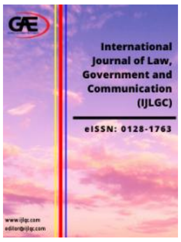
Volume 10 Issue 39 (March 2025) PP. 332-346
DOI 10.35631/IJLGC.1039023


International Journal of Law, Government and Communication (IJLGC)
eISSN: 0128-1763

INTERNATIONAL JOURNAL OF LAW, GOVERNMENT AND COMMUNICATION (IJLGC)
www.ijlgc.com

TRUTH DISTORTED: DEEPFAKES AND THE FIGHT FOR WOMEN'S RIGHTS

Jing Tian Chua[1*], Hasbollah Mat Saad[2]

[1] Faculty of Law, Multimedia University, Malaysia
Email: 1211101083@student.mmu.edu.my
[2] Faculty of Law, Multimedia University, Malaysia
Email: hasbollah.saad@mmu.edu.my
[*] Corresponding Author

Article Info:

Article history:
Received date: 12.02.2025
Revised date: 27.02.2025
Accepted date: 20.03.2025
Published date: 30.03.2025

To cite this document:

Jing, T. C., & Mat Saad, H. (2025). Truth Distorted: Deepfakes And The Fight For Women's Rights. *International Journal of Law, Government and Communication, 10* (39), 332-346.

DOI: 10.35631/IJLGC.1039023.

Abstract:

Artificial intelligence is developing rapidly, and deepfake technology is one of the inventions. This study explores in depth the impact of the misuse of deepfake technology on women's rights. The findings of the study show that deepfake technology has both beneficial and harmful sides, which severely impact on the society. The study aims to analyse existing laws and their ability to protect victims from deepfake abuse, especially women. Notably, deepfakes can produce misleading content, violate privacy and dignity, and lead to reputational damage which mainly targets women and harms their rights. The study identifies gaps in current legal frameworks, which lack focus on specific issues of deepfake abuse. The study also examined existing international legal standards and considered the best practices to propose effective solutions. These solutions aim to protect women from misuse of deepfakes and ensure that perpetrators are held accountable. Overall, the study aims to raise awareness of the challenges posed by the misuse of deepfake technology, which causes unacceptable risks to women, and the urgent need for legal reforms to safeguard women's rights on the global stage.

Keywords:

Artificial Intelligence, Deepfakes, Misuses, Women Rights, Legal Reforms.

Introduction

Background of the Study

Nowadays, technology is faster than the law. Artificial Intelligence ("AI") is the most disruptive area of innovation for the foreseeable future as it differs from ordinary computer algorithms. The rise of AI and machine learning has brought about revolutionary changes. Notably, deepfake technology is one of the most powerful but controversial tools among these advancements.

It is undeniable that deepfake technology benefits creative and commercial innovation. It enables the generation of realistic synthetic content using basic devices. It offers promising applications in areas, such as the film and entertainment industries (Adaobi, 2024). Additionally, deepfake technology is also useful in commercial areas, as it enables multilingual content and virtual try-on features, enhancing customer experiences and potentially increasing conversion rates (Radoslav, 2023). These applications highlight its potential for ethical and beneficial use, showcasing how it can positively contribute to society.

However, while the technology initially gained attention for its potential, its dark side of implications has overridden its benefits in recent years. Like all technologies, deepfake technology has a dual nature. Its high realism and accessibility raise concerns. While the technology itself is not inherently harmful, it has generated widespread discussion about its potential for abuse, which stems from human behaviour.

As deepfakes blur the lines between truth and falsehood, it has become a tool for illegal activities and unethical content creation. Deepfakes can convincingly manipulate video, audio, and image content to show individuals performing actions or saying things that they never actually did. This is further facilitated when social platforms provide abundant facial data that is easy to acquire. The low cost and accessibility of the technology allow almost anyone to use deepfakes. This leads to the rapid production of convincing synthetic content, increasing its potential for harm. The harm caused is frequently irreversible, with victims experiencing long-term damage to their reputations and well-being.

Consequently, deepfakes can undermine fundamental human rights by violating individuals' privacy, dignity, and consent. It jeopardises the security of the online environment and poses a growing threat to social stability. Its abuse has also brought significant harm to targeted vulnerable groups, especially women. Therefore, existing laws that merely address old computer crimes are no longer sufficient. New effective regulatory frameworks to govern ethical AI practices are crucial to avoid the negative implications of deepfakes in their use.

At this very moment, this paper will put a special focus on the concerns that deepfake abuse infringes on women's rights. In the context of humanitarian law, deepfakes expose women to gender-based violence, privacy violations, and reputational damage. Also, this study will examine the rights and dignity of women globally in the context of deepfake issues, while proposing practical and ethical solutions to mitigate the abuse of deepfakes.

Literature Review

This literature review will discuss three key points. Firstly, the issue of the misuse of deepfake technology; secondly, its implications for women's rights; and finally, the existing legal frameworks addressing these issues. This review will primarily focus on discussing the ethical implications of deepfake technology, which its misuse may severely harm women's autonomy and dignity. Also, the review will explore the need for comprehensive legislation to safeguard women's rights under the misuse of deepfakes.

Firstly, the thematic paper by the Platform of Independent Expert Mechanisms on Discrimination and Violence against Women (EDVAW), focused on addressing the digital aspects of violence against women. The article explicitly highlighted that technology has introduced new forms of abuse towards women, particularly in the form of deepfake pornography (Council of Europe, 2022).

The misuse of deepfake technology poses a significant threat to women. These techniques have been utilised for malicious purposes, such as the production of non-consensual pornographic content (Lock, 2023). This technology produces pornographic videos without the consent of the person depicted.

Those affected are mainly women, which perpetuates gender discrimination (Bass & Penning, 2023). The prevalence of deepfake pornography, which accounts for about out 96% of all deepfakes, highlights the disproportionate impact on women. This ruined their dignity to mere objects of sexual exploitation (Nnamdi et al., 2023).

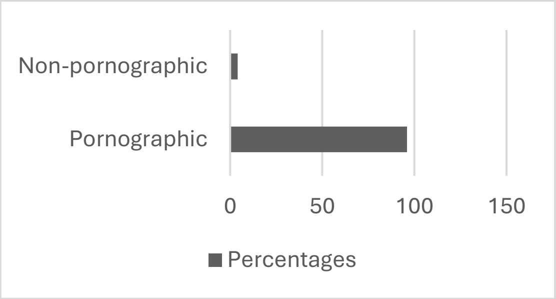


Figure 1: Content of deepfake videos online

Sources: (Deeptracelabs, 2019)

The ethical implications of deepfake technology go beyond the law. The violation of sexual privacy through deepfake pornography undermines the autonomy and dignity of women. It contributes to a culture of sexual humiliation and exploitation (Nour & Gelfand, 2021). Victims of fake pornography often face severe emotional distress, financial loss, and reputational damage (Nnamdi et al., 2023) The rapid distribution of deeply faked content through social media exacerbates the risks, as users can share and modify content with minimal supervision (Boucher, 2021).

Despite recognising these issues, it is found that existing legal frameworks often fail to provide adequate protection for victims. There is a clear need for comprehensive laws that address all aspects of deepfake technology. Current laws tend to focus on specific situations, such as revenge porn. These laws do not cover all aspects of malicious use (Quirk, 2021). This legislative gap allows for the continued utilisation of deepfake technology. It can blur the line between fact and fiction, leading to widespread misinformation and distrust of the media (Boucher, 2021).

The Concept Of Deepfake Technology

The first introduction of deepfake technology was to describe synthetic media in 2017. A user with the pseudonym "deepfakes", in an online platform named Reddit formed a subreddit. Then, the user began posting videos that used face-swapping technology to put celebrities' faces into existing pornographic videos (Somers, 2020).

Generally, the term "deepfake" is a combination of the words "deep learning" and "fake", which refers to an image generation technique based on AI technology. The technique relies on Generative Adversarial Networks (GANs), a deep learning model that uses two neural networks in opposition. This model was developed by Ian Goodfellow and his colleagues in 2014 (Giles, 2018). In its theory, one network will generate the image and the other will determine whether the image is real or not (Maithili, Sakshi, Vaishnavi & Sachin, 2024). After several time of generations, the final generated deepfake content can become so convincing that it is virtually indistinguishable from a real image or video, often deceiving the human eye.

Deepfakes continue to evolve alongside advancements in AI technology. In their early stages, deepfake content was relatively easy to identify. However, as the technology has matured, training data has been refined to create deepfakes that are increasingly difficult to detect. This data is used to generate video and image deepfakes in various ways, including face-swapping, attribute editing to alter specific features of a character, face reenactment to modify facial expressions, and full synthetic creation to design characters that do not exist in reality (Europol, 2022).

Today, with just a single photo or voice recording of a person available online, deepfake technology can manipulate facial features, simulate specific actions, or generate speech, achieving a lifelike effect that blurs the line between real and fake. This technology represents a significant advancement in synthetic media creation, offering unprecedented opportunities while posing substantial risks. Its diverse applications span across video, picture, audio, textual, social media, and real-time modalities, each presenting distinct capabilities and challenges (BasuMallick, 2022). These various types of deepfake technology underscore its transformative potential across multiple domains, making it increasingly difficult to be controlled.

Women's Rights at Stake

At first, deepfake technology rapidly gained prominence due to its use in creating illegal sexual content, which primarily targets women in the generated material (Gosse & Burkell, 2020). It is emphasised in a 2019 report that the majority of deepfake content consists of deepfake pornography videos (Deeptrace, 2019). Therefore, it clearly shows that if deepfake technology is abused, it might become another form of sexual abuse towards women.

Just like walking down the street in the real world, we need to be safe in the virtual world (United Nations Serbia, 2023). Women's rights must be protected both in real life and in the virtual world. The abuse of deepfakes poses a significant challenge to the realisation of these protections. This is because such abuse dehumanises women, treating them not as individuals but merely as sources of data for deepfake content. Perpetrators who exploit deepfakes to create non-consensual content objectify women, using them solely to fulfil their selfish and malicious desires.

Notably, women's rights are human rights (United Nations, 2024). Therefore, such abuse of deepfakes primarily infringes upon women's rights in several key aspects. These infringements have made the online world an increasingly hostile environment for women. Society must reflect on this troubling situation and take necessary action. When encountering deepfake content, particularly those created with malicious intent, individuals should not enable such abuse or remain passive and indulge in voyeurism. Instead, they must firmly oppose and boycott such deepfake pornographic content. Such harmful practices threaten societal well-being and undermine women's rights. Therefore, individuals must actively contribute to the advancement of modern civilisation by eliminating this abusive content to foster a healthier digital environment for women.

Rights To Privacy, Reputation And Dignity

Notably, Article 12 of the Universal Declaration of Human Rights 1948 provides that no one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor attacks upon his honour and reputation (United Nations, 1948). Everyone has the right to the protection of the law against such interference or attacks. This article can be interpreted into two limbs, firstly, everyone shall be subjected to the right to privacy, and secondly, everyone has the right to their honour and reputation.

First and foremost, the concepts of privacy and private life are often used interchangeably. This indicates Article 12 also means that everyone has the right to be protected from interference in their private life. The abuse of deepfake technology can invade people's privacy in ways they may not even realise.

In this cyber society, images of individuals are often easily obtained with a simple click on the internet, especially if they have shared their pictures online or in the public domain. However, sharing their own images online does not mean they are automatically entitled to public use. Individuals, groups, or institutions have the right to decide when, how, and to what extent such information they want to share with others (Westin, 1967). Therefore, in Westin's theory, people can choose what parts of their lives they want to keep private and what parts they are comfortable sharing with others. When women share moments from their daily lives on social media, this does not mean they consent to their images being used in deepfake content or made publicly accessible for malicious purposes. As such, women's right to privacy is infringed when their images are acquired, manipulated, and published in deepfake content without their consent.

Secondly, people who are deceived by deepfake content may believe that the woman engaged in actions or made statements she never did. This form of abuse also directly ruined women's reputations and dignity. If the victim's friends, family, or colleagues were to come across such content, they might misunderstandings her and these circumstances would lead to the act of damaging her reputation among those who know her.

Furthermore, women who are victims of deepfake content often find it impossible to prove their innocence due to the advanced nature of deepfake technology, which can easily deceive the human eye. This can result in them being unjustly victimised. One notable deepfake pornography case occurred at Seoul National University, where the perpetrators were accused of circulating illegal deepfake pornography. In this case, the perpetrators were suspected of using photographs of female victims' faces to create 419 sexually degrading deepfake materials, which they began distributing in 2020. The Seoul Central District Court sentenced the perpetrators to five years of imprisonment. However, even though the perpetrators are now in prison, the victims continue to suffer from the crime, as the deepfake content has already been spread to the public (The Straits Times, 2024). This harm will persist as long as the internet exists. This example demonstrates how the abuse of deepfake technology can destroy an individual's peaceful life by ruining their reputation and dignity.

Freedom Of Speech, Expression And Information

Article 19 of the Universal Declaration of Human Rights highlights the right to freedom of opinion and expression, which is protected under all relevant international human rights treaties. In the context of deepfake abuse, there are two contrasting views on this right (United Nations, 1948). Some oppose making the creation of deepfake content illegal, while others support such a law. Therefore, a detailed discussion of both perspectives is necessary.

Firstly, the American Civil Liberties Union (ACLU) and the Electronic Frontier Foundation (EFF) have expressed disagreement with laws prohibiting the creation of deepfake content, particularly involving politicians (Wang, 2019). They argue that such laws would infringe upon freedom of speech.

However, under Article 7 of the Universal Declaration of Human Rights (United Nations, 1948), it is emphasised that all humans are equal before the law and are entitled to equal protection of the law. Therefore, if these organisations consider it acceptable to create deepfakes of politicians to protect freedom of expression, a loophole may arise. It indicates that regulating deepfake content targeting women politicians might be considered to infringe on the freedom of speech.

On the other hand, another view argues that deepfakes do not qualify for protection under freedom of expression laws because their nature is inherently deceptive (Barber, 2023). It expressed that granting perpetrators the freedom to create deepfake content exploits others' freedom of expression. Every individual has the right to choose whether to express their opinions or remain silent. Deepfake content falsifies reality by attributing words or actions to someone who never expressed or performed them.

Thus, the second view offers a more legally sound and ethical outcome, as it avoids the loopholes inherent in the first perspective. This view also upholds the principle of protecting everyone's freedom of expression. Freedom of expression which encompasses acts, words, and information should not be protected indiscriminately. It must be assessed based on intent, and protection should only apply in the absence of malicious intent and those that do not harm public order. Therefore, the abuse of deepfakes to create pornography is obviously an act that stems from malicious intention. At its core, such abuse is a form of exploitation of women's rights to freedom of expression by slanting their actions and words.

Rights To Life, Liberty And Security Of Person

Article 3 of the Universal Declaration of Human Rights states that everyone has the right to life, liberty and security of person (United Nations, 1948). The issue of deepfake abuse, clearly shows that it results in the infringement of the right to security of women.

It is a generally accepted principle that most of the international human rights that apply in the "offline" world must also be shielded in the "online" world (Schmitt, 2017). Therefore, several laws that protect women's rights in the physical world need to be applied to the same standard as in the online world. General Recommendation No. 35 emphasises that states must protect victims of gender-based violence against women in their law, granting them access to justice and effective remedies. It also underscores sexual assault is a violation of women's rights to personal security (McQuigg, 2017). This clearly shows that sexual assault is an offence which infringes on the right to security of women.

In line with that, non-consensual deepfake pornography is clearly a sexual assault offence in the "online" world. Hence, the abuse of such technology to create pornography by using images of women without consent is a clear violation of women's right to security in their lives. Such gender-based violence against women by using new technology must also be prohibited and punished with the same standard as it is committed in the offline world. The abuse of deepfakes must also be banned to ensure that women can enjoy their right to a secure life.

Global Responses to Deepfake Technology

This section highlights how countries mitigate the risk of misuse of deepfakes by enacting new laws and policies or amending existing ones. Some of the laws and regulations are effective and align with technological advancement, while others have faced criticism for being flawed. This chapter provides an overview of the legal frameworks of several countries in dealing with the misuse of deepfakes, including the United Kingdom, Australia, the United States, and Singapore.

United Kingdom

In combating the problem of misuse of deepfakes, the United Kingdom has actively reacted by reforming its existing laws and introducing new laws. The main law in the United Kingdom which regulates deepfake technology is the Online Safety Bill 2023 (Ramluckan, 2024). This bill outlines comprehensive regulations for online platforms to control harmful content. In detail, the bill highlights that platforms are required by law to verify user identities. This is to ease the process of tracing the perpetrator who creates malicious deepfakes. Additionally, platforms must remove harmful content, or limit access to such content. Also, the bill imposes severe penalties for those creating misleading deepfakes, when the malicious intent of the perpetrators to cause harm is proven.

Moreover, the bill amended the existing Sexual Offences Act 2003, which focuses on the issue of non-consensual pornographic deepfakes. Section 13(4) of the bill stated that all harmful content targeted at adults must be restricted (Sexual Offences Act, 2003). Sometimes, the bill is criticised for its focus only on non-consensual pornography. It is criticised that the bill fails to address other harms caused by deepfakes. However, even though the bill only directs on the issue of non-consensual pornography, it is still a significant step in combating the misuse of deepfakes to protect women, who are the major victims of such issues.

Other than the bill, victims of deepfake misuse have the option to report incidents to the police. The official platform "Police.uk" serves as a channel for reporting such cases (Police.uk, 2024). Victims can submit details of the incident through this platform to initiate an investigation. The authorities then work to ensure the removal of harmful materials, particularly if they violate existing laws, such as those outlined in the Online Safety Bill 2023 (Online Safety Bill, 2023).

Also, law enforcement agencies use the information submitted through Police.uk to trace the creators of malicious deepfakes. The availability of Police.uk simplifies the process for victims to seek help and protection to report deepfake-related abuse.

Australia

In Australia, the government undertakes initiatives to combat the misuse of deepfake technology, aiming to prevent gender violence and sexual abuse in society. It is notable to refer to an Australian case, *eSafety Commissioner v Anthony Rotondo Aka Antonio Rotondo* (eSafety Commissioner v Rotondo, 2023), in this case, it is alleged that Mr Rotondo uploaded “intimate images” of several Australians to a website called MrDeepFakes.com. This website collects, hosts, and displays “deepfake” videos and images. The images in question were allegedly created and posted by Mr Rotondo and show people in intimate situations. These individuals, referred to as the “depicted persons”, are Australian public figures who did not give their consent for the creation of these deepfakes. The eSafety Commissioner issued a removal notice to Mr Rotondo, requiring him to take down the deepfake materials. However, Mr. Rotondo ignored the notice and was eventually charged with contempt of court.

This case raises an important legal consideration. Mr Rotondo was charged with failing to comply with a court order, but he was not directly penalised for misuse of the deepfake technology. This raises the question of whether the current legal framework fully addresses the misuse of deepfake technology. The Australian government has made significant efforts to regulate and prevent the misuse of deepfake technology. However, there may be gaps in the law in terms of directly penalising individuals who create and share harmful deepfake content.

Recently, the government introduced the Criminal Code Amendment (Deepfake Sexual Material) Bill 2024, which will amend the Criminal Code Act 1995 (Hamilton, 2024). The introduction of this bill effectively addresses the gap highlighted in the case mentioned above, by specifically criminalising the creation and distribution of non-consensual deepfake material.

This bill criminalises the creation of sexually explicit material by the other person without their consent. Also, the bill establishes aggravated offences for cases where a person creates deepfake content that is later shared without consent and is subject to three or more civil penalties under the Online Safety Act 2021 (Online Safety Act, 2021). In these cases, the maximum penalty increases to seven years imprisonment.

United States

The United States sought to reduce the harm caused by deepfakes, such as privacy violations, non-consensual pornography, fraud, and election interference through both state and federal legislation. These laws regulate the use of deepfake technology and hold offenders accountable.

Notably, seventeen states in the United States have passed laws specifically targeting online impersonation and the malicious use of deepfakes (NCSL, 2024). In illustrations, California’s law prohibits the creation or sharing of deepfakes meant to harm reputations or influence elections. Not only that, but Texas also criminalises the misuse of deepfakes, which impersonate others for fraudulent purposes or to cause harm. These state laws-imposed liability on the perpetrators, to make them bear the legal consequences of their malicious intent and actions (Tandy, 2024).

Moreover, Federal lawmakers have addressed this issue by introducing The Identifying Outputs of Generative Adversarial Networks Act, which emphasises the importance of establishing technologies that can identify altered media (Identifying Outputs of Generative Adversarial Networks Act, 2020). As a result, it can ensure that manipulated content can be distinguished from real content.

Moving on, there is the Deepfake Report Act of 2019 which imposes an obligation to the Science and Technology Directorate in the Department of Homeland Security. They need to report at specified intervals on the state of digital content forgery technology (Deepfake Report Act, 2019). This law ensures that the government can monitor the situation of misuse of deepfakes, to protect national security, public safety, and individual rights, by this means helping to develop better policies.

Additionally, defending each and every person from false appearances by keeping exploitation subject to the Deepfakes Accountability Act 2023, further requires every creator of deepfake content to put digital watermarks or disclaimers on the content generated by deepfakes (Deepfakes Accountability Act 2023). This ensures that ordinary people can recognise when content has been altered. The Disrupt Explicit Forged Images and Non-Consensual Edits Act of 2024 (Defiance Act) criminalises harmful uses of deepfake technology (Disrupt Explicit Forged Images and Non-Consensual Edits Act, 2024). This law defines offences involving deepfakes and ensures that offenders can be prosecuted accordingly.

South Korea

There were large-scale public demonstrations and international criticism that labelled South Korea as the “global deepfake capital” (McGlynn & Toparlak, 2024). The push for legal action began after mass demonstrations in 2018. These demonstrations highlighted the alarming extent of public voyeurism. To combat this issue, in 2020, South Korea criminalised the creation of sexual digital forgeries with the intent to distribute them. This earlier legislation was also a response to the exposure of extensive online communities engaged in creating and sharing such material (McGlynn & Toparlak, 2024).

Recently, South Korea has implemented comprehensive criminal laws to tackle sexual digital forgeries, particularly deepfake pornography. This action followed new legislation introduced. In October 2024, the Act on the Punishment of Sexual Crimes was revised to make the production of deepfake pornography a criminal offence, even if there is no intent to distribute it (Act on the Punishment of Sexual Crimes, 2024). The legislation was prompted by public outrage over the widespread creation, trading, and sharing of sexual digital forgeries. This included images of young schoolgirls shared across various Telegram channels (Se Eun, 2024).

The new law in South Korea criminalises not only the creation of sexual digital forgeries but also the possession, purchasing, storing, and viewing of such material. All activities related to deepfake sexual abuse are now prohibited under their criminal law. Furthermore, the state is now responsible for helping to remove illegal material and assisting victims. Additionally, new legislation allows investigative agencies to carry out covert investigations into digital sex crimes (Park, H, 2024).

The rapid introduction of this legislation reflects the government's response to a growing sense of crisis. The recent legal changes in South Korea are seen as a continuation of previous efforts to combat deepfake sexual abuse. These efforts had not effectively changed behaviours. Legal measures alone cannot secure societal change. They require accompanying societal support and changes in attitudes.

Recommendations

The misuse of deepfake technology presents significant challenges that require effective solutions. Several international standards can serve as a foundation for these solutions. The European Union Artificial Intelligence Act, published on November 6, 2024, provides valuable insights into addressing the risks associated with deepfakes (European Union Artificial Intelligence Act, 2024). Article 5 of the Act contains a list of prohibitions of AI services placed in the market, which include the use of subliminal techniques or deliberately manipulative or deceptive techniques that materially distort behaviour in a way that causes significant harm. In addition, Article 50(4) of the Act requires content to be labelled if it is generated from deepfake technology to ensure transparency. Furthermore, under the 2024 report published by the United Nations (Governing AI for Humanity, 2024), it is stated that robust verification and detection of deepfake content are important, together with prompt notice and take-down procedures to prevent cause harm to society.

Based on the analysis of these standards, three basic steps were identified, which may effectively address the misuse of deepfakes if implemented as guidelines for law enforcement bodies.

Firstly, countries must establish departments to realise effective deepfake detection mechanisms. This mechanism is essential to distinguish deepfake content from other lawful content. If any illegal deepfake is discovered, all liable parties must be identified. This includes the producers of deeply falsified content and the ISPs that host the platform. The liability of those who produce and distribute harmful content must be established. Not only that, but ISPs who manage the platform must also be held accountable, for failing to maintain and screen the online environment to be safe. Such accountability will also remind the content creators and internet service providers to be responsible for the content published to the public. Next, the law enforcement body must impose appropriate penalties on those who produce or provide harmful deepfake content. This can be achieved by the government enacting new laws or amend existing ones.

Also, international cooperation is necessary to effectively implement those steps mentioned. An “Online Security and Protection (Deep Fake Accountability) Treaty” could be introduced. This treaty could bind the Member States to establish norms in regulating the misuse of deepfakes. Notably, the treaty could make it mandatory for ISPs offering services related to deepfakes to register a licence under relevant regulations. Consequently, ISPs would be required to comply with the terms and conditions of the licences to maintain the security of their online platforms. There must be an organisation under the treaty that conducts regular audits and assessments to verify ISPs’ compliance with the established terms and conditions.

Furthermore, under such a treaty, all deepfake content is to be labelled and categorised appropriately to help users identify deepfake content. It can categorise the content in categories like education, entertainment, and deepfake-generated. In addition, the treaty could implement a system for creators to register identification numbers (ID) and verify the identity of deepfake content before publishing it, which would effectively ensure the traceability of such content.

Each piece of deepfake content would have its ID accompanied by the identity of the creator and the hosting platform. This measure will make it easier to track the origin of the content and hold creators accountable for their actions if such content causes harm to others.

Also, the treaty could promote international cooperation among Member States. The framework would facilitate the sharing of information and best practices related to in-depth counterfeit detection and regulation. In such instances, countries could cooperate to communicate in developing advanced counterfeit detection technologies and share resources for training law enforcement bodies.

Finally, the treaty should outline remedies for the victims. The remedies could include monetary compensation to redress the victim’s reputational damage or mental injury, or mandatory injunctions to remove the illegal content. Specific processes for victims to report abuse and seek redress must also be included in this treaty. This process would ensure that victims receive adequate assistance to ensure their rights are protected. By providing this support, the treaty will help to create a safer online environment for all people.

Conclusion

As artificial intelligence technology continues to advance, existing protections and regulations may fail to keep pace with their rapid development. Such developments make it increasingly tough to enforce investigations and regulatory efforts, especially in non-consensual deepfake pornography cases which stem from human sexual desires and interests that severely infringe on women's rights.

It is challenging to find a balance between the protection of women's rights and at the same time ensure the shield of freedom of expression. While it is necessary to stop the spread of harmful deepfake content, strict regulation may infringe on the right to freedom of expression. Also, most existing laws did not include deepfakes in the content. This oversight has led to gaps in the legal framework and inconsistencies in the enforcement of the law. In light of this, governments need to develop comprehensive legislation in response to such issues, which should be able to effectively address the offences associated with deepfake technology.

Therefore, to reduce the risks posed by deepfake technology, it requires the use of advanced technological solutions. Detection mechanisms that use artificial intelligence and machine learning to analyse digital media for signs of manipulation are important. In addition to technological efforts, public awareness and a sense of responsible media consumption are also fundamental to reducing the risk of deepfakes. Governments shall hold media literacy campaigns aimed at educating the public on how to identify reliable digital content. Such education reduces the possibility that individuals will be easily deceived by false information. Also, digital media platforms should take responsibility for fact-checking, screening, and labelling deeply falsified content. At the same time, they should actively report suspicious media to prevent the spread of deeply falsified content.

Furthermore, the implementation of comprehensive regulatory policies is essential to prevent the threat of deepfakes. It would be necessary to improve the adaptability of the legal framework to comply with future uses of AI technologies.

In conclusion, it is essential to balance technological development with the protection of women's rights in every country. Non-consensual deepfake pornography continues to cause significant harm to women worldwide. This issue needs to be immediately addressed, both now and in the future. Today, the rise of deepfake technology has turned this problem into a nightmare for women, making them feel unsafe in the online environment. Therefore, the use of deepfakes must be carefully regulated to ensure that women's rights are fully protected in this new era of technological advancement.

Acknowledgement

We extend our deepest gratitude to the Faculty of Law, Multimedia University, for providing the academic environment and resources that made this research possible. Our sincere thanks also go to the International Journal of Law, Government and Communication (IJLGC) for granting us the opportunity to publish our work and contribute to the scholarly discourse on this critical issue. We are profoundly indebted to our families and friends for their unwavering support, patience, and encouragement throughout this journey. Their belief in our efforts has been a constant source of motivation. Lastly, we acknowledge each other's contributions, as this collaborative effort was important for the completion of this paper.

References

Adaobi. (2024). *The role of AI in marketing personalization: A theoretical exploration of consumer engagement strategies*. Retrieved from: https://www.researchgate.net/publication/379393342_the_role_of_ai_in_marketing_personalization_a_theoretical_exploration_of_consumer_engagement_strategies

Act on the Punishment of Sexual Crimes, 2024. Retrieved from: https://elaw.klri.re.kr/eng_mobile/viewer.do?hseq=68812&type=sogan&key=9

Anne Dulka. (2023). The use of artificial intelligence in international human rights law. Stanford Law Review, 26 Stan. Tech. L. Rev. Retrieved from: https://law.stanford.edu/wp-content/uploads/2023/08/Publish_26-STLR-316-2023_The-Use-of-Artificial-Intelligence-in-International-Human-Rights-Law8655.pdf

Barber, A. (2023). Freedom of expression meets deepfakes. *Synthese, 202*(40), 1–17. Retrieved from: https://doi.org/10.1007/s11229-023-04303-2

BasuMallick. (2022). What is deepfake? Spiceworks. Retrieved from: https://www.spiceworks.com/it-security/cyber-risk-management/articles/what-is-deepfake/

Boucher, P. (2021)." What if deepfakes made us doubt everything we see and hear?" European Parliament. Retrieved from: https://www.europarl.europa.eu/RegData/etudes/ATAG/2021/690046/EPRS_ATA(2021)690046_EN.pdf

Deeptrace. (2019). The state of deepfakes: Landscape, threats, and impact. Deeptrace Labs. Retrieved from: https://regmedia.co.uk/2019/10/08/deepfake_report.pdf

Deepfake Report Act, 2019. Retrieved from: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.congress.gov/116/crpt/srpt93/CRPT-116srpt93.pdf

Deepfakes Accountability Act 2023. Retrieved from: https://legiscan.com/US/text/HB5586/id/2843792

Disrupt Explicit Forged Images and Non-Consensual Edits Act, 2024. Retrieved from: https://www.govinfo.gov/content/pkg/BILLS-118s3696is/html/BILLS-118s3696is.htm

European Union Artificial Intelligence Act, 2024. (2024). Retrieved from: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

Europol. (2022). Facing reality? Law enforcement and the challenge of deepfakes. Publications Office of the European Union. Retrieved from: https://www.europol.europa.eu/publications-events/publications/facing-reality-law-enforcement-and-challenge-of-deepfakes

Gosse, C., & Burkell, J. (2020). Politics and porn: How news media characterizes problems presented by deepfakes. Critical Studies in Media Communication, 37(5), 497–511. Retrieved from: https://doi.org/10.1080/15295036.2020.1832697

Governing AI for Humanity. (2024). Retrieved from: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf

Hamilton Janke Lawyers. (n.d.). Deepfake sexual material bill. Retrieved from Retrieved from: https://www.hamiltonjanke.com.au/deepfake-sexual-material-bill/

Identifying Outputs of Generative Adversarial Networks Act, 2020. Retrieved from: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.congress.gov/116/plaws/publ258/PLAW-116publ258.pdf

Kadam, M. M., Kate, S. S., Chavare, V., & Bhoite, S. (2024). Comparative analysis of deep learning techniques for deepfake detection: Evaluating threats and opportunities. Grenze International Journal of Engineering and Technology, 10(2). Retrieved from: https://thegrenze.com/pages/servej.php?fn=576_1.pdf&name=Comparative%20Analysis%20of%20Deep%20Learning%20Techniques%20forDeepfake%20Detection:%20Evaluating%20Threats%20andOpportunities&id=2815&association=GRENZE&journal=GIJET&year=2024&volume=10&issue=2

Lock, O. (2023). The legal issues surrounding deepfakes and AI content. Farrer & Co LLP. Retrieved from: https://www.farrer.co.uk/news-and insights/the-legal-issues-surrounding-deepfakes-and-ai-content/

Giles, M. (2018, February 21). The GANfather: The man who’s given machines the gift of imagination. MIT Technology Review. Retrieved from: https://www.technologyreview.com/2018/02/21/145289/the-ganfather-the-man-whos-given-machines-the-gift-of-imagination/

Park, H. (2024). *[ADRN Issue Briefing] South Korea’s rising deepfake sex crimes and recent legal responses*. EAI (Electronics and Telecommunications Research Institute). Retrieved from: https://eai.or.kr/new/en/project/view.asp?intSeq=22816&code=105

Maithili. M, Sakshi. S, Vaishnavi. C and Sachin. B. (2023). Comparative analysis of deep learning techniques for deepfake detection: Evaluating threats and opportunities. *ResearchGate*. Retrieved from: https://www.researchgate.net/publication/383063232_Comparative_Analysis_of_Deep_Learning_Techniques_for_Deepfake_Detection_Evaluating_Threats_and_Opportunities

McGlynn, C. M. S., & Toparlak, R. T. (2024). The new voyeurism: Criminalising the creation of “deepfake porn.” Journal of Law and Society. Retrieved from: https://ssrn.com/abstract=4894256

McQuigg, R. (2021). *The Istanbul Convention, domestic violence and human rights* (1st ed.). Routledge. Retrieved from: https://doi.org/10.4324/9781315652436

National Conference of State Legislatures. (2024). Deceptive audio or visual media (deepfakes) 2024 legislation. Retrieved from: https://www.ncsl.org/technology-and-communication/deceptive-audio-or-visual-media-deepfakes-2024-legislation

Nnamdi, N., Adeyemi, O., and Abegunde, B. (2023). An Appraisal of the Implications of Deep Fakes: The Need for Urgent International Legislations. Retrieved from: https://www.researchgate.net/publication/372611418_An_Appraisal_of_the_Implications_of_Deep_Fakes_The_Need_for_Ur gent_International_Legislations

Nour., N and Gelfand, J. (2021). Deepfakes: A Digital Transformation Leads to Misinformation. Conference Proceedings Insights and Issues that Challenge and Demonstrate the Role of GL. Retrieved from: http://www.greynet.org/images/GL2021_Nour_and_Gelfand_pp.55 65.pdf

NPR. (2024, September 6). South Korea deepfake. Retrieved from: https://www.npr.org/2024/09/06/nx-s1-5101891/south-korea-deepfake

Online Safety Bill 2023. Retrieved from: https://www.legislation.gov.uk/ukpga/2023/50

Park, S. (2024). South Korea's rising deepfake sex crimes. ADRN Issue Briefing. Retrieved from: https://eai.or.kr/avanplus/filedownload.asp?o_file=20241212134520255318530.pdf&uppath=/data/bbs/eng_issuebriefing/&u_file=Park_SouthKorea%E2%80%99sRisingDeepfakeSexCrimes_241212_ADRNIssueBriefing.pdf

Quirk, C. (2021). The High Stakes of Deepfakes: The Growing Necessity of Federal Legislation to Regulate This Rapidly Evolving Technology. Princeton Legal Journal. Retrieved from: https://legaljournal.princeton.edu/the-high-stakes-of-deepfakes-the-growing-necessity of-federal-legislation-to-regulate-this-rapidly-evolving-technology/

Radoslav. (2023). *Deepfake technology's impact on digital marketing*. Retrieved from: https://www.researchgate.net/publication/370830218_Deepfake_technology's_impact_on_digital_marketing

Ramluckan, T. (2024). Deepfakes: The legal implications. *ResearchGate*. Retrieved from: https://www.researchgate.net/publication/379221500_Deepfakes_The_Legal_Implications

Schmitt, M. N. (2017). *International human rights law*. In Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations (pp. 179–208). Cambridge University Press.

Sexual Offences Act 2003. Retrieved from: https://www.legislation.gov.uk/ukpga/2003/42/contents

Somers, M. (2020). Deepfakes, explained. MIT Sloan. Retrieved from: https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained

Tandy, B. (2024). Deepfakes: Identity misappropriation in the digital age. Belmont University College of Law Research Paper No. 2024-20. SSRN. Retrieved from: https://ssrn.com/abstract=5101133

The Straits Times. (2024). South Korea court jails man for 10 years over deepfake porn. Retrieved from: https://www.straitstimes.com/asia/east-asia/south-korea-court-jails-man-for-10-years-over-deepfake-porn

United Nations. (2023). Bodyright campaign continues to raise awareness on technology-facilitated gender-based violence. Retrieved from: https://serbia.un.org/en/233886-bodyright-campaign-continues-raise-awareness-technology-facilitated-gender-based-violence

United Nations. (n.d.). Women's rights are human rights. Retrieved from: https://www.ohchr.org/sites/default/files/Documents/Events/WHRD/WomenRightsAreHR.pdf

United Nations. (n.d.). The digital dimension of violence against women. Retrieved from: https://www.ohchr.org/sites/default/files/documents/hrbodies/cedaw/statements/2022-12-02/EDVAW-Platform-thematic-paper-on-the-digital-dimension-of-VAW_English.pdf

Wang, C. (2019, November 1). Deepfakes, revenge porn, and the impact on women. Forbes. Retrieved from: https://www.forbes.com/sites/chenxiwang/2019/11/01/deepfakes-revenge-porn-and-the-impact-on-women/

Westin AF. (1967). *Privacy and Freedom*. Atheneum, New York. Retrieved from: https://scholarlycommons.law.wlu.edu/wlulr/vol25/iss1/20/