Ismayil Musayev. Legal framework and countermeasures addressing deepfake risks in Europe. International Law and Integration Problems, 2025, №2 (71), pp. 138-149. DOI: 10.61638/ZAPV3676

LEGAL FRAMEWORK AND COUNTERMEASURES ADDRESSING DEEPFAKE RISKS IN EUROPE

Ismayil Musayev*

Abstract

This article critically examines the existing European legal frameworks addressing the challenges posed by deepfake technology. It highlights the significant risks associated with deepfakes, including misinformation, infringements on personal image rights, privacy violations, and threats to democratic processes. Central to the analysis is the European Union's Artificial Intelligence Act (AI Act), set to come into force in August 2025, which introduces the first binding legal definition and regulatory obligations for high-risk AI applications, including deepfakes. The article also considers complementary instruments such as the GDPR, the Digital Services Act, and relevant national legislation within Member States like Germany and France. Despite these regulatory advances, structural gaps remain – particularly in enforcement mechanisms and the inclusion of end-users – that hinder comprehensive governance. The article underscores the necessity for a multifaceted, coordinated approach combining harmonized enforcement, technological oversight, public awareness, and evolving legal measures to effectively regulate synthetic media technologies within the European legal order.

Keywords: *artificial intelligence, deepfake, AI Act, image rights, AI governance in Azerbaijan.*

I. Introduction

This article addresses the pivotal question of whether the existing European legal framework adequately regulates the challenges posed by deepfake technology. Deepfake technology, created through advanced artificial intelligence methods, enables the generation or manipulation of highly realistic synthetic audiovisual content. While these innovations offer valuable applications in sectors like education and entertainment, their misuse raises significant risks including misinformation, violation of personal image rights, reputational harm, and threats to democratic processes. In response, the European Union (the EU) has developed a comprehensive legal framework that addresses these challenges at multiple levels. Central to this is the Artificial Intelligence Act (AI Act)— the world's first binding legal instrument that regulates AI technologies using a structured, risk-based classification system. Scheduled to enter into force in August 2025, the AI Act sets explicit obligations for high-risk AI systems and imposes transparency requirements, particularly regarding deepfake technologies. Importantly, it provides a forward-looking definition of "deepfake", becoming the first legal instrument in Europe—and globally—to explicitly define and regulate this phenomenon within a binding framework. The Act's extraterritorial reach and harmonized enforcement mechanisms make it a pioneering global standard for AI governance. Alongside the AI Act, other crucial European regulations contribute to deepfake governance, including the General Data Protection Regulation (GDPR), the Digital Services Act (DSA), and the e-Commerce Directive. At the national level, Member States such as France, Germany have adopted or are developing specific laws to complement EU-wide measures, addressing issues like defamation, image rights, and electoral integrity. This article explores the technical foundations of deepfake technology, the societal and legal risks it poses, the European Union's regulatory response—focusing particularly on the AI Act—and

* Leading Specialist, Center for Legal Examination and Legislative Initiatives, email: ismayilmusayev05@gmail.com

relevant national legislation within selected Member States. It further analyzes how fundamental rights enshrined in the European Convention on Human Rights interact with these emerging technologies, aiming to provide a comprehensive legal perspective on the governance of deepfakes in Europe.

II. What is deepfake?

In recent years, deepfake technology has emerged as increasingly pressing issue, raising significant legal, ethical, and societal concerns across Europe and beyond. Its rapid development has outpaced existing legal frameworks, challenging regulators to respond effectively.

The emergence of the term "deepfake" traces back to a Reddit user who adopted the same pseudonym, blending "deep learning" with "fake" to describe face-swapping videos [11, p. 2]. Initially, this user's intent was to share pornographic content featuring the faces of well-known celebrities digitally superimposed onto actors' bodies without their consent [11, p. 2]. By releasing the source code behind the technology, the user inadvertently sparked a wave of interest across the Reddit platform, accelerating the mass production and distribution of such manipulated content. Initially, prominent public figures became the primary subjects of these early deepfake creations. Even in its early stages, the misuse of deepfakes revealed significant risks to personal dignity, privacy, and public trust, laying the groundwork for today's more complex legal and ethical debates.

A deepfake refers to any form of content—whether video, audio, or other media—that is either entirely fabricated or has been digitally altered using advanced manipulation techniques [30, p. 1]. This may involve generating synthetic material from scratch or modifying existing recordings to misrepresent reality [30, p. 1]. According to the European Parliamentary Research Service, deepfakes are defined as artificially generated or altered audio-visual content that convincingly imitates real individuals, making them appear to say or do things they never actually said or did [31, p. 4]. These highly realistic synthetic media are created using artificial intelligence technologies, specifically deep learning algorithms and generative adversarial networks (GANs).

Deep learning is a type of machine learning that uses neural networks to analyze data and find patterns by mimicking the human brain's processes [9, p. 7]. It enables computers to learn more effectively from large datasets, which are essential for training generative models [9, p. 7]. The availability of vastly labelled visual datasets on the internet has accelerated the development of deepfake technology. By learning important features and their relationships from data, deep learning algorithms create realistic and convincing deepfake content.

Generative Adversarial Networks (GANs), introduced by Goodfellow et al. in 2014, enable the production of highly realistic synthetic images by optimizing them to be statistically indistinguishable from genuine images, thereby blurring the line between real and artificially generated content [10, p. 3]. GANs function through the interaction of two models: one that generates synthetic content and another that evaluates its authenticity [9, p. 8]. This process allows the system to refine its outputs over time, resulting in synthetic media that increasingly resembles real data. As a result, GANs have significantly improved the effectiveness and accessibility of deepfake creation by streamlining the learning process and enhancing the realism of generated content.

III. The Threats Posed by Deepfakes

Although deepfake technology brings notable advantages, it raises significant concerns due to its potential to erode public trust, spread misinformation, and challenge the distinction between fact and fiction. The advent of deepfake technology has facilitated various forms of disinformation, manifesting in diverse contexts. The implications of such misuse are profound and multifaceted, spanning the political sphere—through election interference and propaganda dissemination; the financial sector—via fraud schemes and market manipulation; and reputational domains—by facilitating defamation, identity theft, and personal or institutional credibility erosion.

During armed conflicts, deepfake technology has emerged as a powerful tool for manipulating public perception and spreading disinformation. On March 2, 2022, a deepfake video falsely showing Ukrainian President Volodymyr Zelenskyy urging soldiers and citizens to surrender briefly appeared on a national news outlet. Although the video closely imitated his appearance and speaking style, Ukrainian authorities quickly dismissed it as inauthentic. This marked the first high-profile use of deepfake technology during an armed conflict, signalling a shift in the nature of information warfare [4, p. 2].

One of the primary legal challenges posed by deepfake technology concerns the infringement of individuals' privacy rights through unauthorized use of their identity, image, or voice. While publicized cases involving celebrities' manipulated likenesses have drawn significant attention, most of the harm disproportionately affects private individuals who are often subjected to blackmail, harassment, or retaliation [1, p. 147]. These issues underscore urgent needs for strengthened legal protections and regulatory mechanisms to safeguard personal dignity and reputation in the digital era. The creation of deepfakes frequently results in violations of human rights by unlawfully processing personal data, as an individual's image constitutes inherently identifiable information. Without the explicit and informed consent of the data subject, the use of their likeness in synthetic media production breaches core data protection principles, since reliance on a purported legitimate interest fails to provide adequate legal justification [1, p. 147]. Accordingly, the unauthorized creation and dissemination of manipulated visual content should be subject to clear and enforceable legal prohibitions to safeguard personal privacy and uphold fundamental data protection rights in the digital environment.

AI-generated deepfakes present a profound threat to personal and professional reputations. As numerous public figures have been targeted with fabricated stories and defamatory remarks, the swift and widespread circulation of such material on social media platforms further exacerbates the resulting damage. Beyond mere reputational harm, these synthetic media can fundamentally erode public trust and be weaponized by malicious actors to propagate misinformation and orchestrate sophisticated deception campaigns.

Deepfake technology undermines intellectual property protections by enabling unauthorized replication and manipulation of protected audio-visual content, thereby complicating the enforcement of copyright and trademark laws. This raises critical legal questions concerning ownership, attribution, and the adequacy of traditional frameworks to address AI-generated synthetic media in the digital age. In 2020, Roc Nation, the U.S.-based entertainment company founded by Shawn Carter, professionally known as Jay-Z — an internationally renowned American rapper and music producer — filed copyright takedown notices against AI-generated deepfake videos [2]. These videos, created by an anonymous YouTuber known as Vocal Synthesis, used artificial intelligence to replicate Jay-Z's voice, featuring him performing Billy Joel's "We Didn't

Start the Fire" and the Hamlet soliloquy from Shakespeare [2]. Roc Nation asserted that the content unlawfully used AI to impersonate their client's voice, which led to the removal of the videos from YouTube. This landmark case highlights the urgent need to modernize and reinforce copyright law to effectively counteract the novel threats posed by AI-driven deepfakes, thereby safeguarding the integrity and value of creative works in an increasingly digital landscape.

Deepfake technologies also pose a significant threat to political systems, particularly during election periods, by enabling the creation and dissemination of deceptive content that can manipulate public opinion and undermine the integrity of democratic processes. In the United States, notable recent cases include a deepfake audio of a Chicago mayoral candidate allegedly endorsing police violence, as well as a robocall mimicking President Joe Biden's voice intended to suppress voter turnout during a primary election [29, p. 5]. These incidents demonstrate the potential of deepfake technology to manipulate public opinion and disrupt electoral participation, thereby raising critical concerns about political disinformation and the integrity of democratic processes.

IV. Europe

The European Union is actively addressing the legal and societal risks associated with deepfake technologies. Recognizing the potential of deepfakes to undermine democratic processes, spread disinformation, and infringe upon fundamental rights such as privacy and personal image, European institutions have started to explore regulatory and policy-based approaches aimed at mitigating these risks and ensuring accountability for the misuse of such technologies.

According to a study conducted by the European Parliamentary Research Service, various categories of risks associated with deepfakes have been systematically identified. The study outlines psychological, financial, and societal harms as key areas of concern. Psychological risks include forms of abuse such as extortion, defamation, intimidation, bullying, and a general erosion of trust [31, p. 4]. In terms of financial harm, the study reveals threats such as extortion, identity theft, fraud (including insurance and payment-related schemes), manipulation of stock prices, brand damage, and broader reputational loss. The societal dimension encompasses risks like the manipulation of news media, threats to economic stability, the integrity of the justice and scientific systems, the deterioration of democratic institutions, election interference, diplomatic tensions, and vulnerabilities in national security. These findings underscore the multifaceted and systemic impact of deepfake technologies across both individual and institutional levels.

Within the European Union, a comprehensive legal framework exists to combat the challenges posed by deepfake technologies, encompassing both hard law and soft law instruments. This dual approach enables European legal experts and policymakers to effectively utilize a variety of regulatory tools in addressing the multifaceted risks associated with deepfakes. Among the key legal instruments employed in this domain are:

- *The Artificial Intelligence Act*, which establishes a risk-based regulatory regime targeting AI systems, encompassing those capable of generating or manipulating synthetic audiovisual content;
- *The European Convention on Human Rights (ECHR),* which safeguards fundamental rights including privacy (Article 8) and freedom of expression (Article 10), both of which may be engaged in the context of deepfake technology;

- *The General Data Protection Regulation (GDPR),* which protects individuals' personal data and biometric information, thereby addressing unauthorized usage within deepfake creations;
- *The Digital Services Act (DSA),* introducing enhanced transparency, accountability, and content moderation obligations for digital platforms to mitigate the spread of harmful or unlawful synthetic media;
- *The European Democracy Action Plan*, focused on fortifying democratic integrity and safeguarding electoral processes from technological manipulation, including the risks posed by deepfake content;
- *The e-Commerce Directive*, setting forth rules on the liability and responsibilities of online intermediaries, including platforms potentially hosting deepfake content;
- *The Audiovisual Media Services Directive (AVMSD),* ensuring regulatory oversight over audiovisual content dissemination, encompassing manipulated or synthetic media;
- *The Code of Practice on Disinformation*, a soft law voluntary framework fostering cooperation among platforms and stakeholders to combat misinformation and deceptive content online;
- *The EU Action Plan Against Disinformation*, outlining strategic measures aimed at enhancing the resilience of public discourse and democratic processes against coordinated disinformation efforts;
- *European Parliament resolutions;*
- *National laws of Member States.*

The Artificial Intelligence Act (AI Act). The European Commission released its Proposal for the Artificial Intelligence Act on April 12, 2021, marking a significant step in global efforts to regulate AI. The Act entered into force on August 1, 2025, establishing a comprehensive legal framework for the ethical and responsible use of AI across the European Union [14, p. 1]. It seeks to position Europe as a preeminent hub for trustworthy AI by instituting unified rules that govern the development, commercialization, and deployment of AI technologies throughout the EU.

The framework sets out the following objectives: 1. to ensure AI systems placed on the Union market are safe and respect fundamental rights and Union values; 2. to provide legal certainty that fosters investment and innovation; 3. to enhance governance and enforce applicable laws on AI safety and fundamental rights; and 4. to promote a unified market for lawful, safe, and trustworthy AI applications, preventing fragmentation [13, p. 2].

The AI Act provides for an extensive territorial scope with respect to its provisions on AI systems. Its provisions apply where an AI system—whether standalone or integrated into a product governed by EU product safety legislation—is placed on the EU market, put into service, or used by a deployer established or located within the Union [3]. It applies regardless of the provider's place of establishment—broadly interpreted in line with other EU laws—and also covers AI-generated outputs created outside the EU but used within it, thus including non-EU providers and deployers [3].

According to Article 3(60) of the AI Act, a 'deepfake' is defined as "AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful" [17, art. 3(60)]. Article 3(60) of the AI Act provides a context-sensitive definition of deepfakes by restricting the term to synthetic audio-visual content, explicitly excluding AI-generated texts, and clarifying that such content is not necessarily limited to human representations [17, art. 3(60)]. This legally precise and operational definition is likely to

serve as a normative benchmark for legal scholars and regulatory authorities, fostering greater conceptual clarity and contributing to the development of consistent and effective legislative and policy responses to the challenges posed by deepfake technologies.

Annex III of the AI Act lists specific AI systems that are considered high-risk due to their potential impact on safety, fundamental rights, or critical sectors. For instance, AI systems used for biometric identification in public spaces, credit scoring that affects individuals' access to financial services, or AI employed in the recruitment and evaluation of job applicants are all classified as high-risk under Annex III [18]. Additionally, AI applications in critical infrastructure, such as electricity supply or water management, are also included due to their potential societal impact [18]. Article 7 of the AI Act grants the European Commission the authority to update the catalogue of high-risk AI applications, providing a critical tool for ensuring the regulation remains responsive to emerging technologies [17, art. 7]. These systems must comply with stricter regulatory requirements to ensure transparency, reliability, and accountability.

However, the AI Act's limited applicability to end users excludes non-professional creation and dissemination of deepfakes, raising serious legal gaps—particularly in cases involving disinformation or non-consensual content [14, p. 12]. Assuming provider-level safeguards are sufficient is questionable, as technical circumvention is feasible and state or non-state actors are unlikely to be deterred. These deficiencies can undermine the effectiveness of the current framework and necessitate broader regulatory measures.

The European Convention on Human Rights (ECHR). Deepfake technology presents significant challenges to image rights, particularly where a person's likeness is digitally manipulated and disseminated without consent. Such interference with an individual's visual identity may constitute a violation of the right to respect for private life under Article 8 of the European Convention on Human Rights, insofar as it undermines the individual's autonomy and control over the use of their image. Since individuals generally do not hold copyright over their own image, copyright law offers limited resources for protecting one's persona in such cases. In the case of Reklos and Davourlis v. Greece, the European Court of Human Rights (ECtHR) affirmed that an individual's image constitutes a fundamental aspect of personal identity, noting that it reflects unique characteristics and serves to distinguish one from others [20, para. 40]. The Court emphasized that the right to control the use of one's image is an essential element of personal development and autonomy under Article 8 of the European Convention on Human Rights. In this light, the creation and dissemination of deepfakes—particularly where a person's likeness is digitally altered and distributed without consent—may constitute a violation of the right to private life.

The legal responses to deepfake technologies vary significantly across jurisdictions, with countries adopting distinct terminologies and imposing both civil and criminal liabilities under different legal frameworks. Only after all domestic remedies have been duly exhausted may individuals bring their claims before ECtHR, in accordance with Article 35 of the ECHR. However, cases involving deepfakes often raise complex questions concerning the infringement of fundamental rights, particularly the right to freedom of expression under Article 10 and the right to respect for private and family life under Article 8 of ECHR.

However, it should be noted that the utilization of deepfake technology presents intricate legal dilemmas, as it may invoke a conflict between the protection of the right to respect for private and family life under Article 8 of ECHR and the safeguarding of

freedom of expression under Article 10, thereby requiring a nuanced and proportionate balancing exercise consistent with the Court's established jurisprudence. Although political deepfakes typically involve the intentional dissemination of false or misleading information, a blanket prohibition based solely on their inaccuracy would raise serious concerns under Article 10 of the European Convention on Human Rights. In one of its judgments ECHR held that Article 10 does not prohibit sharing or discussing information suspected to be false, as doing so would unduly restrict the fundamental right to freedom of expression [25, para. 113]. The resolution of this dilemma requires a careful and context-specific approach that upholds freedom of expression while protecting privacy rights, ensuring that any restrictions on deepfake content are justified, necessary, and proportionate considering the competing interests involved.

Given that the European Court of Human Rights has not yet developed case law or provided interpretation regarding the regulation of deepfakes under Articles 8 and 10 of the Convention, Member States retain discretion to tackle the harmful effects of such content through existing legal frameworks, so long as these measures remain consistent with the safeguards of Article 10. The ECtHR has ruled that interference with the rights to privacy and freedom of expression under Articles 8 and 10 is only in accordance with the law if it is grounded in domestic legislation, accessible to the public, and meets the requirements of foreseeability and the rule of law as defined by the ECtHR [21, para. 151]. The ECtHR has held that a measure is justified in a democratic society only if it addresses a pressing social need, pursues a legitimate aim, and is proportionate [23, para. 76]. The reasons advanced by the state must be relevant and sufficient, and while states enjoy a margin of appreciation, the Court retains ultimate authority to assess the necessity and proportionality of the interference [24, para. 78].

In some instances, rights violations involving the use of deepfake technology may be perpetrated by anonymous users, thereby complicating the identification of those responsible and the pursuit of legal accountability. The European Court of Human Rights examined the tension between online privacy and the need for accountability in a case where a member State of the Council of Europe had failed to require an Internet service provider to reveal the identity of an individual who had posted an inappropriate advertisement involving a minor on a dating platform. The Court acknowledged that while the protection of private communications and the right to freedom of expression are essential, these rights are not without limits. In certain situations - such as preventing crime, safeguarding public order, or protecting the rights of others - these freedoms may be lawfully restricted. Accordingly, States are under a positive obligation to establish a legal framework that carefully balances these competing interests [22, para 49].

The Budapest Convention does not criminalise anonymous communication via computer systems. Its Explanatory Report considers the alteration of traffic data for anonymity or encryption as a lawful exercise of privacy rights. However, Parties may criminalise abuses, such as modifying packet headers to conceal identity in the commission of a crime [5, para. 62]. The Council of Europe's Committee of Ministers, in its Declaration on Freedom of Communication on the Internet, recognised the right to anonymity as a safeguard for freedom of expression and protection against surveillance. Member States are encouraged to respect users' choices to remain unidentified. However, this does not preclude lawful measures to identify individuals involved in criminal conduct, in line with national law, the ECHR, and relevant international instruments.

The provision on the disclosure of subscriber information under the Second Additional Protocol to the Budapest Convention on Cybercrime establishes a crucial legal mechanism for identifying anonymous users involved in illicit online activities, including the dissemination of deepfake content. While this measure does not itself constitute a technical tool for detecting the creators of deepfakes, it empowers competent authorities to request subscriber data directly from service providers across borders, thereby enabling the identification of individuals behind anonymous digital accounts [28, art. 7]. This legal framework is essential for overcoming anonymity in cybercrime investigations and facilitating effective cross-border cooperation. However, technical expertise and additional investigative methods remain necessary to conclusively attribute responsibility in deepfake cases.

The General Data Protection Regulation (GDPR). The creation of deepfakes often involves processing personal data, such as voice samples or images that identify natural persons, thus qualifying as personal data under data protection laws. The processing of an individual's audio-visual data can be considered personal data processing, as defined by the General Data Protection Regulation (GDPR) as "any information relating to an identified or identifiable natural person ('data subject')" [15, art. 4]. The processing of such data is subject to the conditions established by the GDPR, which safeguards individuals' rights to privacy and data protection. Consequently, the GDPR protects the personal data of individuals using deepfake technology-based applications and applies to the development, creation, and dissemination of deepfake software and content.

However, 2 notable legal gaps exist regarding the GDPR's applicability to deepfakes: those concerning deceased individuals and virtual persons. Personal data of deceased people falls outside the regulation's scope, and additionally, the GDPR does not apply to political deepfakes featuring virtual, non-existent individuals, as data protection laws are irrelevant when the generated content does not pertain to any real person [16]. Thus, while the GDPR offers important protections for personal data in the context of deepfake technology, significant gaps remain that limit its effectiveness in addressing different categories of deepfake content.

e-Commerce Directive. Adopted in the early phase of internet commercialization, the EU's e-Commerce Directive serves as a foundational instrument for regulating online content by exempting intermediary service providers from any general obligation to monitor transmitted information or conduct ex ante content verification, with the aim of fostering digital service growth through minimal regulatory burden [31]. Although certain online platforms engage in voluntary detection of deepfake content, there is no legal obligation compelling them to do so. Moreover, service providers are only required to act upon credible indications of unlawfulness, thereby shifting the evidentiary burden onto individuals or authorised flaggers [31]. While the e-Commerce Directive permits the removal of unlawful deepfake content, it lacks a clear definition of "illegal content" and fails to harmonise the criteria for establishing liability, thereby creating legal uncertainty for intermediaries [31]. Overall, the e-Commerce Directive laid important groundwork for regulating online content with a light-touch approach to foster digital growth; however, its lack of clear definitions and harmonised liability criteria creates significant legal uncertainties, particularly regarding the handling of deepfake content.

The Digital Services Act (DSA) introduces additional rules and standards for hosting providers while maintaining the notice-and-action framework. Although it builds upon

the e-Commerce Directive, both legislative acts currently coexist, and the DSA does not fundamentally alter the regulatory approach to deepfake content [30, p. 12].

V. Experience of Selected European Countries

The Republic of Italy. Italy governs deepfakes under established laws protecting personal rights, image, and privacy. The Italian Civil Code forbids unauthorized use of a person's likeness, and in conjunction with the Italian Copyright Law (Articles 96 and 97), these regulations empower individuals to claim compensation when their image is exploited without consent, particularly if it damages their honour or reputation [27, art. 96-97].

Germany. Currently, Germany lacks specific legislation explicitly regulating deepfake technology. Nevertheless, the challenges and harms resulting from the misuse of deepfakes are addressed under existing laws, primarily through provisions in the Strafgesetzbuch (StGB) — the German Criminal Code. Key provisions such as Section 187 StGB on defamation and Section 201a StGB on violations of personal privacy and image rights are applicable to false or manipulated content disseminated via deepfakes [6, art. 187; 201a]. These statutes enable legal action against reputational damage and infringements of personal rights arising from malicious deepfake usage.

In 2024 the German Bundesrat has proposed a new criminal offense (§201b StGB) targeting the creation and distribution of manipulated digital content, including deepfakes used for defamation, non-consensual pornography, or election interference [12]. However, the federal government argues that existing laws (§187 and §201a StGB) sufficiently address these harms and raises concerns about legal uncertainties and reduced protections in the proposal [12].

France. France's SREN Law, adopted on May 21, 2024, amends Article 226-8 of the French Criminal Code and introduces a new provision, Article 226-8-1, to the same Code, explicitly prohibiting the non-consensual creation and sharing of deepfake content, with stricter penalties for pornographic deepfakes [26, art. 15; 21]. The law criminalizes sharing algorithmically generated visual or audio content depicting a person without consent, unless clearly identified as artificial, imposing fines and imprisonment that increase if distributed via public online platforms [26, art. 15; 21]. Despite these advances, the broad definition of "algorithmic processing" and exceptions for obviously generated content create legal uncertainties requiring judicial interpretation [7].

Denmark. Additionally, Denmark has recently introduced a legislation granting individuals copyright over their own likeness to combat the spread of deepfakes and unauthorized digital imitations, emphasizing protection against misinformation and misuse of personal identity [8]. If enacted, this legislation would represent a pioneering legal framework as the first law to explicitly grant individuals copyright protection over their own likeness, specifically targeting the prevention of deepfakes and the unauthorized use of digital representations, thereby addressing emerging challenges in personal identity and misinformation.

Azerbaijan. Although Azerbaijan is geographically located at the crossroads of Eastern Europe and Western Asia, outside the core European Union countries, its membership in the Council of Europe, participation in the European Political Community, adherence to the European Convention on Human Rights, and active cooperation with the EU under the Eastern Partnership programme since 2009 position it as a close partner involved in European legal and political cooperation. While specific legislation on deepfakes remains absent, recent policy initiatives and Azerbaijan's active

engagement with European legal standards demonstrate a growing commitment to addressing the challenges posed by synthetic media technologies.

Existing general legal frameworks in Azerbaijan can provide mechanisms for the legal assessment of cases involving deepfakes. Deepfakes may be employed as tools in the commission of various offenses, potentially triggering criminal liability under the Azerbaijani Criminal Code. For example, under Article 147 (Defamation), disseminating a deepfake that unjustly damages a person's reputation may constitute a punishable offense. Similarly, Article 182 (Extortion) may likewise be invoked in circumstances where manipulated digital content is employed to compel an individual to relinquish property. Furthermore, Azerbaijan is a party to the Budapest Convention on Cybercrime, which provides an international legal framework facilitating cooperation in combating cybercrimes, including offenses that may involve the misuse of technologies such as deepfakes.

Notably, the National Artificial Intelligence Strategy for 2025–2028, adopted by the government, reflects an increasing awareness of such risks at the policy level. While the Strategy is not legally binding, it highlights several critical objectives relevant to the governance of technologies like deepfakes. These include the development of a responsible and ethical regulatory framework for artificial intelligence, the protection of citizens' rights and privacy, the prevention of disinformation and manipulative content, and the safeguarding of national and information security. Specifically, the Strategy acknowledges that technologies such as deepfakes may be used by foreign actors, terrorist organizations, or transnational criminal groups to undermine national interests, spread disinformation, and erode public trust.

This policy-level recognition, although still in an early stage, suggests an emerging institutional readiness to address the legal and ethical challenges posed by deepfake technologies. It also signals the need for Azerbaijan to gradually transition from general legal provisions to more targeted regulatory mechanisms capable of addressing the complexities of synthetic media in the digital age.

VI. Conclusion

Deepfake technology reflects a dual nature: while it offers substantial benefits in areas such as healthcare, accessibility, and entertainment, it simultaneously generates serious legal and ethical challenges. These include disinformation, non-consensual manipulation, reputational harm, and violations of privacy. Addressing such risks requires a regulatory framework that balances the promotion of innovation with the protection of fundamental rights.

The European Union's regulatory response—chiefly through the AI Act—marks a notable milestone. By introducing a legal definition of deepfakes and establishing transparency obligations, the Act lays down a foundational framework. Nevertheless, the regulation suffers from several limitations, particularly its exclusion of end users, weak enforcement mechanisms, and the risk of circumvention. These structural gaps undermine the Act's ability to comprehensively address the evolving threats posed by deepfakes.

While instruments such as GDPR, DSA, and the interpretative guidance of the ECtHR provide complementary layers of protection, the EU's regulatory landscape remains fragmented. The rapid development of generative AI technologies necessitates continuous legal recalibration and stronger coordination among the Member States.

Crucially, the AI Act should not be regarded as a definitive or exhaustive solution. Instead, it must be seen as a transitional regulatory tool requiring iterative refinement.

Member States retain the competence and indeed the responsibility to adopt supplementary national measures. The fact that the Act applies directly should not be understood as preventing the implementation of more ambitious domestic safeguards aimed at strengthening rights protected under Articles 8 and 10 of the ECHR.

To safeguard these rights effectively and ensure that deepfake technologies are not weaponized against the public interest, further steps are needed. These include clearer definitions, enforceable accountability mechanisms, and practical tools for oversight—both at provider and user levels. Ultimately, the effective governance of deepfakes in Europe depends on a multi-faceted strategy: harmonized enforcement, transnational cooperation, technological vigilance, public awareness, and where appropriate, the introduction of new legal instruments specifically targeting synthetic manipulations.

References:

1. Balcerzak, M., Kapelańska-Pręgowska, J., Artificial Intelligence and International Human Rights Law, Edward Elgar Publishing, Paris, 2024. 1st Edition.

2. Beaumont-Thomas, B., Jay-Z takes action against "deepfakes" of him rapping Hamlet and Billy Joel,

URL: https://www.theguardian.com/music/2020/apr/29/jay-z-files-takes-action-against-deepfakes-of-him-rapping-hamlet-and-billy-joel (last access: 29.05.2025).

3. Bird & Bird, European Union Artificial Intelligence Act: A Guide,

URL:https://www.twobirds.com/-/media/new-website-content/pdfs/capabilities/artificial-intelligence/european-union-artificial-intelligence-act-guide.pdf (last access: 29.05.2025).

4. Byman, D.L., Gao, C., Deepfakes and International Conflict,

URL:https://www.brookings.edu/wp-content/uploads/2023/01/FP_20230105_deepfakes_international_conflict.pdf (last access: 29.05.2025).

5. Budapest Convention on Cybercrime, Article 2, Explanatory Report, 2001.

6. Criminal Code of the Federal Republic of Germany, 1949 (in German / *Strafgesetzbuch – StGB, Bundesrepublik Deutschland, 1949*).

7. Coslin, C., France prohibits non-consensual deep fakes,

URL: https://www.hoganlovells.com/en/publications/france-prohibits-non-consensual-deep-fakes (last access: 29.05.2025).

8. Desmarais, A., Denmark Fights Back Against Deepfakes with Copyright Protection. What Other Laws Exist in Europe?

URL: https://www.euronews.com/next/2025/06/30/denmark-fights-back-against-deepfakes-with-copyright-protection-what-other-laws-exist-in-e (last access: 29.05.2025).

9. Europol Innovation Lab, Facing reality? Law enforcement and the challenge of deepfakes.

URL: https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf (last access: 29.05.2025).

10. Goodfellow, I., Generative Adversarial Networks, Advances in Neural Information Processing Systems, V. 3, N. 11, 2014. p. 1-9

11. Kietzmann, J., Lee, L.W., Deepfakes: Trick or Treat?, Business Horizons, V. 63, N. 2, 2019, p. 1–12.

12. Knoke, L., Germany's Legal Debate on Criminal Liability for Misuse of Deepfakes: Navigating Uncharted Waters

URL: https://technologyquotient.freshfields.com/post/102jisu/germanys-legal-debate-on-criminal-liability-for-misuse-of-deepfakes-navigating (last access: 29.05.2025).

13. Kop, M., EU Artificial Intelligence Act: The European Approach to AI, in Transatlantic Technology Law Forum (red.), Transatlantic Antitrust and IPR Developments, Stanford Law School, 2021.

14. Łabuz, M., Deepfakes and the Artificial Intelligence Act—An important signal or a missed opportunity?, Policy & Internet, V. 16, N. 4, 2024. p. 1–18.

15. Regulation (EU) 2016/679, General Data Protection Regulation, 2016.

16. Regulation (EU) 2016/679, General Data Protection Regulation, 2016, Recital 27.

17. Regulation (EU) 2024/1689, Artificial Intelligence Act, 2024.

18. Regulation (EU) 2024/1689, Artificial Intelligence Act, Annex III: High-Risk AI Systems Referred to in Article 6(2), 2024.

19. Regulation (EU) 2024/1689, Artificial Intelligence Act, Art. 7, 2024.

20. The Judgment 1234/05 of the European Court of Human Rights, Reklos and Davourlis v. Greece, 2009.

21. The Judgment 26839/05 of the European Court of Human Rights, Kennedy v. the United Kingdom, 2010.

22. The Judgment 2872/02 of the European Court of Human Rights, K.U. v. Finland, 2008.

23. The Judgment 44647/98 of the European Court of Human Rights, Peck v. the United Kingdom, 2003.

24. The Judgment 64569/09 of the European Court of Human Rights, Delfi AS v. Estonia, 2015.

25. The Judgment 65518/01 of the European Court of Human Rights, Salov v. Ukraine, 2005.

26. The Law of the Republic of France on Securing and Regulating the Digital Space, 2016 (in French / *Loi n° 2016-1321 du 7 octobre 2016 pour une République numérique*)

27. The Law of the Republic of Italy on Copyright, 1941 (in Italian / *Legge 22 aprile 1941, n. 633 – Protezione del diritto d'autore e di altri diritti connessi al suo esercizio*)

28. The Second Additional Protocol to the Budapest Convention on Cybercrime, 2022.

29. The U.S. Copyright Office, Copyright and Artificial Intelligence

URL: https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-1-Digital-Replicas-Report.pdf (last access: 29.05.2025).

30. Van der Sloot, B., Wagensveld, Y., Deepfakes: Regulatory Challenges for the Synthetic Society, Computer Law & Security Review, V. 46, N. 1, 2022. p. 1–15.

31. Van Huijstee, M., van Boheemen, P., Tackling deepfakes in European policy

URL:https://www.europarl.europa.eu/RegData/etudes/STUD/2021/690039/EPRS_STU(2021)690039_EN.pdf (last access: 29.05.2025).

Date of receipt of the article in the Editorial Office: 02.06.2025
Date of acceptance for publication: 25.12.2025