

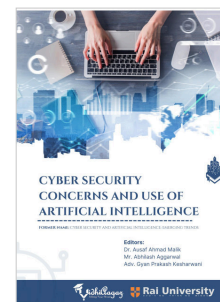
CHAPTER 18

Shifting Realities: Addressing the Legal and Ethical Challenges of Deep Fakes in Indian Jurisprudence

Shubhangi Pandey

in *Cyber Security Concerns and Use of Artificial Intelligence*, edited by Ausaf Ahmad Malik, Abhilash Aggarwal and Gyan Prakash Kesharwani (VidhiAagaz 2024), pp. 150–158

ISBN 978-81-947620-5-8



CITATIONS

BLUEBOOK (21ST ED.)

Shubhangi Pandey, *Shifting Realities: Addressing the Legal and Ethical Challenges of Deep Fakes in Indian Jurisprudence*, in *CYBER SECURITY CONCERNS AND USE OF ARTIFICIAL INTELLIGENCE* 150 (Ausaf Ahmad Malik, Abhilash Aggarwal & Gyan Prakash Kesharwani eds., VidhiAagaz 2024).

OSCOLA (4TH ED.)

Shubhangi Pandey, ‘Shifting Realities: Addressing the Legal and Ethical Challenges of Deep Fakes in Indian Jurisprudence’ in Ausaf Ahmad Malik, Abhilash Aggarwal and Gyan Prakash Kesharwani (eds), *Cyber Security Concerns and Use of Artificial Intelligence* (VidhiAagaz 2024) 150.

CHICAGO (17TH ED.)

Pandey, Shubhangi. “Shifting Realities: Addressing the Legal and Ethical Challenges of Deep Fakes in Indian Jurisprudence.” In *Cyber Security Concerns and Use of Artificial Intelligence*, edited by Ausaf Ahmad Malik, Abhilash Aggarwal and Gyan Prakash Kesharwani, 150–158. Maihar: VidhiAagaz, 2024. <https://vidhiaagaz.com/books/cyber-security-concerns-and-use-of-artificial-intelligence/chapter-18/>.

APA (7TH ED.)

Pandey, S. (2024). Shifting realities: Addressing the legal and ethical challenges of deep fakes in Indian jurisprudence. In A. A. Malik & A. Aggarwal & G. P. Kesharwani (Eds.), *Cyber security concerns and use of artificial intelligence* (pp. 150–158). VidhiAagaz. <https://vidhiaagaz.com/books/cyber-security-concerns-and-use-of-artificial-intelligence/chapter-18/>

SOURCE

<https://vidhiaagaz.com/books/cyber-security-concerns-and-use-of-artificial-intelligence/chapter-18/>

RIGHTS

Free to read online. © VidhiAagaz 2024. All rights reserved: no part of this publication may be copied, reproduced, adapted or transmitted in any form or by any means without the prior written permission of the publisher.

Citations are given as a guide; check them against the style manual you follow. This page was added by VidhiAagaz to the downloaded file. The chapter follows, with the page numbers of the printed book.

SHIFTING REALITIES: ADDRESSING THE LEGAL AND ETHICAL CHALLENGES OF DEEP FAKES IN INDIAN JURISPRUDENCE

Shubhangi Pandey²³

ABSTRACT

Artificial intelligence and machine learning have given birth to a very significant evil of the digital world known as Deepfakes. It is a form of manipulated media that allows for the creation of fake and fabricated videos, photos, and audios. Two major forms of deepfakes that are common in society are political and pornographic fakes. Political deepfakes defame the political leaders, which can lead to manipulation of the electoral results in the post-truth era, while pornographic deepfakes often consist of face-swapping, wherein a person's face is superimposed onto another body, creating the delusion of sexual activity. With the increase in the advent of penetration of social media, there is a growing potential for deepfakes to intrude into various professional and personal domains. It threatens the privacy and security concerns of the masses, including the celebrity figures of different domains. Deepfakes have also disproportionately affected women, particularly in the form of non-consensual pornography and harassment. The political dispensation across the globe is endeavouring to take action against it, with the likes of India banning TikTok while the U.S. exploring the privacy concerns of the aforesaid platform. Current Indian laws like the Information Technology Act, 2000, and Bhartiya Nyaya Sanhita, 2023, fail to address the rising deepfake issues imbibed in the country. This paper throws light on the need for a regulatory framework to protect infringement of privacy and manipulation of public opinion at large. This paper further explores legislative reforms and highlights the importance of raising public awareness and ethical measures to mitigate the risks posed by deepfakes. Through a targeted series of legislative and consultative approaches to promote digital awareness, India can work toward creating a safer digital environment.

Keywords: Deepfakes, privacy violations, Information Technology Act, 2000, Bhartiya Nyaya Sanhita, 2003, Legislative Reforms

INTRODUCTION

AI and machine learning have benefitted our society in several manners, including a wide spectrum of industries, namely entertainment, business, and healthcare. However, on the flip side, these artificial technologies have also paved the way for darker applications, including deep fakes. Deep fakes are created through artificial intelligence that in turn manipulates or fabricates images, videos, audios, etc. The main aim of the technology is to recreate gestures, voice, and face to make it seem real.

If we talk about the entertainment industry, this deep fake technology improves the effects through aiding and language dubbing for movies. It is utilised primarily to curb public opinion, cyber harassment, and violate the privacy of an individual. Major deep fakes are happening in politics and pornography, where they are exploited to defame political leaders or generate non-consensual explicit content. The penetration of social media with its current state of being less regulated has provided a breeding ground for deepfakes. It poses a threat to the security and privacy of every individual, including politicians, celebrities, etc. This paper throws light on the ethical, legal, and regulatory

²³ Assistant Professor at SAGE University, Indore, India.

challenges associated with the current legislation governing this evil impliedly, namely, the Information Technology Act, 2000, Bhartiya Nyaya Sanhita, 2023, the Copyright Act, 1957 etc.

The current situation is knocking doors for ethical guidelines, public awareness, and legislative reforms to combat the potential risks of deepfakes. The world has witnessed the infamous video of Donald Trump being insulted by Barack Obama, which resulted in the spread of misinformation quickly and rapidly into the public domain. Unfortunately, this evil has yet to garner the legal attention as required, in spite of the fact that the cases related to revenge porn and political manipulation have transcended mainstream society.

TYPES OF DEEP FAKES

- **Political Deepfakes and Manipulation of Public Opinion**

It is one of the significant tools used by individuals to malign the image of political opponents and manipulate election mandates. This technology is used to create realistic videos of politicians with false or damaging statements. This leads to profound influence upon voters, including social unrest and influencing voters.

The infamous deep fake video of Barack Obama was debunked eventually; however, its circulation highlighted the ease of spreading misinformation. This particularly alters the pathway of geopolitical events and political campaigns. The emotionally motivated storylines have displacing the authentic and manipulated content in the post-truth era, and it is quite difficult to distinguish between the truth and false narratives.

The institutions of democracy have been severely compromised with the diminishing trust in public figures and rapid spread of misinformation. Even the voter perceptions are getting affected by the deepened political polarisation and interference into the functioning of democracy. In comparatively young democracies, the political deepfakes lead to strong social and ethnic disharmony leading to the civil war and toppling of the government, while in the case of mature democracies, it is used to violate freedom of speech and expression, thereby suppressing the dissent of the common people. Chesney and Citron (2019) have long been termed Deepfakes as weapons of mass deception and can have a significant impact on global security, particularly during election cycles and periods of political unrest.

- **Pornographic deepfakes and gender-based harassment**

The non-consensual pornographic content is a major concern, and misuse of deepfake technology, which primarily targets women, It is usually done through the superimposition of a person's face into explicit sexual content without the knowledge or consent of the victim of deepfake. This leads to tarnishing of image, defamation, and misuse of copyright and personality rights associated with the victim, and the impossibility to control the widespread use of content uploaded online turns into a misery and tragic scenario for the same.

Often, celebrity women are being targeted through pornographic deepfakes. One of the incidents in the past involved actress Scarlett Johnson, wherein she expressed her helplessness to combat the ongoing generation of fake explicit videos related to her image and video. Adult content websites are filled with

the generation of such deep fakes and thereby lead to widespread abuse, contrary to their public opinion of being a niche tool.

Non-consensual porn not only leads to violation of the privacy and dignity of the victim but also facilitates a pathway of cyber harassment of the victim. Law authorities and victims face a major challenge in removing the content, often across international boundaries, as it makes it hard to identify the perpetrators. In the case of ordinary women, the situation becomes worse as they even lack the resources to choose the legal recourse.

In research done by Studies by Stroud, AI-generated pornography has been considered the primary source of objectification of women, thereby making it a major source of digital sexual violence. Women are often left with no legal recourse and adequate protection to address these deepfakes.

PUBLIC TRUST AND MEDIA INTEGRITY

One of the most common concerns of the deepfake technology is that it takes away public trust and establishes a sense of prejudice in the minds of the viewers. In this digital age, individuals are dependent largely upon the visual and auditory cues to assess the authenticity of the content. It also disrupts the credibility of media institutions.

The conventional way of journalism through relying upon video footage, audio, and witness testimonies has been compromised severely. It no longer guarantees the authenticity of the content, and it seems very difficult to do credible reporting by the media institutions. The risk of information nihilism emerges with the advancement of AI-generated content, wherein individuals discard the authenticity of any media content irrespective of its legitimacy. Rini (2020) terms this phenomenon an ‘epidemic backstop,’ which means our potential inability to form a shared understanding of inability that is imminent to democratic processes and electoral reforms.

Digital media has further increased the vulnerability of deepfakes as people are more glued to sensational, inflammatory, and shocking content. These leave an indelible stigma on public perception and misinformation.

DEEPAKES THROUGH LEGAL LENS

The legal framework in India is not sufficient enough to combat the penetration of deepfakes prevalent in our digital network. The legislation lacks explicit provisions to cover the wide ambit of deepfakes, as there is no evident legal provision to combat AI-generated content relevant to the industry.

- **Information Technology Act, 2000**

Various provisions of the Information Technology Act, 2000, including Section 67A, cover some aspects of deepfake-related content, which criminalises the distribution of obscene material. Similarly, Section 66E of the Act talks about the privacy of an individual violated by the unauthorised recording, publication, or transmission of an individual’s images in mass media using deepfake technology.

On the aforesaid lines, an individual who maliciously deceives or impersonates someone through communication devices or computer resources is combated by Section 66 D. The penalty and

imprisonment provisions are imposed upon the said violation. However, its outdated language takes away the penetration to cover the intricacies of deep fakes within its ambit.

- **Information Technology Rules 2021**

The Ministry of Electronics and Information Technology has released advisories to social media intermediaries to combat deep fakes. One of the advisories dated 7th Nov 2023 released by the aforementioned ministry to identify misinformation and deep fake comprised suggestions pertaining to devising protocols to not allow posting anything that is forbidden by Rule 3(1)(b) of the IT Rules. In case of any violation, there will be legal repercussions. Any content that is prohibited by the IT Rules needs to be made apparent to users.

Another advisory dated 26th December 2023 was released by MeitY to comply with existing IT Rules, which dealt with reasonable efforts that needed to be identifying misinformation and deep fakes. It also focused upon due diligence by social media intermediaries that violate user agreements, rules and regulations, and other relevant guidelines.

The said advisory prompted users to refrain from hosting more misinformation, content, or deep fakes. It also urged the platform to remove any content that is reported within 36 hours of the incident, and access to the content that is in violation of IT Rules 2021 needs to be disabled.

These advisories make it evident that the casual approach from the platforms and intermediaries could result in them being blocked. The platform under Rule 3(1)(b) of the IT Rules 2021 is supposed to comply with the guidelines pertaining to the obligations of the same.

- **Indian Penal Code, 1860**

The Indian Penal Code (applicable before Bhartiya Nyaya Sanhita, 2023) used to contain provisions like forgery (section 465), defamation (section 509), criminal defamation (section 499), and spreading hate on communal lines (section 153) that may be invoked for cybercrimes, including the offence of Deepfakes. However, the language and intention of these sections could not anticipate the evils of Deepfakes.

- **Copyright Act, 1957**

The Copyright Act, 1957, is comprehensive legislation relating to the protection of intellectual property rights of an individual. Deepfakes not only leads to cyber harassment but also violates the copyright of an individual associated with such content. It uses photos, videos, audios, etc. of a person to impose it upon some other content without the consent of the individual. Section 51 of the Act impliedly deals with such incidents as it comprises various offences, including copyright infringement. It tries to provide protection to the original work of the content, including music, films, and creative content.

- **Personality Rights**

Personality rights are associated with mostly celebrities or well-known individuals relating to certain acts or a gesture that makes them synonymous. In the recent past, actor Anil Kapoor obtained a protection order from Delhi High Court pertaining to the misuse of his name, likeness, voice, and

distinctive features from being misused by Deep Fake's technology. Similarly, Amitabh Bachchan was granted similar protection when his public image was used without his consent, resultant to the dangers posed by deepfake technology and the necessity of protecting people from its misuse.

However, there is an evident need to put a stop to the propagation of false and fabricated information about public figures and misinformation around them. The legislators need to find balance between the fundamental and intellectual property rights and legislation to curb deep fake technology.

LEGAL CHALLENGES RELATED TO DEEP FAKES

The widespread outrage of deep fake technology has outpaced legal frameworks around the globe, including India, and has posed some significant challenges pertaining to it. The legislation has been enacted to curb things done by humans largely through conventional methods.

- **Defamation and privacy violations**

This is one of the most significant legal challenges posed by deep fakes. It particularly maligns the image of the victim in the eyes of common people and leads to reputational damage. Section 499 of the erstwhile penal code comprises defamation, while Section 500 of the code deals with the punishment up to two years. In addition to that, the privacy standards as recognised by the Hon'ble Supreme Court in Justice K.S. Puttaswamy v. Union of India under Article 21 of the Constitution are getting violated by the deep fakes, as it does not particularly misuse images and videos used through this technology.

- **Manipulation of Evidence**

It poses audiovisual evidence as the fakes lead to fabricate evidence, wrongful accusations, or acquittals. Further, it jeopardises the forensic tools and is capable of detecting deep fake content. It is the inability to manipulate video footage or audio recordings that makes it difficult to test the authenticity of such evidence in court.

- **Forgery and Fraud**

Section 468 of the IPC, along with Section 66 D of the IT Act, 2000, fails to address the forgery and fraud done with the complexities of deep fakes that involve sophisticated AI techniques and are not easily detectable by the intention of the language used in the provision.

- **Consumer Protection and Data Privacy**

The deep fakes are even affected by the terms and conditions of the Zao app, as it allows perpetual rights over users's images. These images may be potentially damaging the privacy of consumers of the applications.

INTERNATIONAL LEGAL APPROACHES

There are various approaches taken by different countries in order to combat deep fakes and manipulation that should be witnessed closely in order to witness its effectiveness.

- **United States**

The United States has adopted the law against deep fakes both at the federal and state levels, such as California's AB 730 law, which takes care of the prevention of deep fakes within 60 days of an election to deceive voters. It seems to be one of the significant approaches to lessen the penetration of manipulation during the election time through AI-generated content.

Even the Deepfake Report Act of 2019 highlighted the issue of national security and privacy concerns in the U.S. However, these laws have not stood the test of time, as they are at the nascent stage of evolving to keep pace with rapidly advancing technology.

- **European Union**

In the European Union, the GDPR provides various remedies for victims of deep fakes. It provides the users with the right to request the removal of content that infringes on their privacy. However, the fakes are created anonymously or hosted on servers outside the EU, and the enforcement of such requests always remains a challenge.

- **China**

At the onset of the pandemic, China has come up with a stringent regulatory approach in 2020 pertaining to deep fakes. It puts misinformation under the check through the Cyberspace Administration of China's implemented rules requiring AI-generated content. Their approach and understanding of the regulation is a comparatively stricter way to address the issue, as it also points concerns about freedom of expression and state censorship.

ETHICAL CHALLENGES OF DEEP FAKES

Other than legal challenges, there are numerous ethical challenges related to deep fakes that have an immense effect on internet users. Deep fakes deal with various intricacies, including truth, consent, and privacy, which form the basis of this ethical dilemma. It gives us the moral compass to devise the development of AI and its use.

- **Consent and Autonomy**

It is one of the important ethical challenges that is considered mainstream in most of the deep fake incidents. Like in the case of deep fake pornography, there is an evident absence of consent and an undeniable harm. Even an individual's consent to the use of their likeness in one context can be misused in another context using deep fake technology.

The lack of consent of the victim in most of the deep fake incidents violates the principle of autonomy, wherein the image and likeness of the victim are used. It raises ethical concerns about the right to monetising deep fake content, distribution, and creation.

In the context of political figures, there is a debate around the ethical question of whether the public's right to satire or political commentary should outweigh an individual's right to control how their image is portrayed.

- **Privacy Violations**

Privacy is one of the other ethical challenges that are faced by the deep fakes. The major concern regarding the deep fakes is that they use the personal images and videos shared by the individuals on social media platforms. Even the CCTV cameras and the availability of personal data on public platforms make individuals vulnerable to such deep fakes violating the privacy norms. The laws related to privacy deal with the protection of traditional ways of privacy. However, it falls short when the realistic images or videos of individuals without their knowledge blur the boundaries of privacy.

- **Harm Principle and Responsibility**

Deep fakes pose a serious threat to the harm principle, according to which the individuals and organisations are supposed to not cause significant harm to them. It leads to emotional distress through activities like pornographic deepfakes or political misinformation, and the potential for harm is evident. Social media intermediaries like Facebook, Twitter, or YouTube even struggle to combat the ethical challenge of the harm principle as the volume of content makes it difficult to effectively police. Another important concern is to address the responsibility of such deep fakes between the individual who creates such deep fakes or such platform that hosts the same.

- **Manipulation and Deception**

This is another ethical challenge for deep fakes, as it manipulates not only individual perceptions but also collective beliefs. The deceiving nature of deep fakes, as it will become difficult for individuals to identify between reality and fake. It leads to the inability to make informed decisions based on reliable information.

- **Impact on Free Speech**

Deep fakes pose a threat to a balance between the freedom of speech and expression and harm caused to individuals, as often deep fakes are used for satirical and political commentary. One potential way to address this problem is to adopt a graded approach wherein different types of deep fakes are subject to different levels of scrutiny.

WAY FORWARD & CONCLUSION

The deep fake phenomenon is critical to the legal, ethical, and regulatory frameworks of the digital content. It possesses required potential that can take away with public trust bestowed upon them. Further, it leads to violations of privacy and cyber harassment. The specific legislation in India is supposed to match the pace with the immense development of AI technology. The age-old legislation needs to be revamped with the inclusion of the mandate of the AI.

On the same note, a singular set of legal responses would not be able to cater to the issue of deep fake. Even ethical guidelines need to be updated accordingly to promote the public good and individual rights. In addition to it, certain technological solutions need to be explored that can immediately block out the unauthorised use of images and videos through blockchain-based verification.

At the end, an integrated approach is necessary to combat the impact of deepfakes through working with various agencies of governments, tech companies, and civil society so that a secure and safer digital environment can be ensured in a digital era. Thus, adoption of a combination of solutions would help

in addressing the risks posed by deep fakes so that AI in the true sense can assist the human kind and not harm the same.

BIBLIOGRAPHY

1. The Constitution of India, 1950
2. Information Technology Act, 2000
3. Bhartiya Nyaya Sanhita, 2023
4. Copyright Act, 1957
5. What is deepfake technology? <https://www.techslang.com/what-is-deepfake-technology/>
6. Raina Davis, Chris Wiggins, Joan Donovan, Deepfakes, SPRING 2020 SERIES, Rev, 1, (2020) https://www.belfercenter.org/sites/default/files/files/publication/Deepfakes_2.pdf
7. JM Porup, How and why deepfake videos work — and what is at risk, CSO INDIA , <https://www.csoonline.com/article/3293002/deepfake-videos-how-and-why-they-work.html>
8. Westerlund, Mika. The Emergence of Deepfake Technology: A Review. Technology Innovation Management Review. 9.. https://www.researchgate.net/publication/337644519_The_Emergence_of_Deepfake_Technology_A_Review
9. Francis Navarro, Deepfake porn videos are now being used to publicly harass ordinary people, KIM KOMANDO.COM, <https://www.komando.com/security-privacy/deepfake-pornvideos-are-now-being-used-to-publicly-harass-ordinary-people/526877/>
10. Clare McGlynn, Erika Rackley & Ruth Houghton, Beyond 'Revenge Porn': The Continuum of Image-Based Sexual Abuse, Fem Leg Stud 25. Rev.25, 46 (2017). <https://doi.org/10.1007/s10691-017-9343-2>
11. Stroud, Clare. "AI-Generated Pornography and the Objectification of Women: A Feminist Perspective." Journal of Gender Studies, vol. 27, no. 3, 2018,
12. Press Infomation Bureau, MeitY issues advisory to all intermediaries to comply with existing IT rules,
13. <https://pib.gov.in/PressReleseDetailm.aspx?PRID=1990542&ref=static.internetfreedom.in>
14. Union Government issues advisory to social media intermediaries to identify misinformation and deepfakes, <https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1975445>
15. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code). Rules, 2021
16. Rashmi Senthilkumar, Defamation law in India, LEGALSERVICEINDIA.COM <http://www.legalserviceindia.com/legal/article-2224-defamation-law-in-india.html>

17. Parth Tyagi and Achyutam Bhatnagar, Deepfakes and the Indian legal landscape, INFORM.ORG, <https://inform.org/2020/07/03/deepfakes-and-the-indian-legal-landscape-parth-tyagi-and-achyutambhatnagar/>
18. Justice K.S. Puttaswamy v. Union of India, (2017) 10 SCC 1
19. Anil Kapoor vs Simply Life India & Ors CS(COMM) 652/2023
20. Grace Shao and Evelyn Chen, The Chinese face-swapping app that went viral is taking the danger of ‘deepfake’ to the masses, CNBC, <https://www.cnbc.com/2019/09/04/chinese-faceswapping-app-zao-takes-dangers-of-deepfake-to-the-masses.html>
21. Niles Christopher, We’ve Just Seen the First Use of Deepfakes in an Indian Election Campaign, VICE.COM [https://www.vice.com/en_in/article/jgedjb/the-first-use-of-deepfake s-in-indian-election-by-bjp](https://www.vice.com/en_in/article/jgedjb/the-first-use-of-deepfake-s-in-indian-election-by-bjp)
22. Human rights in the age of Artificial Intelligence, <https://www.accessnow.org/cms/assets/uploads/2018/11 /AI-and-Human-Rights.pdf>.
