

DEEPFAKE TECHNOLOGY: LEGAL CHALLENGES AND REGULATORY RESPONSES

AUTHOR – SREE PARVATHAVARTHINI SK* & DR. SANTOSH KUMAR TIWARI**

* LLM (STUDENT) AT SCHOOL OF LAW JUSTICE AND GOVERNANCE GAUTAM BUDDHA UNIVERSITY

** ASSISTANT PROFESSOR AT SCHOOL OF LAW JUSTICE AND GOVERNANCE GAUTAM BUDDHA UNIVERSITY

BEST CITATION – SREE PARVATHAVARTHINI SK & DR. SANTOSH KUMAR TIWARI, DEEPFAKE TECHNOLOGY: LEGAL CHALLENGES AND REGULATORY RESPONSES, *INDIAN JOURNAL OF LEGAL REVIEW (IJLR)*, 6 (9) OF 2026, PG. 774-787, APIS – 3920 – 0001 & ISSN – 2583-2344.

Abstract

The advent of artificial intelligence (ai) has changed the way that people communicate digitally, how we can create media, and how information is delivered. One of the greatest innovations in the use of ai is deepfake technology. Deepfake technology utilizes advanced machine-learning based algorithms to produce highly-realistic synthetic visual imagery (still images), video imagery, and audio-recordings. There are many benefits to using deepfakes including the potential for enhanced forms of entertainment, educational experiences, medical treatment options, and accessibility opportunities. However, there are a number of serious legal, ethical, and social concerns with respect to the use of deepfakes. Since their emergence as a viable tool for deception, individuals and organizations have employed them to further misinformation campaigns; distribute non-consensual pornographic images of individuals; influence elections; commit financial crimes such as stock market manipulation and account takeovers; steal identities; and engage in other forms of cyber-crime.

There are many potential problems that India will encounter when it attempts to regulate DeepFake technology. While India has implemented a number of laws (the Information Technology Act, 2000; the Digital Personal Data Protection Act, 2023; the Bharatiya Nyaya Sanhita, 2023) which partially address the issue of regulation of DeepFakes, none of these laws deal specifically with the problem of DeepFakes. The international community is beginning to develop legislation and regulatory schemes to mitigate the dangers posed by DeepFakes created using artificial intelligence.

This research evaluates the legal impact of deepfakes in terms of Indian constitutional law, criminal law, civil law and international law. It assesses whether current laws and regulations are sufficient to address this new form of deception and suggests an all-inclusive framework for regulating it in India. The analysis concludes that a balance is needed; one that will protect individual rights (of privacy, dignity and democracy) but also allow for creativity and free speech.

Keywords: Deepfake Technology, Artificial Intelligence, Privacy, Cybercrime, AI Governance, Data Protection, Digital Rights.

1. Introduction

Deepfake technology is a manifestation of the AI revolution which has transformed the world over the past two decades by giving us artificial intelligence. This is due to advancements in AI as

well as its sub disciplines (machine learning, neural networks) and computational processing capabilities. The end result has been the development of very advanced digital content

that can mimic human speech, appearance and behavior with an impressive degree of realism.

In combining "deep learning," as well as "fake," we get the name "deepfake." Deepfakes are synthetic media generated by way of artificial intelligence (AI) algorithms which either modify existing visual and/or audio content or create new visual and/or audio content to fake what someone is saying/did. Initially designed to entertain and advance academic studies on AI-generated visuals/audio/video, today's deepfakes have evolved into a power tool to influence public opinion, destroy personal reputation(s), and degrade confidence in digital information. As deep-fake technology has proliferated at a rate much faster than the law and regulation, there are now many serious challenges facing both government agencies and private entities who are tasked with addressing how to prevent the distribution of fake videos/photos/images/recordings.

Deepfakes have the power to cause harm that goes far beyond personal injury. Misinformation produced through deepfakes could disrupt democratic elections, promote violence, influence stock market activity and threaten national security. Synthetic media has the capability to blur lines of truth and deception and therefore is a significant threat to legal systems which depend on the reliability of evidence contained in digital files. The problem has received great attention in India after numerous high profile cases of deepfakes were created using well known actors, actresses, celebrities and politicians. When images of famous actress Rashmika Mandanna and many others were created through manipulated video, they began an India-wide discussion about whether current laws are capable of protecting citizens from AI-generated misinformation. This article will consider the various dimensions of law that may be affected by deepfakes and assess how best governments in both India and internationally can establish regulations to

mitigate the problems created by this technology.

Statement of Problem

Deepfake technology is developing so quickly that it has surpassed development in regulation and laws globally. Laws regarding privacy, defamation, obscenity, cybercrime, and copyright were developed prior to the creation of AI-created sophisticated synthetic video and audio. There are many issues that have yet to be resolved:

- 1.Does current legislation protect against harm caused by deepfakes?
- 2.To whom will responsibility fall when a creator distributes or makes available a deep fake through an online platform (e.g., Facebook, Twitter)?
- 3.Do Constitutional First Amendment Free Speech provisions apply to digitally created content?
- 4.Can privacy and dignity rights be preserved from the unauthorized reproduction of individuals in digital form?
- 5.Can there be adequate regulatory measures to foster innovation while protecting public safety?

Without clarity in these areas of the law, there is much confusion with regard to what action(s) law enforcement may take, how victims may recover damages for violations of their privacy, how courts may interpret the law, and how technology companies may operate.

2. Research Questions

The purpose of this study is to provide answers to the following research questions:

1. What exactly is deepfake technology and how does it work?
2. What legal/constitutional issues do deepfakes generate?
3. Are there enough laws in India to address the harm caused by deepfakes?
4. How did other countries respond to deepfakes?
5. What regulatory changes will be needed for the government to effectively regulate deepfake technology in India

3. Hypothesis

H1: Current indian legal framework does not provide a comprehensive legal regulation of the technology of Deep Fakes and protects individual rights against harm caused by AI generated Synthetic Media.

H2: An exclusive AI governance framework providing obligation for the Transparency on the use of AI Technology in the platforms; Obligation for the Platforms (Social Media etc.) to be accountable for the use of AI Technology; and Obligation to Protect Victims will enable effective regulation of DeepFakes and Protection of Constitutional Freedoms.

4. Objectives of the Study

1. To understand the technological basis of deep fake technology.
2. To assess the legal implications for creating and disseminating deep fakes.
3. To evaluate how well current Indian law is equipped to address the deep fake issue.
4. To compare the legal frameworks that other countries have in place.
5. To develop an overarching regulation plan for India.

5. Research Methodology

This study will use a Doctrinal Research Methodology which will be based on:

Primary Sources

- Constitution of India
- Information Technology Act, 2000
- Digital Personal Data Protection Act, 2023
- Bharatiya Nyaya Sanhita, 2023
- Judicial precedents
- International regulatory instruments

Secondary Sources

- Books

- Research articles
- Law review publications
- Government reports
- Policy papers

6. Literature Review

Deepfake technology has grown as an area of study in academia since its emergence roughly one decade ago. Robert Chesney and Danielle Citron point out that deepfakes represent a new risk to democratic systems, individual privacy, and national security. The authors further contend that synthetic media is undermining the general public's trust in authentic information; and creating previously unimaginable potential for manipulation.¹⁴⁶⁰

As stated by Danielle Citron, deepfake pornography is causing disproportionate harm to women and represents a severe infringement on their autonomy, privacy and dignity.¹⁴⁶¹

Bobby Chesney discusses how the "national security" aspects of deep fakes are especially dangerous (deep fakes can be used to spread disinformation and propaganda; they can be used for military deception and/or to cause diplomatic conflict).¹⁴⁶²

Cathy O'Neil has shown that while most algorithmic systems will likely reinforce or exaggerate already present social biases; they also have the capability to produce new forms of discrimination/injustice if used with inadequate oversight.¹⁴⁶³

Scholars within the Indian context have been increasingly examining the implications of AI-created content on constitutional rights, cybercrime regulations, and data protection law. The current literature is aware that present statutory frameworks lack adequate regulation and there is a disagreement about which should be the most suitable model of regulation.

¹⁴⁶⁰ Robert Chesney & Danielle Keats Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, 107 Calif. L. Rev. 1753 (2019).

¹⁴⁶¹ Danielle Keats Citron, Sexual Privacy, 128 Yale L.J. 1870 (2019).

¹⁴⁶² Robert Chesney & Danielle Keats Citron, Deep Fakes and the New Disinformation War, 98 Foreign Aff. 147 (2019).

¹⁴⁶³ Cathy O'Neil, *Weapons of Math Destruction* 3–15 (2016)

8. Research Gap

The number of areas of study for which research is needed to fully address the impact of deep-fakes includes:

1. Deep fake regulatory issues within Indian Law.
2. The interaction between Deep Fakes and Constitutional Rights in India.
3. Comparative Studies of Indian and International Legal Approaches that are Comprehensive.
4. Discussion regarding Victim Centered Remedies (compensation) Mechanisms.
5. Platform Accountability and Governance of AI.

This research aims to fill in this gap through a comprehensive review of how deepfakes are regulated at the Constitutional, Criminal, Civil and Comparative levels.

9. Understanding Deepfake Technology

A Deep Fake is Artificial Intelligence used for creating synthetic media (images/video/audio/behavior etc.) that has a realistic imitation of another person's visual identity, voice or behavioral patterns.

Mostly uses:

1. Deep Learning
 2. Neural Network Systems
 3. Machine Learning Algorithm Generative Adversarial Networks (GANs).
- A GAN is composed of Two Neural Networks; one called the Generator(creates synthesized data), the other called the Discriminator(attempts to recognize if produced data is real or created).

10. Types of Deepfakes

A. Face-Swap Deepfake : The face of a person is placed on the body of another person by digital means.

B. Voice-Cloned DeepFake : An AI system replicates the voice of an individual with minimal amounts of that individual's audio.

C. Synthetic Video Creation : Videos in their entirety can be generated from scratch without any contribution of the subject.

AI Generated Photos : Photographs generated

entirely by artificial intelligence programs that generate realistic images.

Legitimate Uses of Deepfakes

- Film producers will be able to create film representations of historical figures and make actors appear younger in films.
- Interactive educational simulations will help improve the education results for students.
- Synthetic patient data could allow researchers to conduct studies using protected private information in the healthcare industry.
- Voice synthesis technology helps people with disabilities.

CONSTITUTIONAL DIMENSIONS AND LEGAL CHALLENGES OF DEEPPAKE TECHNOLOGY

11. Constitutional Dimensions of Deepfake Technology

Deep Fake Technologies Present a Major Constitutional Challenge for Democracy While Artificial Intelligence (AI) has the potential to enhance freedom of expression as well as promote innovation, it also has the capability to be used against individuals privacy, dignity, reputation, election integrity and national security. Thus, the primary issue of constitutional law regarding Deep Fakes is how to balance out the competitive right of protection under Articles 19 and 21 of the Indian Constitution.

There are two main questions that the debate on Deep Fake Technology and its application relates to:

1. Does the creation and distribution of Deep Fake Videos/Photos represent an act of free speech that is protected by Article 19(1)(a)?
2. Can victims of Deep Fake Videos/Photos claim protection under their right to Privacy, Dignity, Autonomy and Reputation under Article 21?

Ultimately, the solution will come from finding a way to balance technological advancement with protecting our constitutional rights and preventing people from being harmed.

A. Article 19(1)(a): Freedom of Speech and Expression

- At face value, some examples of Deepfakes can be classified as "expression" or "speech", for which Article 19(1)(a) grants protections to all citizens. Examples of such categories include:

Artistic	Expression
Political	Communication
Satire/Parody	
Creative Works	(in general)
Digital Communications	

While initial appearances would indicate that certain types of "deep faked" content will qualify as "protected" speech/expression, there are limitations placed on this protection. Article 19(2), while permitting reasonable restriction(s) on free speech/expression, limits those freedoms in several areas:

- Creation of "Deepfake Pornography": Such material could reasonably be considered an infringement upon decency/morality. Misinformation related to elections, or other politically relevant information: Such information could reasonably be viewed as threatening the public order. Creation and dissemination of false video materials potentially damaging another's good name/reputation: Such actions could reasonably be interpreted as defamation.

Judicial Position

While recognizing the need for regulating speech with a legal impact in the case of *Shreya Singhal v. Union of India*, (2015) 5 SCC 1, the Court held that regulations on internet speech must meet the test of being reasonable as per the Constitution. The Court ruled out section 66(A) of the IT act to restrict the freedom of expression via Internet. Also, while ruling that this restriction is too broad, the court stated that it can regulate speech that causes cognizable harm to another person or entity under article 19(2). Therefore, any regulations regarding Deepfakes should be narrow and proportional.

B. Article 21: Right to Life and Personal Liberty

The right to life under Article 21 has expanded over time into an even broader guarantee of all interests directly threatened by Deepfake technology.

- **Right to Dignity;** Article 21 is based upon human Dignity. The moment someone's image or voice is altered without that person's consent, they have lost control of who they are and how others represent them publicly. People impacted by deepfakes often suffer from humiliation, psychological trauma, social isolation and professional loss.
- **Right to Reputation;** The Supreme Court has repeatedly held that a person's Reputation is an important aspect of Article 21. A fabricated video that depicts a crime, corrupt activity, or moral failure, can ruin an individual's Reputation within hours. The quick spread of information on social media further exacerbates this problem.

12. Deepfakes and the Right to Privacy

Privacy is the single biggest constitutional issue that arises from deepfakes. In fact, in *Justice K.S. Puttaswamy vs. the Union of India*, the Supreme Court recognized Privacy as a Fundamental Right.

A. Informational Privacy

Informational privacy relates to how much you have control over your own personal data. In order to generate the synthetic images, videos, audio recordings etc. that are part of deepfake technology, you need Images, Videos, Audio Recordings & Biometric Data. A large portion of this type of data is obtained without the users' consent. Therefore, when someone uses your biometric or other personal data to create synthetic media (i.e., fake images/videos) it directly violates your informational privacy rights.

B. Decisional Autonomy

Individuals possess the right to determine how they present themselves to society. Deepfakes

undermine this autonomy by creating fabricated representations beyond the individual's control.

C. Identity Rights

An individual's face and voice form essential aspects of personal identity. The replication of these attributes without consent threatens personal autonomy and self-determination.

13. Deepfake Pornography and Gendered Harm

One of the most harmful applications of deepfake technology involves non-consensual synthetic pornography. Studies consistently indicate that women constitute the overwhelming majority of victims. The process generally involves:

The creation and dissemination of an image or video featuring a person's likeness without their permission is a violation of their personal rights as well as their right to privacy. One type of this manipulation that has serious implications for those who have been victimized by such tactics, include the use of deep fakes in the creation of non-consensually generated pornographic images/videos. There are multiple studies that support that the vast majority of victims of this tactic were female. The process used typically includes:

- Gathering photos of the subject from public sources.
- Digitally superimposing the faces of subjects into the explicit imagery.
- Distributing the altered images/video using digital platforms.

A. Violation of Dignity

Beginning with the fact that the creation of deepfake pornography transforms an individual's body and identity for the purposes of exploitation; it can cause much more than damage to one's reputation. As a result of this type of harassment victims are often plagued by Anxiety, Depression, Social Isolation, Career Setbacks. In extreme cases victims have

reported experiencing Suicidal Ideation and Long-Term Psychological Trauma.

B. Violation of Sexual Privacy

In addition to having a legal basis in the right to control how images of you in sexually explicit situations are used, your right to sexual privacy also includes a social expectation that these representations will be private. Although the actual photograph itself may be created using technology (i.e., "deepfakes"), if people perceive them to be authentic then they can still create considerable amounts of harm.¹⁴⁶⁴

14. Electoral Manipulation and Democratic Governance

Deepfakes present a serious threat to democratic institutions. Historically, misinformation campaigns relied upon edited videos and fabricated narratives. Deepfakes significantly enhance these capabilities by producing realistic synthetic content.

A. Threat to Electoral Integrity

Deepfakes in political advertising are an example of how "Fake News" could be created to sway public opinion. The use of social media and mobile devices has increased the speed at which fake news is distributed as well as the number of individuals that view it. For this reason, the creation of deepfakes in the context of political advertising is particularly concerning as they could potentially damage a campaign's reputation, create distrust among potential voters for a particular candidate and ultimately affect voting decisions. Deepfakes could also help to create divisions among certain social groups by spreading misinformation about candidates. This could lead to a lack of confidence in the electoral process.

B. The "Liar's Dividend"

Liar's dividend" represents a particular concern when it comes to deep fakes. With an increased number of deep fakes available, people will no longer know what evidence to believe. In other words, politicians accused of misconduct could

¹⁴⁶⁴ Danielle Keats Citron, Sexual Privacy, 128 Yale L.J. 1870, 1875–80 (2019).

say that real video or audio recordings of them were made using artificial intelligence. Therefore, there may be less confidence in the actual evidence used in their cases.

C. Indian Context

The India election has several aspects which make it vulnerable to being manipulated by AI-generated campaign material. For example, its massive voter base and extensive use of social media present some challenges. Regulators need to determine how much they want to limit the potential for electoral fraud while protecting the right of free political speech.

15. Financial Fraud and Economic Crimes

Deepfake technologies are creating new forms of fraudulent activity in addition to traditional schemes. Voice cloning is an example of a deep fake system that provides a highly realistic replica of executives, government employees, or family members.

A. Corporate Fraud

Criminals who clone voices may be able to authorize transfers of funds; make changes to bank accounts; and obtain confidential information. Several major companies worldwide have reportedly lost tens of millions of dollars due to this form of artificial intelligence (AI)-enabled impersonation fraud.

C. Banking and Consumer Fraud

Deepfakes are a threat to people who have been victims of fraudulent schemes because they could be contacted by a "customer service" person calling themselves from an organization you know. They might also receive "approval for loans," or other types of notices which ask them for their private information. People who fall victim to these scams will likely experience deceitful actions including having some of their private info compromised. The loss of money is possible too as well as the distrust of all forms of communication with computers.

D. Identity Theft

Synthetic identity creation is one of several methods by which Criminals can bypass Biometric Identification and Financial Services Fraudulent Use. Current Cybersecurity Frameworks are challenged by the risk posed by Deep Fake Systems.

16. National Security Implications

Government agencies around the world are recognizing Synthetic Media as a National Security Threat and as an Information Warfare tool.

A. Military Deception

Military Announcements, Surrender Statements and Strategic Communications in general, can potentially be created using fake media. In this context, if disseminated during a military conflict it could lead to Confusion for those involved.

B. Diplomatic Consequences

Diplomatic Crises may result from fabricated video footage of Political Leaders making inflammatory comments.

C. Terrorism and Extremism

Deep fakes are being used by extremist organizations to disseminate propaganda, attract new followers, and instigate violence against their intended audience. Because the synthesized information looks and sounds so real, it can be more believable and influential than traditional misinformation. Deep fakes could therefore increase the effectiveness of radicalizing individuals as well as create serious security threats.

17. Recent Indian Deepfake Incidents

Rashmika Mandanna Deepfake Controversy: In 2023, an edited (manipulated) video falsely claiming to be of actress Rashmika Mandanna went viral across various forms of Social Media. As such it generated widespread National Debate over Privacy Rights/Consent for Digital Content; Digital Platforms' Accountable Use of User-Generated Data; Artificial Intelligence Regulation/Use.

Alia Bhatt Deepfake Incident: The Indian government subsequently increased its emphasis on taking a much harder line when enforcing legislation prohibiting dangerous Synthetic Video/Media.

Sachin Tendulkar Deepfake Scam : Videos that claimed to depict actress Alia Bhatt performing actions she had never done illustrated the rapidly increasing accessibility of deep fake technology.

INDIAN LEGAL FRAMEWORK, INTERMEDIARY LIABILITY, AND COMPARATIVE

REGULATORY APPROACHES

18.Existing Legal Framework in India

Currently in India, there is no legislation that will specifically regulate Deep Fakes. As such, victims or enforcement bodies have to use various fragmented constitutional, civil, criminal and cyber law provisions which are partially protective but were passed prior to the advent of Advanced Generative AI technologies; and do not deal with all the specific issues created by Deep Fake Technologies.

Information Technology Act, 2000

The Information Technology Act, 2000 is India's principal statutory authority in relation to addressing cybercrime. Despite being passed well prior to the advent and common utilization of deep fake technologies; many of the ITA provisions are applicable to illegal acts related to synthetic media (i.e. fraud, identity theft, spreading false information and/or unauthorized access/processing of private data).

A. Section 66C – Identity Theft

Section 66C of India's IT Act prohibits fraudulently using an individual's electronic signature; a password; or a unique identifier. Fraudulent use of one's face (facial characteristics) in deep fake videos is often done without permission. Using this synthetic information for deception can be considered as stealing someone else's identity.

B. Section 66D – Cheating by Personation

The provision in Section 66D addresses the use of computer systems for the purpose of fraudulent acts (i.e., cheating), which includes voice-cloning scams, as well as synthetic customer service calls, deep-fake investment schemes, and impersonating a public official. Although the statute was created prior to the development of artificially generated content using artificial intelligence, it could be difficult to interpret how this will apply to situations involving deepfakes.

C. Section 67

The use of deepfakes to create pornography would likely be covered under section 67 of the Information Technology Act because the creation of pornographic videos using fake images or video is illegal. Therefore if you were to create and distribute a pornographic video created with a deepfake you could face serious legal consequences.

D. Section 67A

This provision has particular relevance to deepfake pornography, revenge pornography and other forms of non-consensual intimate images and is particularly significant due to the fact that many offenders remain anonymous and/or operate out of jurisdiction. The law provides a very useful tool for responding to these crimes but it is also difficult to enforce as the anonymity and multi-jurisdictional nature of offending make it very difficult to identify the perpetrator.

E. Limitations of the IT Act

Modern regulations have numerous shortcomings to deal with the emerging threat of deepfakes. There is no specific definition of deepfake as it relates to the Information Technology Act; there are no new liability provisions regarding artificial intelligence (AI); and the use of synthetic media is subject to no legal obligation for producers to disclose that their products are artificially produced. Furthermore, a lack of clear mandates and an absence of standards for detecting deepfakes

contribute to complications when attempting to enforce laws and ensure accountability across international borders and jurisdictions.

Digital Personal Data Protection Act, 2023

The DPDP Act, 2023, provides India with its first law that comprehensively governs the protection of personal data. In terms of deepfakes, the Act is very relevant since the creation, development, training, and operation of deepfake systems are based upon collecting a large amount of an individual's personal information to be processed.

A. Personal Data and Deepfakes

Personal data is defined in the DPDP Act as being any information that relates to an identified or identifiable natural person. Images, video files, audio recordings, and biometrics are all common input data for deepfakes. As a result, because these data types allow direct identification of individuals, the Act will protect such data.

B. Consent Requirements

Most unauthorized (deep) fakes have been produced as such, in order that they can be used for illegal activities, such as identity theft or impersonation. This activity will likely include the violation of the Data Protection Directive (DPDP), which states that personal data can only be collected if an individual has given free, unambiguous, specific and informed consent.

C. Rights of Data Principals

Under the DPDP Act, an individual will have a number of fundamental rights that include; the right to obtain access to his/her own personal data, the right to correct inaccurate personal data, the right to request for erasure of personal data (also known as 'right to be forgotten'), and the right to file a complaint. Such provisions would assist victims in obtaining the removal of unlawful created deepfakes or synthetic content.

D. Regulatory Gaps

Deepfakes are important to the law of the DPDP Act yet there is no specific definition of

Deepfakes in the DPDP Act, nor do they provide regulations around how AI training datasets should be created; nor do they specify what information needs to be disclosed about synthetic media; nor do they outline clearly where the responsibility rests with digital platforms that host deepfake images. Therefore, although the DPDP Act provides some good data protection protections it does not cover all the special legal and regulatory issues associated with the use of deepfakes.

Bharatiya Nyaya Sanhita, 2023

Deepfake content can be considered defamatory if it is created from video or audio recordings with false information portraying an individual committing crimes, corrupting others, behaving immorally, speaking offensively, etc., resulting in damage to their reputation. In addition, creating and distributing fake synthetic media could potentially violate provisions regarding forgeries. Courts are increasingly recognizing AI-created content as forms of digital forgery. Moreover, deepfakes have the potential to enable financial scams through the use of deception for obtaining wrongful gains or losses. Furthermore, deepfakes could provide opportunities to commit extortion, blackmail, harassment and/or criminal intimidation.

Copyright Implications of Deepfakes

A. Unauthorized Use of Copyrighted Material

Many of today's deep fakes utilize the copyrighted materials that include movies, photographs, music, TV shows, etc. In some cases the unauthorized duplication, alteration, and/or utilization of this material for the creation of artificial media is a violation of the copyright holder(s) rights. Therefore creators/distributors of deep fakes can be liable in accordance with the applicable copyrights.

B. Ownership of AI-Generated Content

A major outstanding question is what are the rights of authors/owners of created by an artificial intelligence (AI) "deepfake." Important

issues in this area include who would have the right to claim title to such a work? Would such AI-generated works be eligible for copyright protection? To whom would liability fall -- to the user of the deepfake; the developer of the technology that generated the deepfake; or both? Today, India's copyright statute does not address the rapidly evolving new legal issues presented by AI-generated content.

Intermediary Liability and Platform Responsibility

Social Media Plays an Important Role in Creating and Distributing Deepfakes. The reason social media is so important to create and share fake videos is that they have such large audiences. Therefore, it is critical that any law regulating how deepfakes are created, distributed or used has some level of intermediary liability. Liability can be met by having good content moderators who will remove fake content when reported, having ways to detect fake content and removing them quickly, and providing information about fake content to authorities.

A. Safe Harbour Protection

Safe Harbour provisions contained within Section 79 of India's IT Act are designed to protect Internet Service Providers (ISPs) from liability for the content that users post online. ISPs will be protected by these provisions when they operate strictly as an intermediary – i.e., do not edit the user-submitted content; reasonably monitor the user submitted content; and immediately take down or restrict access to the offending material once they receive either a formal court order or have received notice of a claim.

B. Information Technology Rules, 2021

With the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, there are a number of responsibilities placed upon intermediaries. For example, they will have to develop procedures for resolving grievances as well as providing users with mechanisms to submit complaints regarding their usage of such services. Additionally, once an intermediary receives a valid request from

another party that they remove certain content from one of their services; it shall be required to do so in accordance with those terms. Furthermore, each intermediary is also expected to designate a person who is responsible for ensuring that the intermediary complies with all applicable laws and regulations. These types of regulation can be used by governmental bodies to identify harmful deepfakes, track them down, and reduce or prevent them from being distributed via the internet.

C. Challenges for Platforms

Deepfake regulation is also plagued with many challenges. With modern deepfakes being as sophisticated as they are today, it has become very hard for automated detection systems to accurately identify them. In addition, the sheer volume of digital content that is posted each day makes detecting or even just monitoring deepfakes a difficult task. Lastly, jurisdiction can be an issue if the content is being created in another country (outside of India) and can cause potential problems.

23. Comparative Regulatory Approaches

A. European Union

The European Union (EU) has developed an extraordinary regulation framework in the world that regulates artificial intelligence by means of the EU Artificial Intelligence Act. This legislation establishes new standards for transparency as it obliges producers of AI-generated content to indicate when they have used AI to generate some or all of the content. In addition, the Act applies a risk-based methodology and categorizes AI systems into four levels of risk: unacceptable; high; low/limited; and minimum. Each category will correspond to different regulatory measures. Finally, the EU AI Act takes seriously the need to protect the fundamental rights of individuals, especially their right to data protection, human dignity, equality/non-discrimination, and democratic integrity. Therefore, the Indian government may see merit in adapting the EU AI Act to fit the needs and culture of India but using aspects such as

disclosure/identification of AI-generated content, regulation based on risk, and emphasis on protection of fundamental rights, among other features, could provide great insight to how best to create regulations for AI use in India.

B. United States

The United States does not currently have a comprehensive federal law governing deepfakes, with regulation occurring mainly at the state level. Texas prohibits deceptive deepfakes designed to influence elections, aiming to safeguard electoral integrity. California has enacted laws targeting election-related deepfakes and deepfake pornography, allowing victims to seek civil remedies. This state-based approach encourages regulatory innovation and victim protection; however, it also creates a fragmented legal landscape and lacks nationwide consistency in addressing deepfake-related harms.

C. China

China is also implementing a highly restrictive approach to regulating synthetic media. A regulatory framework for the use of deep synthesis technology will include all users who create or distribute AI-created material to be required to clearly label that it was created using artificial intelligence; users creating/distributing AI-created material must have their identities verified; and services providing access to the creation/distribution of AI-created material must undergo an independent assessment of potential security risks. The focus of this regulatory structure are government oversight; public safety; and the prevention of misinformation as a means to mitigate the perceived threats from synthetic media. Critics of these restrictions feel they could severely inhibit freedom of speech in addition to increasing the extent to which the Chinese government exercises control over online communication.

D. United Kingdom

The U.K. is using an approach to digital content governance called "platform governance."

Platform Governance will focus on creating a safer online environment for its citizens, ensuring platforms are held accountable in their actions (on and off their platforms), and reduce harms associated with user experience.

24. Judicial Trends and Emerging Challenges

Although there is a growing number of court cases across jurisdictions dealing with issues related to Artificial Intelligence (AI), Jurisprudence related specifically to Deep Fakes has been very limited. However, many developing trends can be identified. Increasingly, Courts are granting protections for individuals regarding their right to protect images; their right to privacy relative to biometrics; and their right to informational autonomy that will provide stronger protections from those who would create or utilize unauthorized Deep Fakes. In addition, Courts are becoming increasingly focused on the issue of Platform Accountability, requiring Intermediaries to take action in response to harmful content in timely fashion with the potential of being held liable if they fail to do so. Finally, Future Legal Developments are likely to recognize an Individual's Face, Voice and other Biometric Characteristics as Protected Elements of Personal Identity and Personality Rights.

25. Critical Evaluation of the Existing Framework

Several major concerns continue to exist with regard to regulation of deepfakes; including, fragmented legal provisions, lack of clarity on statutory definitions, difficulty enforcing laws across international borders, few options available for remediation for victims and technology is developing faster than detection methods to identify increasing levels of sophistication in synthetic media. As demonstrated through the research, deepfakes cause multi-dimensional harm in terms of privacy, reputation, dignity, identity, election integrity, national security and rights as an owner of Intellectual Property.

RECOMMENDATIONS

Recommendation 1: Enact a Comprehensive Deepfake Regulation Law

India needs to enact a specific piece of legislation on Synthetic Media and Deep Fake regulation and artificial intelligence accountability to deal with increasing threats posed by fake videos and photos. An independent law will give legal clarity for what constitutes a deep fake, establish standards for when companies are liable, regulate how much artificial intelligence can generate video/photos that could be misleading or false and set up procedures for enforcing these new regulations.

Recommendation 2: Establish a Clear Legal Definition of Deepfakes

The proposed legislation should define "deepfakes" with as much detail and breadth as possible in order to provide an acceptable degree of clarity for those who will enforce it (courts, regulatory bodies, police etc.) and/or use it (technology providers). Clear definitions would enable the identification and handling of harmful synthetic media content; clarify ambiguous laws so that they can be enforced; and improve the effectiveness of both enforcement actions, compliance efforts and adjudications.

Recommendation 3: Introduce Mandatory Labelling and Disclosure Requirements

All artificial intelligence (AI)-generated synthetic content should require obligatory disclosure. Using visible labels or machine-readable "watermarks" on synthetic media can assist end-users in identifying media that have been manipulated and differentiating it from authentic, unaltered media. As such, greater transparency within synthesized media will help promote truthfulness; accountability for producers/creators/distributors of synthetic media; build credibility/trust with the general public concerning digital communication; and foster responsible use/development of artificial intelligence (AI) technology.

Recommendation 4: Criminalize Malicious Deepfake Activities

The law should clearly make it a crime for people to use harmful methods such as deepfakes (especially deep fake pornography, manipulating elections, committing financial fraud, or threatening national security) and enhanced penalties could be applied in cases that result in serious harm. Clearer criminal laws would discourage the abuse of this type of tool, hold those who have abused this type of tool accountable, and give law enforcement better tools to prevent and prosecute crimes related to the use of these types of tools.

Recommendation 5: Strengthen Victim-Centric Remedies

Legal remedies for victims who are harmed by deepfakes, such as immediate removal of fake content, injunctions (to stop further harm) and monetary damages should be available. The monetary remedy should include loss of reputation, financial loss and mental health/psychological loss. Ensuring a 'victim centered' response will provide immediate relief from ongoing harms, assist in maintaining public trust that there is a responsive justice system.

Recommendation 6: Establish an Artificial Intelligence Regulatory Authority of India (AIRAI)

India could create a new body called the "Artificial Intelligence Regulatory Authority of India" to ensure both regulation of artificial intelligence in India and compliance with it. The AIRAI will be required to investigate and determine if AI systems are being used responsibly as well as require that all regulatory standards for use of AI systems are developed and implemented by the agency. The AIRAI will also provide a means to develop global standards for the development and implementation of AI systems.

Recommendation 7: Promote Public Awareness and Digital Literacy

Legal actions will be insufficient to provide an effective solution for the issues of "deepfake"

technology. To that end, it is recommended that governmental agencies establish educationally focused public awareness campaigns, media literacy training programs and community outreach programs in schools, universities and local communities. Improved digital literacy will enable citizens to identify synthetic content online; to evaluate digital sources critically; to protect themselves from misinformation; and to engage more safely in all aspects of the digital world.

CONCLUSION

The use of deep fake technology is a major legal/regulated issue resulting from the fast development of artificial intelligence (AI). Synthetic media has many advantages for entertainment, education, health care, accessibility and innovation. However, when used inappropriately it can create harm to individuals' rights to privacy; their dignity; their reputations; democratic governance and national security. In addition to showing the above as true, the research conducted by this paper shows that current Indian law provides few protections or remedies related to the issues discussed in this paper. In particular while some provisions in the information technology act; digital personal data protection act; and bharat nyay sanhita could potentially be utilized against forms of deep fake related misconduct, they are generally not suited to the unique threats created through the use of synthetic media generated using AI.

Comparative research has shown that countries around the world have started to use an increasing number of transparent, accountable, disclosing and responsible for platforms (platforms) as their regulatory approach. The EU's Artificial Intelligence (AI) Act, China's Deep Synthesis Regulations, and growing legislative efforts in America show a clear trend toward global regulation of artificial intelligence. Therefore, India needs to proactively build a strong and integrated regulatory system which protects technological innovations while

protecting rights under the Indian Constitution. That would include such requirements for mandated disclosures; digital watermark standards; mechanisms for accountability on the part of platforms; legal recourse for victims; and sufficient measures for enforcement.

Ultimately, the challenge presented to us by Deepfakes is one that cannot solely be addressed technologically; it has significant constitutional implications for our democracy as well as social implications. In order to protect truth, trust, privacy, dignity, and democratic legitimacy in the face of this growing threat, we need to develop a comprehensive approach which involves all of these stakeholders – lawmakers, judges, regulators, tech companies, and civil society. It will be through working together (collaborative governance) that we are able to fully realize the potential of AI while minimizing its risks.

BIBLIOGRAPHY

A. Cases

Justice K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1.

Shreya Singhal v. Union of India, (2015) 5 SCC 1.

Subramanian Swamy v. Union of India, (2016) 7 SCC 221.

Anuradha Bhasin v. Union of India, (2020) 3 SCC 637.

Kharak Singh v. State of Uttar

Pradesh, AIR 1963 SC 1295.

R. Rajagopal v. State of Tamil Nadu, (1994) 6 SCC 632.

B. Statutes

The Constitution of India.

Information Technology Act, No. 21 of 2000 (India).

Digital Personal Data Protection Act, No. 22 of 2023 (India).

Bharatiya Nyaya Sanhita, No. 45 of 2023 (India).

Copyright Act, No. 14 of 1957 (India).

Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.

United Nations High Commissioner for Human Rights, The Right to Privacy in the Digital Age (2024).

C. Books

Stuart Russell & Peter Norvig, *Artificial Intelligence: A Modern Approach* (4th ed. 2021).

Cathy O'Neil, *Weapons of Math Destruction* (2016).

Danielle Keats Citron, *The Fight for Privacy* (2022).

Luciano Floridi, *The Ethics of Artificial Intelligence* (2023).

D. Journal Articles

Robert Chesney & Danielle Keats Citron, Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security, 107 Calif. L. Rev. 1753 (2019).

Danielle Keats Citron, Sexual Privacy, 128 Yale L.J. 1870 (2019).

Danielle Keats Citron & Mary Anne Franks, Criminalizing Revenge Porn, 49 Wake Forest L. Rev. 345 (2014).

Robert Chesney & Danielle Keats Citron, Deep Fakes and the New Disinformation War, 98 Foreign Aff. 147 (2019).

Lisa Macpherson & Sara Collins, Deepfake Accountability and Regulation, 75 Fed. Comm. L.J. 121 (2023).

E. Reports and International Documents

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

OECD Recommendation on Artificial Intelligence (2019).

UNESCO Recommendation on the Ethics of Artificial Intelligence (2021).

World Economic Forum, Global Risks Report 2025.