms traditional text-based messages.

The mentioned practices indicate deepfakes serve as tools to boost current fraud methods while enabling the creation of novel criminal schemes. The criminological value of deepfakes becomes more pronounced when social and economic operations shift to digital platforms, which requires legal systems to adapt. Research shows deepfake technology has evolved from theoretical threats to actual operational tools used in fraudulent schemes. The corporate world demonstrates this threat clearly because attackers use fake audio and video recordings of executives to carry out unauthorized financial transactions. In one 2019 case, £243,000 was stolen from a British energy company through the use of voice-cloning technology, where perpetrators convincingly imitated the German parent company's chief executive officer to instruct a subsidiary manager to authorize a payment [9].

Such incidents highlight vulnerabilities in corporate governance under conditions of digital communication. The impersonation of private individuals through synthetic images or videos also extends beyond corporate contexts. Falsified profiles created with deepfakes have been used in social networks and online platforms for fraudulent purposes, including romance and commercial scams, financial exploitation, and disinformation campaigns. The emergence of "synthetic identities" compounds the risks associated with conventional identity fraud, since their digital traces often appear authentic. Reports on deepfake threats have frequently documented these concerns. For example, Europol's Internet Organised Crime Threat Assessment (IOCTA) 2020 emphasized the increasing risks linked to technologies capable of producing realistic identity simulations and falsified audio-visual content [10].

Of particular concern is the integration of deepfakes into phishing mechanisms. The attackers have used fake videos that mimic banking staff and government personnel to trick victims into revealing their sensitive information. The attacks based on video deception prove more successful than text-based scams, according to both observational data and cybersecurity evaluations.

The research shows that deepfake fraud has evolved from single incidents into a widespread pattern that affects economic and legal systems. The growing number of these threats requires an assessment of current legal systems to determine their ability to handle emerging security risks. The evaluation of international legal frameworks becomes essential because they define minimum requirements for cybercrime prevention and establish how countries can work together to fight cross-border online scams.

The international legal framework serves as the main authority for creating standardized solutions to handle worldwide problems, including deepfake technology-based scams. The worldwide nature of digital space makes national efforts insufficient, so universal and regional legal instruments become essential for effective regulation. At the same time, the existing framework remains fragmented and lacks specific provisions directly addressing abuses involving synthetic media.

The most significant universal instrument is the Council of Europe's Convention on Cybercrime (Budapest Convention, 2001), which established comprehensive state obligations to criminalize computer-related fraud, unauthorized access, data manipulation, and misuse of devices [11]. Article 7 ("Computer-related forgery") specifies liability for the creation of "inauthentic data," while Article 8 ("Computer-related fraud") addresses property loss through manipulation of computer data or systems. Although deepfakes are not explicitly mentioned, these provisions can be applied when synthetic media are used as instruments of conventional cybercrime. The Second Additional Protocol (2021) expanded cross-border cooperation in obtaining electronic evidence, but likewise did not directly address generative technologies.

Within the United Nations framework, digital crime is referenced in the UN Convention against Transnational Organized Crime (2000), which sets out general obligations for combating cross-border criminality and promoting international cooperation [12]. Additional relevance is provided by documents concerning data protection and human rights, including the General Assembly resolutions "The Right to Privacy in the Digital Age" (A/RES/68/167, 2013; A/RES/71/199, 2016), which emphasize the need to safeguard fundamental rights in online contexts [13].

Regional mechanisms also remain important. In the European Union, Directive 2013/40/EU on attacks against information systems and the General Data Protection Regulation (GDPR, 2016) establish standards of security and protection of personal data [14; 15]. While neither instrument explicitly references deepfake, both provide a foundation for addressing abuses linked to synthetic media. Other regional organizations have likewise advanced initiatives: the African Union adopted the Malabo Convention on Cybersecurity and Personal Data Protection (2014) [16], and the Asia-Pacific Economic Cooperation (APEC) has undertaken efforts to harmonize approaches to digital regulation.

Beyond binding treaties, soft-law instruments also exert influence. UNESCO's Recommendation on the Ethics of Artificial Intelligence (2021), endorsed by all member states, outlines principles of transparency, accountability, and non-discrimination in AI development and use [17]. Similarly, the OECD AI Principles (2019) emphasize safety, reliability, and human rights protection in the application of artificial intelligence [18]. Viewed collectively, these soft-law norms reflect a trend toward developing general approaches to the governance of generative technologies, including deepfake, through ethical frameworks and interdisciplinary dialogue. Although not legally binding, these instruments shape national legal systems and contribute to the eventual inclusion of specific provisions in binding agreements.

Nevertheless, limitations remain evident. First, no existing international framework contains explicit provisions on synthetic media. Most conventions and protocols focus on traditional cybercrime, requiring expansive interpretation when applied to deepfake and reducing legal certainty. Second, the transnational nature of deepfake-enabled fraud – where offenders, servers, and victims may reside in different jurisdictions – renders territorial jurisdiction ineffective. The lack of unified extradition standards and electronic evidence-sharing mechanisms further complicates enforcement, as documented by law enforcement agencies. Third, even among states party to the Budapest Convention, levels of harmonization remain low, with differences in the definitions of computer-related fraud and related offenses limiting recognition of judgments and joint investigations.

Taken together, these factors suggest that the existing international legal framework does not provide sufficient adaptability to challenges associated with deepfake technologies. This creates the need for additional codification and the development of specialized mechanisms aimed at addressing synthetic media as tools of online fraud.

Analysis of national legal systems provides insight into the readiness of states to counter deepfake-enabled fraud. While universal obligations exist, national criminal law ultimately determines how offenses are classified and prosecuted. Comparative evidence demonstrates considerable variation: some jurisdictions have introduced new provisions and initiatives, while others rely on conventional fraud, forgery, or image-rights statutes.

In the United States, regulation has emerged at the state level. California's AB 730 (2019) prohibits the dissemination of manipulated audio or video content within 60 days of an election for the purpose of deceiving voters [19]. Texas enacted SB 751 to prohibit the use of deepfake in political advertising [20]. In 2025, Texas also passed SB 20 ("Stopping AI-Generated Child Pornography Act"), which criminalizes the production and distribution of AI-generated child pornography. At the federal level, the Take It Down Act was signed in May

2025 to facilitate the removal of non-consensual intimate images and deepfake content from online platforms.

In the United Kingdom, regulation remains based on the Fraud Act 2006 and the Computer Misuse Act 1990, neither of which explicitly references deepfake. Nevertheless, in 2025, the Ministry of Justice announced plans to criminalize the creation and dissemination of non-consensual sexually explicit deepfake images [21]. Reports by the Law Commission (2022–2024) similarly emphasized the need to reform legislation to address unauthorized creation and sharing of intimate images, including deepfake [22].

In the European Union, regulation remains largely indirect. The GDPR (2016) provides legal grounds for protecting individuals against the misuse of personal data, including voice and image. Directive 2013/40/EU establishes liability for computer-related crime. Additionally, the proposed AI Act, currently in the final stages of consideration, sets transparency and accountability requirements for generative technologies. In the context of deepfake, these initiatives lay the groundwork for the eventual introduction of explicit criminal law provisions across EU member states.

Legal systems in the post-Soviet region are characterized by a conservative approach to the classification of offenses involving deepfake technologies. Unlike the United States and several European Union member states, where steps toward specialized provisions have been undertaken, regulation in Russia, Kazakhstan, and other states of the region relies on general provisions of criminal law.

In the Russian Federation, the use of synthetic media in fraudulent schemes is addressed under general provisions on fraud in Article 159 of the Criminal Code, and may also fall under statutes related to the unlawful use of personal identifiers (Articles 272, 273, 274 of the Criminal Code) [23]. When images infringing personal rights are created or disseminated, civil law provisions concerning honor, dignity, and business reputation (Article 152 of the Civil Code) and rules regarding image rights are applied. No independent provision specifically criminalizes the use of deepfake, forcing practitioners to rely on analogy or broad interpretation.

A similar situation is observed in Kazakhstan. Article 190 of the Criminal Code defines liability for fraud [24], while the use of deepfake for forgery or deception falls under general provisions addressing property crimes and information security. Legislation regulating digitalization and personal data likewise does not include specific references to deepfake, creating difficulties for enforcement. Legislative proposals have been introduced to amend Article 190 by including the use of artificial intelligence and deepfake as an aggravating factor or as a separate offense [25].

Other states in the region, including Belarus, Armenia, and Uzbekistan, appear to lack provisions explicitly addressing deepfake as a distinct legal category. Practitioners instead rely on existing articles concerning fraud, forgery, personal rights, and reputation protection.

Judicial practice reflects the absence of systematic treatment of deepfake-related fraud. While case law remains limited, individual cases highlight emerging issues. Internationally, high-profile examples have involved voice synthesis in financial crimes. In 2021, in the United Arab Emirates, a transfer exceeding $35 million was executed following a phone call in which perpetrators imitated a company director’s voice and provided falsified correspondence [26].

Although no completed trials have yet established precedent, such incidents have guided law enforcement in classifying these offenses as fraud involving new technical means [27].

In the United Kingdom, case law involving deepfake has been largely limited to image rights violations and dissemination of false materials. For an extended period, courts applied general provisions on fraud and defamation. With the enactment of the Online Safety Act in 2023, and the criminalization in January 2025 of non-consensual sexually explicit deepfake images carrying penalties up to imprisonment [21], the legal framework has begun to reflect the specificity of the technology, though case law remains limited.

In the post-Soviet region, fraud-related cases involving deepfake remain isolated. In 2023, the Moscow Arbitration Court reviewed a case concerning the unauthorized use of Keanu Reeves's likeness in a deepfake advertisement, awarding compensation for infringement of copyright and related rights [28]. Fraud-related uses of the technology, however, continue to be classified under general criminal provisions such as fraud or unlawful use of personal data, without recognition of deepfake as a qualifying element.

In sum, these examples indicate that judicial practice does not yet provide a consistent approach to the classification of deepfake-related offenses. The absence of specific provisions compels courts to rely on expansive interpretation of existing statutes, reducing predictability in enforcement and hindering the development of coherent judicial doctrine.

Comparison of national legal approaches demonstrates substantial variation in the degree to which criminal law has been adapted to address fraud involving deepfake technologies. In Anglo-Saxon jurisdictions such as the United States and the United Kingdom, as well as within the European Union, there is a discernible trend toward the development of specialized provisions and broader strategies for regulating artificial intelligence. Within these systems, deepfake is increasingly recognized as a distinct regulatory object requiring refined classification and targeted sanctions.

By contrast, in post-Soviet jurisdictions, criminal law remains primarily rooted in traditional categories. Regulation continues to rely on general provisions concerning fraud, document forgery, and violations of image rights, without explicit recognition of the technological specificity of generative algorithms. As a result, enforcement frequently depends on an expansive interpretation of existing provisions, which complicates the establishment of consistent judicial practice.

The level of preparedness for digital evidence handling, together with the presence or absence of specialized norms, determines the extent of regulatory asymmetry. Jurisdictions that modernized their systems developed institutional frameworks to accept digital evidence and perform forensic analysis for detecting tampering. The establishment of effective prosecution faces challenges because states with underdeveloped regulatory systems are still working to create these procedures.

The different approaches between jurisdictions prove that deepfake regulation lacks worldwide standardization. The existing differences between jurisdictions create legal ambiguities while criminals use these gaps to conduct international criminal operations because they find vulnerabilities in national systems. The situation demands immediate development of standardized international rules and unified national legal systems.

The analysis of comparative data shows that international commitments and domestic programs have not successfully transformed the current legal system to effectively combat deepfake-related fraudulent activities. The legal system faces three major weaknesses because it lacks dedicated provisions and fails to implement technical detection tools in forensic work and lacks standardized international procedures for exchanging information and evidence. The current limitations prevent the development of reliable enforcement methods that enable transnational criminal activities to expand.

A particularly serious challenge involves the difficulty of establishing the use of synthetic media in specific cases. Contemporary generative algorithms can produce images, video, and audio recordings that are nearly indistinguishable from authentic sources. Traditional forensic methods, designed to identify conventional document or file forgeries, are ill-suited to detecting artificially generated material. Research in digital forensics has found that even specialized detection algorithms demonstrate limited effectiveness, with problems of generalization, dataset variability, and performance speed relative to generative models. For example, a review in Deepfake video detection: challenges and opportunities [29] highlights that "limited quality and diversity of labeled data, high dependency on computational resources, and practical reliability all create significant obstacles" for robust detection of deepfakes.

The absence of standardized forensic methodologies complicates evidentiary processes in court. In most jurisdictions, digital evidence must satisfy criteria of reliability and reproducibility, yet these criteria are often unmet in cases involving deepfake. Further challenges arise in international cooperation, where differences in national standards hinder recognition of forensic findings.

These considerations suggest a need for unified approaches to forensic identification of deepfake. Effective solutions require collaboration among states, the research community, and the private sector. Without such mechanisms, the capacity of criminal law to address fraud involving synthetic media remains severely constrained.

One of the most significant challenges for enforcement remains the absence in most national systems of provisions explicitly establishing criminal liability for the use of deepfake technologies in fraudulent activity. Unlike traditional forms of forgery or deception, where legal constructs are relatively settled, synthetic media generate novel circumstances that do not align with established categories. Practitioners are therefore compelled to rely on analogy or expansive interpretation, reducing the predictability of judicial outcomes and weakening the principle of legal certainty.

Difficulties also arise in distinguishing deepfake-enabled fraud from related offenses. Depending on the circumstances of a case, conduct may be classified as document forgery, unlawful use of personal data, identity theft, or dissemination of false information. The inconsistent application of legal standards makes it difficult to establish a unified approach, which hinders the creation of a consistent judicial doctrine.

The majority of criminal codes lack detailed regulations about generative algorithms, which makes it difficult to apply existing legal provisions to synthetic media. The legal standards that were created for traditional forgery cases fail to work effectively when used to handle synthetic media. Even in jurisdictions where normative adaptation is underway, no clear definition of "deepfake" or criteria for its use in a criminal law context exist. This ambiguity impedes the establishment of offense elements and hinders international harmonization.

Findings suggest that addressing these gaps requires the development of specialized legal norms that take into account both the technological features of deepfake and the need to balance prosecution with protection of human rights, including freedom of expression and privacy. Without such refinement, criminal law systems risk remaining insufficiently effective under conditions of rapid digital transformation.

Beyond legal gaps, the use of deepfake in fraudulent contexts raises a broad range of ethical and human rights concerns. The dual-use character of the technology – capable of serving both innovation and abuse – intensifies the need for balanced approaches that protect against misuse while preserving legitimate applications.

Central to this debate is the protection of image rights and personal integrity. Manipulation of an individual's appearance or voice without consent can produce severe consequences for reputation, professional activity, and psychological well-being. In digital environments, such harms acquire a transnational dimension, as synthetic content is not constrained by territorial borders, complicating the enforcement of protective rights.

Another significant issue is the balance between combating abuse and safeguarding freedom of expression. Deepfake may be employed in art, journalism, and political satire, where overly restrictive regulation risks unwarranted interference with creativity and information exchange. Accordingly, criminal law frameworks must account not only for crime-control objectives but also for human rights obligations, including those enshrined in the European Convention on Human Rights (1950).

The ethical framework includes two main aspects, which deal with discrimination and public opinion manipulation. Deepfake technology enables users to spread fake information and create artificial evidence, which damages public trust in judicial bodies and political systems and news organizations. The breakdown of information reliability threatens democratic principles and creates social instability in relationships between people.

Multiple factors demonstrate the requirement for legal systems to establish mechanisms that unite criminological knowledge with technical expertise and ethical standards. A complete solution that combines all necessary elements will protect fundamental freedoms while providing effective solutions for emerging threats.

Online deepfake fraud requires international collaboration between countries for effective prosecution and evidence sharing because it operates across national borders. Unlike conventional offenses, which often remain within the jurisdiction of a single state, fraudulent activity involving synthetic media frequently displays a distributed structure: offenders may operate from one jurisdiction, servers hosting or distributing the content may be located in another, and victims may reside in a third. Under these conditions, traditional principles of territorial jurisdiction are insufficient to ensure effective enforcement.

A central difficulty arises from the fragmented character of national approaches to the legal classification of deepfake-related offenses. In the absence of unified definitions and qualifying elements, requests for mutual legal assistance and extradition encounter significant obstacles. Even within the Budapest Convention framework, which remains the primary instrument addressing cybercrime, no clear provisions explicitly regulate synthetic media. This gap reduces the efficacy of inter-state cooperation, limiting opportunities for joint investigations and prosecutions.

Differences in procedural standards for admitting electronic evidence represent an additional challenge. In some jurisdictions, strict admissibility and authenticity criteria are applied, whereas in others such determinations remain subject to judicial discretion. Consequently, evidence collected in one jurisdiction may be excluded in another, diminishing the likelihood of successful transnational prosecutions.

Responsibility allocation between states and transnational digital platforms presents a further unresolved issue. Platforms play a central role in the dissemination of synthetic media, yet their obligations to prevent abuse lack uniform international codification. The absence of a global mechanism for regulating platform responsibilities creates additional barriers to the prevention and suppression of deepfake-enabled fraud.

The domain needs enhanced institutional power and standardized legal frameworks for international cooperation to succeed. The absence of unified norms and procedures makes states susceptible to transnational criminal activities that generative technologies enable.

Conclusion

Digital technology advancements combined with generative artificial intelligence have revolutionized criminal activities by creating advanced online fraud schemes that utilize deepfake technology. The ability to generate realistic audio and video content through these technologies makes fraud more deceptive and harder for law enforcement to detect. As demonstrated by recent cases, deepfakes are used to impersonate executives, falsify video communications, fabricate documents and “synthetic identities,” and reinforce phishing campaigns, thus posing a serious threat to individuals, organizations, and broader societal stability.

The analysis has shown that existing legal frameworks at both the international and national levels remain insufficiently adapted to these developments. International instruments such as the Budapest Convention do not contain provisions specifically addressing deepfake, which forces reliance on expansive interpretation and reduces legal certainty. National jurisdictions demonstrate wide variation: while some legal systems, notably those of the United States and the United Kingdom, have initiated the development of specialized norms, others – including states of the post-Soviet region – continue to rely on general provisions on fraud, forgery, and image rights. The lack of standard definitions and qualifying criteria for deepfake offenses creates difficulties for judicial operations while blocking the development of unified enforcement strategies.

The transnational nature of deepfake-enabled fraud creates additional difficulties because it crosses borders and makes it difficult to apply traditional jurisdictional rules effectively. The detection of falsified content becomes challenging for forensic practice because standard methods fail to identify sophisticated synthetic media and specialized detection algorithms struggle with reliability and practical implementation and generalization. The use of deepfake technology creates multiple ethical problems, which include protecting personal images and voices and maintaining privacy rights and finding a balance between crime prevention and free speech protection.

A complete solution to these problems needs to unite legal aspects with institutional frameworks and technical solutions and ethical considerations. The development of exact legal definitions for deepfake technology requires the simultaneous implementation of specific laws that penalize deceptive deepfake applications. The international community needs to create standardized legal frameworks while developing better systems for cross-border cooperation through mutual legal assistance and extradition and electronic evidence sharing and digital platform responsibility definitions. States and the scientific community and the private sector need to work together to develop forensic methods that detect synthetic content reliably. Legal frameworks need to protect fundamental human rights while implementing deepfake regulations because such measures should not restrict artistic or journalistic or political uses of synthetic media.

These combined actions will help nations update their laws and build stronger defenses against new criminal activities while establishing unified global strategies to combat deepfake technology threats.

This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan (Grant No. AP26103625 «Online fraud using deepfake-technologies and social engineering: problems of criminal law counteraction, prospects for legislative regulation»).

The contribution of the authors

K.M. Beaver has done research and written the entire article.

References

1 Europol. Internet Organised Crime Threat Assessment (IOCTA) 2024. Publications Office of the European Union, Luxembourg, 2024. – Режим доступа: https://www.europol.europa.eu/publication-events/main-reports/internet-organised-crime-threat-assessment-iocta-2024 (дата обращения: 15.04.2025).

2 Federal Bureau of Investigation. Internet Crime Report 2021. – Режим доступа: https://www.ic3.gov/AnnualReport/Reports/2021_ic3report.pdf (дата обращения: 20.03.2025).

3 Chesney R., Citron D. Deepfakes and the New Disinformation War. 2018. – Режим доступа: https://www.foreignaffairs.com/articles/world/2018-12-11/deepfakes-and-new-disinformation-war (дата обращения: 10.02.2025).

4 Financial and cybercrimes top global police concerns, says new INTERPOL report. 19.10.2022. – Режим доступа: https://www.interpol.int/en/News-and-Events/News/2022/Financial-and-cybercrimes-top-global-police-concerns-says-new-INTERPOL-report (дата обращения: 05.04.2025).

5 The Deepfake Menace: Legal Challenges in the Age of AI. TRT World Research Centre, 2024. – Режим доступа: https://researchcentre.trtworld.com/wp-content/uploads/2024/03/The-Deepfake-Menace_v2.pdf (дата обращения: 12.03.2025).

6 Singh T. Cybercrime and International Law: Jurisdictional Challenges and Enforcement Mechanisms // African Journal of Biomedical Research. 2024. Vol.27. No3S. Pp. 697–708. https://doi.org/10.53555/AJBR.v27i3S.2101

7 Cybercrime. Module 2: General Types of Cybercrime. Key Issues. Computer-related offences. United Nations Office on Drugs and Crime (UNODC). – Режим доступа: https://www.unodc.org/e4j/zh/cybercrime/module-2/key-issues/computer-related-offences.html (дата обращения: 18.01.2025).

8 Citron D.K., Chesney R. Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security // California Law Review. 2019. Vol.107. Pp. 1753–1819. – Режим доступа: https://scholarship.law.bu.edu/faculty_scholarship/640 (дата обращения: 25.04.2025).

9 Damiani J. A Voice Deepfake Was Used To Scam A CEO Out Of $243,000. 03.09.2019. – Режим доступа: https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/ (дата обращения: 22.02.2025).

10 Europol. Internet Organised Crime Threat Assessment (IOCTA) 2020. European Union Agency for Law Enforcement Cooperation, 2020. – Режим доступа: https://www.europol.europa.eu/cms/sites/default/files/documents/internet_organised_crime_threat_assessment_iocta_2020.pdf (дата обращения: 14.03.2025).

11 European Treaty Series – No. 185. Convention on Cybercrime. Budapest, 23.XI.2001. – Режим доступа: https://rm.coe.int/1680081561 (дата обращения: 12.02.2025).

12 United Nations. Convention against Transnational Organized Crime and the Protocols thereto. Adopted by the UN General Assembly: 15 November 2000, Resolution 55/25. – Режим доступа: https://www.unodc.org/unodc/en/organized-crime/intro/UNTOC.html (дата обращения: 10.01.2025).

13 Office of the High Commissioner for Human Rights (OHCHR). Privacy in the digital age. – Режим доступа: https://www.ohchr.org/en/privacy-in-the-digital-age/international-standards (дата обращения: 20.03.2025).

14 Directive 2013/40/EU of the European Parliament and of the Council of 12 August 2013 on attacks against information systems and replacing Council Framework Decision 2005/222/JHA. – Режим доступа: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32013L0040 (дата обращения: 05.02.2025).

15 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). – Режим доступа: https://eur-lex.europa.eu/eli/reg/2016/679/oj (дата обращения: 15.04.2025).

16 African Union. African Union Convention on Cyber Security and Personal Data Protection. Date of adoption: 27.06.2014. – Режим доступа: https://au.int/en/treaties/african-union-convention-cyber-security-and-personal-data-protection (дата обращения: 28.01.2025).

17 UNESCO. Recommendation on the Ethics of Artificial Intelligence. SHS/BIO/PI/2021/1. – Режим доступа: https://unesdoc.unesco.org/ark:/48223/pf0000381137 (дата обращения: 18.03.2025).

18 OECD. About the OECD AI Principles. – Режим доступа: https://www.oecd.org/en/topics/sub-issues/ai-principles.html (дата обращения: 25.02.2025).

19 California State Legislature. Assembly Bill 730. 2019–2020 Regular Session. – Режим доступа: https://legiscan.com/CA/text/AB730/id/2055885 (дата обращения: 30.01.2025).

20 Guidry, T. Texas did it first: Texas was the first to enact state legislation on the use of deep fakes in elections. TCPAWorld, 28.05.2024. – Режим доступа: https://natlawreview.com/article/texas-did-it-first-texas-was-first-enact-state-legislation-use-deep-fakes-elections (дата обращения: 10.04.2025).

21 Government crackdown on explicit deepfakes. Ministry of Justice; Davies-Jones A. – Режим доступа:https://www.gov.uk/government/news/government-crackdown-on-explicit-deepfakes (дата обращения: 20.01.2025).

22 Law Commission. Taking, making and sharing intimate images without consent. – Режим доступа: https://lawcom.gov.uk/project/taking-making-and-sharing-intimate-images-without-consent/ (дата обращения: 05.02.2025).

23 The Criminal Code of the Russian Federation of June 13, 1996 № 63-ФЗ. – Режим доступа: https://www.consultant.ru/document/cons_doc_LAW_10699/8012ecdf64b7c9cfd62e90d7f55f9b5b7b72b755/ (дата обращения: 15.03.2025).

24 Penal Code of the Republic of Kazakhstan. The Code of the Republic of Kazakhstan dated 3 July 2014 №226-V of the Law of the Republic of Kazakhstan. – Режим доступа: https://adilet.zan.kz/eng/docs/K1400000226 (дата обращения: 10.02.2025).

25 Criminal liability for deepfakes may be introduced in Kazakhstan. – Режим доступа: https://bluescreen.kz/ugholovnuiu-otvietstviennost-za-dipfieiki-moghut-vviesti-v-kazakhstanie/ (дата обращения: 20.04.2025).

26 Brewster T. Fraudsters cloned company director's voice in $35 million heist, police find. – Режим доступа: https://www.forbes.com/sites/thomasbrewster/2021/10/14/huge-bank-fraud-uses-deep-fake-voice-tech-to-steal-millions/ (дата обращения: 15.11.2024).

27 Financial Crimes Enforcement Network (FinCEN). FinCEN alert on fraud schemes involving deepfake media. FIN-2024-Alert004. 13.11.2024. – Режим доступа: https://www.fincen.gov/system/files/shared/FinCEN-Alert-DeepFakes-Alert508FINAL.pdf (дата обращения: 05.01.2025).

28 Russian court imposes first-ever compensation order in unauthorized Keanu Reeves deepfake use case. 06.12.2023. – Режим доступа: https://theins.ru/en/news/267363 (дата обращения: 18.01.2025).

29 Kaur A., Noori Hoshyar A., Saikrishna V., et al. Deepfake video detection: Challenges and opportunities // Artificial Intelligence Review. 2024. Vol.57. P. 159. https://doi.org/10.1007/s10462-024-10810-6

К.М. Бивер

Флорида штатының университетінің Криминология және қылмыстық сот төрелігі колледжі, АҚШ, Флорида

(e-mail: kevinmbeaver@hotmail.com)

Онлайн алаяқтық пен дипфейк технологияларының халықаралық және ішкі мемлекеттік құқықтық шеңберлері: қылмыстық заңнаманы салыстырмалы-құқықтық талдау

Аңдатпа: Зерттеудің өзектілігі цифрлық технологиялардың жылдам дамуы мен генеративті жасанды интеллекттің енгізілуімен айқындалады. Бұл жағдай онлайн-алаяқтықтың жаңа түрлерінің, соның ішінде дипфейктерді қолдану арқылы жасалатын қылмыстардың пайда болуына алып келді. Шынайы контенттен ажырату қиын аудио және бейнематериалдарды құра алатын бұл технологиялар қылмыстық схемалардың манипулятивтік әлеуетін күшейтіп, олардың әшкереленуін қиындатады. Қолданыстағы халықаралық және ұлттық құқықтық тетіктер мұндай сын-қатерлерге жеткілікті бейімделмеген.

Мақаланың мақсаты – дипфейктерді қолданатын онлайн-алаяқтыққа қарсы іс-қимылды реттейтін халықаралық және ұлттық құқықтық негіздерді салыстырмалы талдау, құқықтағы олқылықтарды анықтау және оларды жоюдың бағыттарын ұсыну. Зерттеу нысаны – онлайн-алаяқтықтағы дипфейк технологиялары мен әлеуметтік инженерия, зерттеу пәні – оларды шектеудің қылмыстық-құқықтық тетіктері.

Әдістемелік база жүйелі және салыстырмалы талдау, құқықтық актілер мен академиялық еңбектердің контент-талдауы, кейс-стади, сондай-ақ классификация және жалпылау әдістерінен тұрады.

Нәтижелер халықаралық құжаттарда, мысалы, Будапешт конвенциясында, дипфейктерге қатысты тікелей нормалардың жоқтығын көрсетті. Бұл құқықтық айқындықты төмендетеді. Ұлттық құқықтық жүйелерде айтарлықтай айырмашылықтар бар: АҚШ, Ұлыбритания және ЕО арнайы нормаларды әзірлеуде, ал посткеңестік елдер негізінен алаяқтық пен қолдан жасауға қатысты жалпы ережелерге сүйенеді. Сот тәжірибесінің талдауы қылмыстарды саралауда, синтетикалық контентті анықтауда, сараптамалар жүргізуде және құқық қолданудың біріздiлігін қамтамасыз етуде қиындықтарды айқындайды.

Мақалада арнайы нормаларды әзірлеудің, халықаралық тәсілдерді үйлестірудің және ынтымақтастықты институционалды нығайтудың қажеттілігі атап өтіледі. Зерттеудің практикалық маңызы – ұлттық құқықтық жүйелерді жаңғыртуға және трансұлттық цифрлық алаяқтыққа қарсы күрестің тиімділігін арттыруға бағытталған ұсыныстарды қалыптастыру.

Түйін сөздер: дипфейк-технологиялар, онлайн алаяқтық, қылмыстық құқық, құқықтық негіздер, жасанды интеллект.

К.М. Бивер

Колледже криминологии и уголовного правосудия Университета штата Флорида,
США, Флорида
(e-mail: kevinmbeaver@hotmail.com)

Международные и внутригосударственные правовые рамки онлайн-мошенничества и технологий дипфейк: сравнительно-правовой анализ уголовного законодательства

Аннотация: Актуальность исследования определяется стремительным развитием цифровых технологий и внедрением генеративного искусственного интеллекта, что обусловило появление новых форм онлайн-мошенничества, в частности с использованием дипфейков. Эти технологии, способные создавать аудио- и видеоконтент, практически неотличимый от оригинала, значительно усиливают манипулятивный потенциал преступных схем и затрудняют их выявление. Существующие международные и национальные правовые механизмы оказываются недостаточно подготовленными к таким вызовам.

Цель статьи заключается в проведении сравнительного анализа международных и национальных правовых рамок, регулирующих противодействие онлайн-мошенничеству с применением дипфейков, а также в выявлении пробелов и выработке направлений их устранения. Объектом исследования выступают технологии дипфейков и социальная инженерия в контексте мошенничества, предметом – уголовно-правовые механизмы их пресечения.

Методологическая база исследования включает системный и сравнительный анализ, контент-анализ правовых актов и академических публикаций, метод кейс-стади, а также классификацию и обобщение.

Результаты показывают, что международные документы, такие как Будапештская конвенция, не содержат прямых положений, касающихся дипфейков, что снижает правовую определённость. Национальные правовые системы демонстрируют значительные различия:

США, Великобритания и ЕС разрабатывают специализированные нормы, тогда как страны постсоветского региона полагаются на общие положения о мошенничестве и подлоге. Анализ судебной практики выявляет сложности квалификации деяний, установления факта использования синтетических медиа, проведения экспертиз и обеспечения единообразия правоприменения.

Выводы статьи акцентируют необходимость разработки специализированных норм, гармонизации международных подходов и институционального укрепления сотрудничества. Практическая значимость исследования заключается в формировании рекомендаций для модернизации национальных систем и повышения эффективности борьбы с транснациональными угрозами цифрового мошенничества.

Ключевые слова: дипфейк-технологии, онлайн-мошенничество, уголовное право, правовые рамки, искусственный интеллект.

Information about the author:

Beaver K.M. – PhD, Judith Rich Harris Professor of Criminology, Florida State University, 222 S. Copeland Street, Tallahassee, FL 32306, USA

Бивер К.М. – PhD докторы, криминология профессоры (Judith Rich Harris), Флорида штаты университеті, 222 S. Copeland Street, Таллахасси, FL 32306, АҚШ

Бивер К.М. – доктор философии (PhD), профессор криминологии (Judith Rich Harris), Университет штата Флорида, 222 S. Copeland Street, Таллахасси, FL 32306, США