EUROPEAN ACADEMIC RESEARCH
Vol. XIII, Issue 2/ May 2025

Impact Factor: 3.4546 (UIF)
DRJI Value: 5.9 (B+)

ISSN 2286-4822
www.euacademic.org

AI, Deepfakes, and Sexual Harassment. Legal and Ethical Perspectives on a New Frontier of Sexual Violence

Ph.D (c) INA VELESHNJA
Department of Criminal Law
Faculty of Law, University of Tirana
Email: ina.veleshnja@fdut.edu.al

Abstract

Sexual harassment has evolved considerably over time, with technological advancements playing a critical role in its transformation. Technology has not only created new environments such as, for example the Internet that facilitates harassment, but has also introduced challenges in the identification and prosecution of offenders through the use of anonymity-enhancing tools. This article examines the convergence of sexual harassment and digital technology, analyzing the ways in which online platforms enable misconduct, the legal barriers to effective prosecution, and the ethical complexities inherent in regulating internet behavior. Recent developments in artificial intelligence (AI) governance and content moderation are critically assessed, with particular attention to the rise of non-consensual deepfake pornography and other AI-facilitated abuses. The article concludes by proposing policy recommendations, including the implementation of AI watermarking technologies, the development of specialized legal frameworks, and the strengthening of enforcement mechanisms to better safeguard victims in the digital era.

Key words: sexual harassment, deepfakes, internet, AI

INTRODUCTION

Sexual harassment is recognized in numerous national legal frameworks as a form of sexual offense involving the imposition of unwelcome sexual advances by a perpetrator upon one or more victims, either simultaneously or over time. It is crucial to emphasize that victims do not consent to these advances and may express their discomfort or refusal either verbally or nonverbally, whether directly to the perpetrator or to others within their social environment. Given the wide range of behaviors that can constitute sexual harassment, accurately identifying and categorizing such acts can often present significant challenges.

One example comes from the Albanian Criminal Code, where sexual harassment is stipulated in Article 108/a and defines sexual harassment as the "*Commitment of actions of sexual nature which infringe the dignity of a person by any means or forms, by creating a threatening, hostile, degrading, humiliating or offensive environment…".* (Criminal Code of the Republic of Albania 2024). Within the primary criminal legislation of Albania, particular emphasis is placed on the *contextual environment* in which the offense transpires. According to the legal definition, the perpetrator engages in a series of sexually motivated acts intended to undermine the dignity of the victim by any means or in any form, thereby fostering a threatening,

hostile, degrading, humiliating, or offensive atmosphere. The establishment of such an environment is a fundamental component of the Albanian legal framework regarding sexual harassment. In judicial proceedings, it is incumbent upon the prosecution to establish, through admissible evidence, that the perpetrator's conduct materially contributed to the creation of the harmful environment, thereby satisfying the elements of criminal liability as prescribed by the Albanian Criminal Code.

It is important to note that, under the same legal framework, if the perpetrator's conduct escalates to the point of non-consensual sexual intercourse, Article 108/a of the Albanian Criminal Code ceases to apply. In such instances, the act of sexual harassment is considered a precursor to sexual assault, and the perpetrator is prosecuted under other, more severe offenses outlined within the Criminal Code. Although the Albanian criminal framework does not explicitly reference the use of technology as a modus operandi for the commission of sexual harassment, the broad scope of its definition permits its inclusion as a means through which such acts may be perpetrated. With technology becoming increasingly embedded in everyday life, offenders are exploiting digital tools to further their criminal objectives.

Sexual harassment may also manifest in non-verbal forms. This type of harassment encompasses a range of behaviors wherein the perpetrator's actions, rather than verbal statements, convey inappropriate sexual intentions. Although the specific acts constituting non-verbal harassment are numerous and difficult to exhaustively enumerate, some of the most recognizable examples include prolonged staring, winking, blowing kisses, and following the targeted individual (What is Sexual Harassment?, 2024). A significant contemporary challenge in addressing sexual harassment lies in its intersection with technology. The proliferation of digital platforms has facilitated the commission of such offenses, enabling perpetrators to reach victims across geographical boundaries with relative ease. Consequently, individuals worldwide, regardless of their relationship to the offender, are increasingly vulnerable to these forms of abuse.

SYNTHETIC MEDIA AND THE DEEPFAKE THREAT

A pivotal development within the digital revolution has been the integration of artificial intelligence (AI) into various facets of daily life. While AI has facilitated numerous technological advancements, it has also been exploited to inflict harm. One particularly concerning misuse involves the application of deepfake technology to produce non-consensual pornographic material. The emergence of this phenomenon can be traced back to 2019, marked by the release and widespread accessibility of a program known as "DeepNude." This software enabled users to upload photographs of individuals and, through AI processing, automatically generate manipulated images depicting the subjects in a state of undress, thereby producing pornographic content without the consent of the individuals portrayed. (Bateyko and McCammon 2019). The program utilized machine learning to create very realistic pornographic images of other people. The likeness in the images was remarkably accurate.

Deepfake technology, which utilizes artificial intelligence to generate hyper-realistic yet fabricated images and videos, has profoundly reshaped the landscape of sexual harassment. By digitally superimposing individuals' faces onto explicit content without their consent, perpetrators are able to produce convincing but entirely fictitious depictions, resulting in significant psychological trauma and reputational damage for the victims. Recent incidents have underscored the pervasive nature of this threat. Notably, reality television personality Vicky Pattison participated in a Channel 4

documentary in which a deepfake pornographic video of her was created to raise public awareness about image-based abuse. This initiative ignited widespread discussions concerning the ethical ramifications of deepfake technology and the severe psychological toll it imposes on affected individuals. (Cross 2025). The legal framework addressing deepfake pornography is undergoing significant development. In the United Kingdom, legislative efforts have been initiated to criminalize the creation of non-consensual deepfake imagery. Specifically, an amendment to the Online Safety Bill seeks to render the production of such content illegal, signaling an increasing recognition of the urgent need for legal reforms to confront this emerging form of abuse. (Governmnet Bill 2024)

It is even more concerning that sometimes the images created in this manner, were of underage children thus effectively producing child pornography. The usage of deepfakes didn't stop with "DeepNude", even though this program is no longer active. Unfortunately, there are millions of similar programs on the Internet that essentially do the same thing. Since most of the people behind these programs are veiled in anonymity, it's very hard for law enforcement agencies to prosecute the responsible people. A study by the Internet Watch Foundation (IWF) identified over 20,000 AI-generated images on a dark web forum within a single month, with more than 3,000 depicting criminal child sexual abuse activities. This underscores the urgency for comprehensive strategies to combat the creation and dissemination of such material. (Internet Watch Foundation IWF 2024). This lack of accountability for the creator of these images or videos, affects their victims directly, creating problems in their family or professional lives. There are a number of issues when it comes to deepfakes, stemming into two main branches the technical and legal vulnerabilities.

Technical dimensions

The creation of deepfakes typically involves the collection of substantial volumes of images, videos, and audio recordings of the individual to be replicated. Artificial intelligence systems then analyze and map the person's distinct features, subsequently applying them to generate new synthetic content. In contexts such as de-aging actors for entertainment purposes or producing humorous filters on social media, deepfake technology represents a cutting-edge tool capable of achieving highly realistic and innovative results. (Stouffer 2023)

However, the core concern with deepfake technology lies in its potential for malicious use. In cases involving the creation of illegal pornographic material, artificial intelligence systems are often trained on vast datasets comprising millions of explicit images and videos in order to replicate human anatomy and facial features with high precision. As deepfake technology continues to evolve, so too do the challenges associated with identifying, investigating, and prosecuting those who misuse it. Offenders frequently employ multiple layers of obfuscation technologies, such as encryption, VPNs, and anonymization tools in order to conceal their identities and evade detection. Encryption not only shields the perpetrator's identity but also obscures their geographical location and any additional illicit activity carried out across digital platforms.

A study done by (Security Hero 2023), concludes that deepfake pornography makes up 98% of deepfake video content online. 99% of deepfake pornographic videos target women and 53% of the individuals represented in these deepfake videos, are entertainers from South Korea and lastly 7 out of 10 pornographic websites, hosts deepfake pornographic videos. The data is quite concerning, especially if we take under

consideration, who the main target of deepfake pornography is. It becomes even more concerning when the people represented in these types of videos are underage, thus effectively constituting into child pornography.

Although the victim is neither involved in the creation of the deepfake content nor often aware of its existence, the potential psychological and reputational harms inflicted upon them must not be overlooked. With modern technological tools, perpetrators no longer require physical proximity to harass their victims; harassment can now be conducted remotely, often from the comfort of the offender's own residence, while employing sophisticated methods to conceal their identities. This anonymity significantly complicates the efforts of law enforcement agencies to investigate and prosecute such crimes effectively.

Moreover, a critical challenge arises in the realm of internet regulation, posing both technical and legal dilemmas. A persistent question concerns the extent to which online spaces should be regulated. Historically, the internet was perceived as a domain beyond the reach of traditional legal frameworks, offering users near-total freedom to act without oversight. While such a perception may have been partially justified during the early development of the Internet, it has become increasingly evident that, given the vast and transformative capabilities of digital platforms, cyberspace cannot remain outside the scope of law and governance. Effective regulation is necessary to balance individual freedoms with the protection of fundamental rights and the maintenance of public order in the digital environment.

Legal challenges

In addressing the legal dilemmas arising from this discussion, two primary hypotheses can be identified. The first concerns the degree of control that should be exerted over the Internet. If there is a consensus that moderation is necessary to filter illegal content, important questions arise regarding the appropriate limits of such moderation and its potential repercussions on freedom of expression. Maintaining the delicate balance between protecting free speech and implementing effective Internet regulation remains an extraordinarily complex and contentious challenge. Numerous cases worldwide have involved individuals using artificial intelligence to generate provocative or controversial imagery, subsequently defending their actions under the pretext of exercising free speech. However, the creation and dissemination of such content online can have significant consequences, foremost among them the risk of causing harm to others, particularly to vulnerable populations such as children potentially even resulting in long-term psychological damage.

On the other hand, heavy moderation of the interspace can produce other issues such as limiting free speech and being used as an oppressive tool in countries that adopt high levels of speech restrictions. Thus, it is imperative to find the right balance between the two, to not only protect the rights and interests on users online but also to provide for a platform that won't censor free thought and speech. Nonetheless, content moderation is necessary, and various countries around the world are working to create clear guidelines, especially with the introduction of AI and its impact in nonconsensual pornography. There have been instances where specific countries have prohibited the usage of AI completely. The prohibition of AI altogether, although considered a positive step in protecting the rights of victims that suffered damages, in my opinion is far the best solution in a world that is spearheading into the digital era. While full scale AI regulations are uncommon, collective efforts around the world are being made in order to combat misinformation, sexual harassment done through AI and

more. One clear example on AI content regulation can be observed in China. The state authorities have proposed new guidelines on AI content, aiming to protect rights of its citizens and legal persons in the country. The measures proposed by the "Cyberspace Administration of China" aim to make AI generated content easily verifiable and explicitly labeled. (Chu and Chen 2024). Clear labels, that the content is AI generated, should be in compliance with the law, regardless if the material is being downloaded or exported. If a provider fails to adhere with these security measures, severe penalties will be incurred on a case-by-case system. The strategy is currently undergoing the public feedback process until mid-November 2024, and its effects are set to begin in 2024 too.

Japan has also undertaken significant steps in AI content moderation. More specifically Japan follows a set of principles on AI, that must ensure that basic human rights should not be infringed, protection of privacy, fairness, transparency and accountability should be clearly implemented. (Nayak 2024). At the moment of writing of this article there are a number of laws in Japan that involve the area of AI such as the "Personal Information Protection Act", that highlights the risks of artificial intelligence especially in data protection. Laws protecting personal data, ensuring for a parallel implementation of strict regulation on AI content in Japan, is being considered although the Japanese government hasn't come up with a clear AI strategy, like the case of China. (Nakazaki 2024)

While countries such as China and Japan have recognized the risks associated with artificial intelligence and have introduced proposals for regulatory frameworks, this approach has not been universally adopted. In Albania, for instance, there remains a significant absence of legal guidelines specifically addressing AI-generated content, highlighting a substantial regulatory gap in the national legal framework.Although a "*Strategy for the National Cybernetics Security 2020-2025*" (AKSK 2024 2020-2025) Although the strategy exists, topics such as artificial intelligence and deepfakes are entirely absent from its scope. This omission is understandable, given that the strategy was developed in 2020, at a time when experts could not have fully anticipated the rapid emergence of AI technologies and the profound impact that deepfakes would soon exert on society.

Other European countries have followed through with the legal discourse concerning AI in various ways. EU member countries have created a number of strategies and acts on the benefits and risks of AI in our lives. Although AI has the potential to fundamentally change our lives, the technology should be sustainable and centered around humans. (EUR-Lex 2024). The EU act on artificial intelligence divides and highlights that AI can have various levels of threats to humans. High-risk AI are considered, those technologies that pose a threat to law enforcement, generative AI and more.

In the United States there is a lack of federal regulation concerning AI, leaving it into the hands of the individual states and industry self-regulation. (WhiteCase 2024). This fragmentation within legislative frameworks presents significant challenges, particularly in ensuring comprehensive and effective legal protection for victims of the criminal misuse of artificial intelligence. A further key distinction between the United States' and the European Union's approaches to AI regulation lies in the EU's establishment of clear risk categorization guidelines. Under the EU framework, member states are mandated to strictly regulate and assess identified risks. Additionally, the Artificial Intelligence Act provides for the creation of the AI Board, an entity tasked with overseeing and facilitating the uniform

implementation of the Act across the European Union. (European Commission 2024). It is clear that the EU legal framework and institutional response is much more comprehensive and prepared as compared to the US's legal preparedness.

LEGAL AND REGULATORY TRENDS ADDRESSING AI-GENERATED SEXUAL HARASSMENT

The right of having a personal and private life, is considered a basic and fundamental human right. Provisioned on Article 8 of the European Convention of Human Rights, it states that everyone has the right to respect his private, family life, home and correspondence. In the digital era, the private family life has never been more attacked. Now with the introduction of generative AI and especially deepfakes, an offender can generate illegal and offensive content of anyone.

This type of sexual harassment that takes the form of nonconsensual pornography, is on the rise and the Internet is the perfect place for perpetrators to hide and disseminate these materials. Cases of nonconsensual or revenge pornography examined by the European Court of Human Rights, under the light of Article 8 are many, for example (Volodina v Russia n.d.) or (Ismayilova v Azerbaijan 2019) just to name a few. In both of these cases, the applicants were harassed with intimate images or videos of them, that were shared without their consent to the public, by people they had had a previous relationship.

Even though, it is abundantly clear that the perpetrators acted on malice, wishing to cause significant damage to the reputation and character of their victims and being quite successful, things have only gotten worse with the implications of AI and deepfakes. There have been numerous cases of famous entertainers around the world that have been targeted and AI-generated pornographic content was produced. Taylor Swift (Gibson 2024) and Megan Thee-Stallion (Zhou 2024), were just two of the best-known celebrities to suffer the consequences of AI generated pornography.

It is quite accurate to assume that legislative calls are needed to regulate the usage of AI in generating pornographic images or videos without the consent of the party depicted in them. It should be noted once more that the perpetrators responsible for such acts, usually act out of malice, wanting to slander and attack the personality and character of the person depicted. Typically, women are the primary target of AI pornographic deepfakes, however, no one is really immune to this form of criminality. The Congress of the USA, has acknowledged the risk deepfakes pose for society in general and especially deepfakes utilized for this purpose. In the US there are no federal laws regulating AI in general and often the laws on the state level are not enough. (Sarnoff 2024). Some US states that provide for legal regulations on generative AI pornography, include New York, California, Florida etc.

With regard to EU member states, as previously noted, the European Union has adopted the Artificial Intelligence Act, which regulates the use of AI across member states and addresses the associated legal implications. Nevertheless, individual countries remain in the early stages of developing comprehensive legislation specifically addressing the use of generative AI for the purposes of sexual harassment. To ensure full legal preparedness, AI must not only be recognized as a potential threat, but its use in creating pornographic images or videos without the consent of the individuals depicted should be explicitly classified as a criminal offense. While revenge pornography is already criminalized in the legal frameworks of several EU countries, the evolving nature of AI-driven offenses necessitates amendments to existing laws to

explicitly encompass the production of non-consensual deepfake pornography. Legal provisions should be particularly stringent in cases involving minors; the creation of deepfake pornographic content featuring minors constitutes child sexual abuse material and must be addressed promptly and rigorously by the relevant law enforcement authorities.

Lastly, regarding the case of Albania, the criminal legal framework lacks any specific provisions addressing these forms of offenses. The Albanian Criminal Code does not regulate non-consensual pornography, let alone the emerging issue of AI-generated non-consensual pornography. The country's legislative framework in this area is outdated and has repeatedly failed to adequately protect the rights of victims. Nevertheless, some positive steps have been taken, including growing acknowledgment by legislators and other key stakeholders of this new form of criminality. Efforts are currently underway to amend and modernize the existing Criminal Code to better address these challenges.

CONCLUSIONS

The intersection of technology and sexual harassment has transformed both the scope and methods of these offenses, making them more pervasive and difficult to regulate. The rise of AI-driven tools, deepfake pornography, and online anonymity has allowed perpetrators to exploit digital platforms, often evading detection and legal consequences. These technological advancements have widened the scale of victimization, making sexual harassment a borderless crime.

One of the most concerning developments is the misuse of deepfake technology, which has led to the creation of non-consensual pornographic content. This is particularly dangerous when underage victims are involved, effectively constituting child sexual exploitation material. While some countries, such as the UK and certain U.S. states, have moved to criminalize deepfake pornography, others, including Albania and many developing nations, lack specific legal frameworks, leaving victims without adequate protection or recourse.

Addressing these challenges requires a multifaceted and coordinated response. First, comprehensive legal frameworks are needed at national and international levels to classify AI-generated non-consensual pornography as a distinct criminal offense. The EU AI Act serves as a proactive model, offering a risk-based approach to AI regulation, while the U.S. remains fragmented, relying on state laws and industry-led policies that lack uniformity. Bridging these regulatory gaps is crucial for effective enforcement. Beyond legislation, technological solutions play a key role. AI watermarking, content authentication systems, and enhanced tracking mechanisms must be developed to prevent the spread of non-consensual deepfakes and to hold perpetrators accountable. Additionally, law enforcement agencies must be better equipped to trace, identify, and prosecute offenders, despite challenges posed by encryption and jurisdictional limitations.

At the societal level, raising public awareness about digital sexual harassment is essential. Education programs should focus on preventing victimization, promoting responsible AI usage, and empowering users with tools to report and remove harmful content. Social media platforms and tech companies must take greater responsibility in moderating content and implementing stricter policies against AI-generated abuse.

In conclusion, while technology has revolutionized the way we communicate and interact, it has also introduced new avenues for sexual harassment, privacy violations, and digital exploitation. Policymakers, legal professionals, and technology experts must work collaboratively to implement effective solutions that strike a balance between innovation and the protection of fundamental human rights. The fight against AI-facilitated harassment requires global cooperation, proactive regulation, and ethical technology development to create a safer digital environment for all.

BIBLIOGRAPHY

1. AKSK 2024. 2020-2025. https://aksk.gov.al/wp-content/uploads/2024/01/Permbledhje-Plani-i-Veprimit-2024-2025-shqip-1.pdf.
2. Bateyko, Dan, and Muira McCammon. 2019. "Columbia Journalism Review." September 11. https://www.cjr.org/watchdog/deepnude-ai-synthetic-media-ethics.php.
3. Chu, Jonathan, and Kelly Chen. 2024. "China proposes new regulations on AI-generated content labelling." November 14. https://cms-lawnow.com/en/ealerts/2024/09/china-proposes-new-regulations-on-ai-generated-content-labelling.
4. 2024. "Criminal Code of the Republic of Albania ." October 17. https://www.warnathgroup.com/wp-content/uploads/2017/11/Albania_CC_1995_am2015_en.pdf.
5. Cross, Felicity. 2025. *DEEPFAKE DANGER Vicky Pattison shares deepfake porn clip of herself as she warns of dangers on C4 doc.* January . https://www.thesun.ie/tv/14568513/vicky-pattison-deepfake-channel-4-documentary/?utm_source=chatgpt.com.
6. EUR-Lex. 2024. *Regulation (EU) 2024/1689* . https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689.
7. European Commission . 2024. https://digital-strategy.ec.europa.eu/en/policies/ai-board.
8. Gibson, Kate. 2024. "CBSNews." https://www.cbsnews.com/news/taylor-swift-artificial-intellignence-ai-4chan/.
9. Governmnet Bill . 2024. *UK Parliament* . https://bills.parliament.uk/bills/3137.
10. Internet Watch Foundation IWF. 2024. *What has changed in AI CSAM landscape?* . July.
11. Ismayilova v Azerbaijan. 2019. https://hudoc.echr.coe.int/eng#{%22itemid%22:[%22001-188993%22]}.
12. Nakazaki, Takashi. 2024. *AI Governance Trends in Japan.* July 31. https://iclg.com/practice-areas/data-protection-laws-and-regulations/02-trends-in-ai-governance-in-japan.
13. Nayak, Rohit. 2024. "Japan's AI regulations: Agile governance in action." July 18. https://www.diligent.com/resources/blog/japan-ai-regulations.
14. Sarnoff, Leah. 2024. https://www.congress.gov/118/meeting/house/116778/documents/HHRG-118-JU03-20240202-SD002.pdf.
15. Security Hero. 2023. *State of Deepfakes 2023.* https://www.securityhero.io/state-of-deepfakes/#key-findings.
16. Stouffer, Clare. 2023. "What are deepfakes? How they work and how to spot them." November 1. https://us.norton.com/blog/emerging-threats/what-are-deepfakes.
17. Volodina v Russia . n.d. https://hudoc.echr.coe.int/fre#{%22itemid%22:[%22001-194321%22]}.
18. 2024. "What is Sexual Harassment? ." *Women Watch UN.* October 17. https://www.un.org/womenwatch/osagi/pdf/whatissh.pdf.
19. WhiteCase. 2024. https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-united-states?utm_source=chatgpt.com.
20. Zhou, Li. 2024. https://www.vox.com/culture/354798/megan-thee-stallion-deepfakes.