The Indian Journal for Research in Law and Management

Open Access Law Journal – Copyright © 2025

Editor-in-Chief – Dr. Muktai Deb Chavan; Publisher – Alden Vas; ISSN: 2583-9896

Regulation of Deepfakes in India: Balancing Innovation, Privacy and Freedom of Speech

By Virat Choudhary

Abstract

The rapid advancement of Artificial Intelligence (AI) and generative technologies has led to the emergence of deepfakes—highly realistic synthetic audio, video, and image content generated using deep learning techniques. While deepfakes have transformative potential in sectors such as entertainment, education, healthcare, and digital communication they simultaneously pose serious threats to privacy, democracy, reputation, and public trust. India, being one of the world's largest digital democracies, faces a critical challenge in regulating deepfake technology without stifling innovation or undermining constitutional guarantees of free speech and expression.

This paper critically examines the legal and constitutional dimensions of deepfake regulation in India. It analyses the applicability of existing laws, including the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023, the Digital Personal Data Protection Act, 2023, and intermediary liability frameworks under the Information Technology Rules, 2021. The paper further explores constitutional concerns relating to privacy under Article 21 and freedom of speech under Article 19(1) (a) of the Constitution of India. By comparing India's legal position with regulatory models adopted in the European Union, the United States, and China, the study identifies major regulatory gaps and proposes a balanced rights-oriented framework for deepfake governance in India.

The paper argues that India requires a dedicated and technologically adaptive legal framework that protects individual dignity and democratic integrity while encouraging responsible AI innovation. It concludes that effective deepfake regulation must be based on

transparency, accountability, consent, platform responsibility, and constitutional proportionality.

Keywords- Deepfakes, Artificial Intelligence, Privacy, Freedom of Speech, Generative AI, Digital Governance, Synthetic Media, Constitutional Law.

1. Introduction

Artificial Intelligence has significantly transformed the digital ecosystem by enabling machines to replicate human cognition, speech, and creative expression. Among the most disruptive developments in AI is the emergence of "deepfake" technology. The term "deepfake" combines "deep learning" and "fake" and refers to manipulated or synthetically generated media that realistically imitates the appearance, voice, gestures, or actions of a real person.

Deepfakes are primarily created using machine learning models such as Generative Adversarial Networks (GANs), auto encoders, and neural rendering systems. Initially regarded as experimental technological tools, deep fakes have now become accessible to ordinary internet users through publicly available software and mobile applications. The increasing realism of AI-generated content has blurred the distinction between authentic and fabricated media, thereby creating significant legal, ethical, and societal concerns.

Although deep fake technology has legitimate applications in cinema, digital accessibility, education, virtual communication, and entertainment, it has also become a tool for disinformation, identity theft, cyber fraud, political propaganda, financial scams, and non-consensual pornography. The widespread misuse of deep fakes has raised alarms globally regarding electoral integrity, public trust in digital media, and individual autonomy.

India presents a unique regulatory challenge in this regard. As one of the largest internet markets in the world, India witnesses massive digital engagement through social media platforms, messaging services, and AI-driven applications. At the same time, the constitutional commitment to freedom of speech and expression under Article 19(1)(a) requires the State to ensure that regulation does not become a mechanism for excessive censorship or suppression of legitimate expression.

The legal debate surrounding deepfake regulation in India therefore lies at the intersection of three competing interests-

1. Protection of innovation and technological growth;

2. Protection of privacy, dignity, and reputation; and

3. Preservation of freedom of speech and democratic discourse.

This paper seeks to analyse whether the existing Indian legal framework adequately addresses the harms associated with deep fakes and whether India requires dedicated legislation specifically governing synthetic media and generative AI.

2. Research Questions

The present research attempts to answer the following questions:

1. What are deep fakes and how do they affect society, privacy, and democratic institutions?

2. Whether the existing legal framework in India is sufficient to regulate deepfake technology?

3. How can India balance technological innovation with privacy and freedom of speech?

4. Whether India requires a separate and dedicated law for deepfake regulation?

5. What lessons can India learn from international regulatory approaches?

3. Hypothesis

This paper is based on the hypothesis that the existing legal framework in India is inadequate to effectively address the challenges posed by deepfake technology. Although current laws provide partial protection against cyber offences, privacy violations, and obscene content, they do not specifically regulate AI-generated synthetic media. Therefore, India requires a balanced and rights-oriented regulatory framework that protects privacy and democratic values without unnecessarily restricting innovation and freedom of speech.

4. Research Objectives

The present study seeks to

1. Examine the concept, functioning, and societal impact of deepfake technology

2. Analyse the constitutional implications of deepfakes in India

3. Evaluate the adequacy of existing Indian laws in addressing deepfake-related harms

4. Compare India's legal framework with international regulatory models

5. Identify the regulatory and enforcement challenges associated with deepfakes and

6. Propose a balanced and rights-oriented legal framework for deep fake regulation in India

5. Research Methodology

This paper adopts a **doctrinal and analytical method of research**. The study relies upon primary and secondary legal sources, including statutes, judicial decisions, government advisories, scholarly articles, policy papers, journal publications, and comparative international legal frameworks.

Primary sources include the Constitution of India, the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023, and the Digital Personal Data Protection Act, 2023. **Secondary sources** include academic literature on artificial intelligence governance, digital privacy, platform accountability, and comparative cyber law.

The research further adopts a comparative approach by examining the regulatory responses of jurisdictions such as the European Union, the United States, and China.

The study mainly relies upon doctrinal analysis and interpretation of constitutional principles, cyber law provisions, and AI governance frameworks. Efforts have been made to maintain an objective and analytical approach throughout the paper.

6. Literature Review

Several scholars and policy researchers have examined the legal and social impact of deepfake technology in recent years. Danielle Citron and Robert Chesney [1]have argued that deepfakes pose serious threats to privacy, democracy, and national security. Their work highlights the misuse of synthetic media for disinformation, harassment, and manipulation.

Mika Westerlund[2] discusses how deepfake technology has evolved rapidly and why governments are struggling to create effective regulatory responses. Similarly, Jan Kietzmann

[1] Robert Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 California Law Review (2019).
[2] Mika Westerlund, *The Emergence of Deepfake Technology: A Review*, 9 Technology Innovation Management Review 39 (2019).

[3]and other researchers have examined both the positive and negative implications of deepfakes in business and media environments.

In the Indian context, legal scholarship on deepfakes is still developing. Existing studies mainly focus on the applicability of the Information Technology Act, intermediary liability, and privacy concerns after the Puttaswamy judgment. However, there remains limited detailed research regarding the constitutional balance between innovation, privacy, and freedom of speech in India.

This paper attempts to contribute to the existing literature by combining constitutional analysis, comparative study, and policy-oriented recommendations within the Indian legal framework.

7. Understanding Deepfakes and Synthetic Media

➢ Growth of Deepfake Technology: Statistical Trends and Data Analysis

The rapid increase in deepfake content over the last few years demonstrates how quickly generative AI technologies are expanding across the world. Earlier, deepfakes were limited to experimental research projects and small online communities. However, with the rise of publicly available AI tools and mobile applications, the production and circulation of synthetic media has increased significantly.

Several international studies and cyber security reports indicate that deepfake-related incidents have grown at an alarming rate year by year.

Year-wise Growth of Deepfake Technology

Year	Major Development	Estimated Trend
2017-	Emergence of deepfake software on online forums	Initial public exposure
2018-	Growth of AI face-swapping applications	(Rapid experimentation phase)
2019-	Increase in political and celebrity deepfakes	(Global policy concern begins)

[3] Jan Kietzmann et al., *Deepfakes: Trick or Treat?* 64 Business Horizons 135 (2020).

2020- Rise of deepfake pornography and misinformation during COVID-19 (Significant misuse reported)

2021- Expansion of AI voice cloning tools (Increase in cyber fraud risks)

2022- Generative AI tools become commercially accessible (Mainstream adoption of AI content creation)

2023- Deepfake scams and election-related manipulation increase globally (Regulatory debates intensify)

2024- Governments and technology companies introduce AI labelling and transparency measures (Beginning of formal AI governance frameworks)

2025- Increased integration of generative AI into social media and communication platforms (Need for stronger legal safeguards)

According to multiple cyber security and AI governance reports, deepfake videos online have increased exponentially since 2019. Studies have shown that a substantial portion of harmful deepfake content consists of non-consensual explicit material, with women being disproportionately targeted.

Financial fraud through AI voice cloning has also increased in recent years. Cyber security agencies across different jurisdictions have reported incidents where scammers used cloned voices to impersonate company executives, government officials, and family members in order to obtain money or confidential information.

In India, concern regarding deepfakes significantly increased after several manipulated videos involving celebrities and public figures circulated on social media platforms in 2023 and 2024. These incidents triggered public debate regarding AI regulation, platform accountability, and digital ethics.

The increasing accessibility of generative AI tools suggests that deepfakes are no longer limited to technologically advanced users. Today, realistic manipulated media can be produced with minimal technical expertise using commercially available applications and online software.

This statistical and technological growth highlights the urgent need for a clear and balanced legal framework capable of addressing the risks associated with synthetic media.

7.1 Meaning and Nature of Deepfakes

Deepfakes refer to digitally manipulated or synthetically generated media created using AI systems capable of mimicking human appearance, speech, behaviour, or identity. Deepfake systems rely on large datasets consisting of facial images, voice samples, video recordings, and behavioural patterns to train machine learning algorithms.

The most common techniques used in deepfake creation include:

- Generative Adversarial Networks (GANs)
- Auto encoder-based facial mapping
- Neural voice cloning
- Motion synthesis systems and
- Diffusion-based generative models
- Deepfake technology can create
- Face-swapped videos
- AI-generated voices
- Synthetic political speeches
- Fake celebrity endorsements
- Digitally altered photographs and
- Fully AI-generated virtual humans

With the rapid development of generative AI tools, synthetic media has become more realistic than before and, in many cases, difficult for ordinary users to identify.

7.2 Legitimate Uses of Deep-fake Technology

Deep-fakes are not inherently harmful. In many sectors, synthetic media technology has beneficial applications

(**a) Entertainment and Cinema**

Film industries increasingly use AI-based facial reconstruction and voice replication for dubbing, de-aging actors, and recreating historical characters.

(b) **Accessibility and Education**

Synthetic voices assist individuals with speech disabilities, while AI-generated simulations improve educational and training experiences.

(c**) Digital Communication**

AI avatars and virtual assistants facilitate multilingual communication and customer support.

(d) **Historical and Cultural Preservation**

Museums and educational institutions use AI-generated recreations of historical figures for interactive learning.

These positive applications demonstrate that deepfake regulation should not prohibit the technology itself but rather target harmful and malicious uses.

7.3 Harmful Uses of Deepfakes

Despite legitimate uses, deepfakes have increasingly been weaponized.

(a**) Non-Consensual Pornography**

One of the most serious abuses involves the creation of sexually explicit deepfake content without consent. Women are disproportionately targeted, resulting in severe emotional and reputational harm.

(b**) Political Manipulation and Disinformation**

Deepfakes may be used to fabricate speeches or videos of political leaders, potentially influencing elections and public opinion. In recent years, concerns regarding election-related misinformation have increased globally because manipulated content can spread very quickly through social media platforms before proper verification takes place.

(c**) Financial Fraud and Cybercrime**

AI-generated voice cloning has been used to impersonate executives, government officials, and family members for fraudulent purposes.

(**d) Defamation and Character Assassination**

Deepfakes may spread false narratives and damage reputations by depicting individuals engaging in acts they never committed.

(e) Erosion of Public Trust

The existence of deepfakes undermines trust in digital evidence and creates what scholars describe as the "liar's dividend," where genuine evidence can also be dismissed as fake.

8. Constitutional Dimensions of Deepfake Regulation in India

8.1 Right to Privacy under Article 21

The right to privacy was recognized as a fundamental right by the Supreme Court of India in **Justice K. S. Puttaswamy v. Union of India**.[4] The Court held that privacy is intrinsic to life, liberty, dignity, and personal autonomy under Article 21.

Deepfakes directly implicate several dimensions of privacy

- Informational privacy
- Bodily autonomy
- Decisional autonomy
- Reputational privacy and
- Psychological integrity.

The unauthorized use of a person's image, facial features, or voice amounts to a violation of informational self-determination. In cases involving explicit or defamatory deepfakes, the harm extends beyond mere data misuse and affects dignity and mental well-being.

The constitutional recognition of privacy therefore provides a strong normative basis for regulating malicious synthetic media.

8.2 Freedom of Speech and Expression under Article 19(1)(a)

Article 19(1) (a) guarantees freedom of speech and expression to all citizens. Any attempt to regulate deepfakes must therefore comply with constitutional limitations.

Not all deepfakes are unlawful. Certain forms of synthetic media may constitute:

- Artistic expression
- Satire and parody
- Political commentary

[4] *Justice K. S. Puttaswamy v. Union of India*, (2017) 10 SCC 1.

- Cinematic expression
- Educational experimentation or
- Creative innovation.
- Excessive regulation may result in
- Prior censorship
- Chilling effects on online speech
- Over-removal of legitimate content by intermediaries; and
- Suppression of artistic creativity

At the same time, Article 19(2) permits reasonable restrictions in the interests of:

- Defamation
- Public order
- Decency and morality
- Sovereignty and integrity of India and
- Security of the State

Therefore, the main constitutional challenge is to regulate harmful deepfakes without unnecessarily restricting legitimate speech and creative expression

8.3 Right to Reputation and Human Dignity

Indian constitutional jurisprudence recognizes reputation as an aspect of dignity under Article 21. Deepfake technology can irreparably damage personal and professional reputation.

Victims of manipulated media often experience:

- Social ostracism;
- Professional loss;
- Psychological trauma;
- Harassment; and
- Public humiliation.

Accordingly, any legal framework governing deepfakes must integrate dignity-based protections alongside free speech safeguards.

9. Existing Legal Framework Governing Deepfakes in India

At present, India does not have a separate law specifically dealing with deepfakes or AI-generated synthetic media. Existing legal responses are fragmented across cyber laws, criminal statutes, data protection laws, and intermediary regulations.

9.1 Information Technology Act, 2000

The Information Technology Act, 2000[5] remains India's primary cyber law legislation.

- **Section 66C – Identity Theft**

This provision criminalizes fraudulent use of electronic signatures, passwords, and unique identification features. Deepfake impersonation involving facial or voice replication may fall within its scope.

- **Section 66D – Cheating by Personation**

This section penalizes cheating through personation using computer resources. AI-generated impersonation scams may be prosecuted under this provision.

- **Section 66E – Violation of Privacy**

Section 66E punishes capturing or transmitting images of private areas without consent. It may apply to intimate deepfake content.

- **Sections 67 and 67A**

These provisions criminalize the publication and transmission of obscene and sexually explicit material in electronic form.

9.2 Limitations of the IT Act

Despite partial applicability, the IT Act suffers from several limitations-

- It predates generative AI technology
- It does not define deepfakes or synthetic media
- It lacks consent-based standards for AI-generated identity manipulation, and
- It does not address AI training datasets and algorithmic accountability.

9.3 Bharatiya Nyaya Sanhita, 2023

[5] Information Technology Act, 2000.

The Bharatiya Nyaya Sanhita (BNS), 2023[6] contains provisions relevant to deepfake-related offences.

These include

- Cheating
- Forgery
- Criminal intimidation
- Defamation
- Publication of obscene content; and
- Impersonation.

Deepfake-generated misinformation, fraudulent communication, and defamatory synthetic media may be prosecuted under these provisions.

However, the BNS similarly lacks-

- A definition of synthetic media
- Technology-specific offences and
- AI-related evidentiary standards.

9.4 Digital Personal Data Protection Act, 2023

The Digital Personal Data Protection Act, 2023[7] (DPDP Act) establishes a consent-based framework for personal data processing.

Deepfake systems frequently process:

- Facial data
- Voice samples
- Biometric identifiers
- Photographs and
- Behavioural information

Unauthorized use of such data for synthetic media generation may amount to unlawful data processing.

[6] Bharatiya Nyaya Sanhita, 2023.
[7] Digital Personal Data Protection Act, 2023.

However, the DPDP Act has limitations in addressing deepfakes because

- Publicly available data may still be used for AI training
- The Act does not specifically regulate AI-generated identity replication, and
- Enforcement mechanisms remain underdeveloped.

9.5 Information Technology Rules, 2021

The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021[8] impose due diligence obligations on intermediaries and social media platforms.

Platforms are required to-

- Remove unlawful content upon notice
- Maintain grievance redressal mechanisms
- Exercise due diligence, and
- Assist law enforcement authorities.

Government advisories have increasingly emphasized platform responsibility for addressing AI-generated misinformation and deepfake content.

Nevertheless, concerns remain regarding-

- Vague takedown obligations
- Over-censorship by intermediaries
- Lack of procedural safeguards, and
- Threats to online free expression.

10. Regulatory Challenges in Governing Deepfakes

10.1 Absence of a Clear Legal Definition

One of the major challenges in India is the absence of a statutory definition of "deepfake" or "synthetic media." Without definitional clarity, enforcement agencies face difficulties in identifying prohibited conduct.

10.2 Technological Sophistication

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

Deepfake technology evolves rapidly, often outpacing detection mechanisms. AI-generated content has become increasingly realistic, making forensic verification difficult.

10.3 Cross-Border Enforcement Problems

Deepfake creators may operate anonymously across jurisdictions, complicating investigation and prosecution.

Digital platforms hosting harmful content may also be headquartered outside India.

10.4 Electoral and Democratic Risks

Political deepfakes threaten electoral integrity by spreading fabricated speeches, manipulated interviews, and misinformation campaigns.

Such content can distort democratic discourse and reduce public trust in institutions.

10.5 Chilling Effect on Free Speech

Broad or vague regulation may encourage social media platforms to remove lawful content merely to avoid legal liability. In practical situations, platforms may prefer removing content even when the legality of such content is uncertain. This can negatively affect satire, parody, political criticism, documentaries, and other forms of legitimate expression.

This may adversely affect:

- Satire
- Artistic creativity
- Political criticism
- Journalism; and
- Academic research.

10.6 Gendered Impact of Deepfakes

Women are disproportionately targeted through non-consensual explicit deepfakes.

This raises important feminist and human rights concerns relating to:

- Digital consent
- Bodily autonomy

- Online violence; and
- Gender-based cyber abuse

11. Contemporary Deepfake Incidents and Their Impact in India

The discussion regarding deepfake regulation in India became more serious after multiple incidents involving manipulated videos and AI-generated content started circulating widely on social media platforms. These incidents demonstrated that deepfakes are no longer merely theoretical technological concerns but practical challenges affecting privacy, public trust, and digital governance.

One of the most discussed incidents involved manipulated videos of well-known Indian celebrities that went viral online. Such videos created confusion among viewers and raised concerns regarding consent, digital impersonation, and misuse of AI tools. Public debate intensified after several actors and public personalities expressed concern over the misuse of their facial identity and online likeness.

Apart from celebrity-related incidents, there has also been an increase in online fraud involving AI-generated voice cloning. Reports from cybersecurity agencies suggest that scammers have used cloned voices to impersonate family members, company executives, and public officials in order to deceive individuals and obtain money or sensitive information. These incidents indicate that deepfake technology can create both financial and emotional harm.

Another important concern relates to political misinformation. During election periods, manipulated videos and misleading edited content can spread rapidly across social media platforms. Even if such content is later identified as fake, the damage to public discourse may already have occurred. In a democratic society like India, where digital media plays an important role in shaping public opinion, the spread of AI-generated misinformation creates serious challenges for electoral integrity.

The issue also affects journalism and public trust in digital evidence. As deepfake technology becomes more sophisticated, people may begin doubting even authentic audio or video recordings. This phenomenon is sometimes referred to as the “liar’s dividend,” where genuine evidence can be dismissed as fabricated. Such a situation may weaken public confidence in media institutions and investigative reporting.

The growing number of deepfake incidents demonstrates that India requires both legal preparedness and technological awareness. Merely relying on traditional cyber law provisions may not be sufficient in the long run. Public awareness campaigns, digital literacy programs, and ethical AI practices are also necessary to address the social impact of synthetic media.

12. Comparative International Approaches

12.1 European Union

The European Union has adopted a rights-based and transparency-oriented approach toward AI regulation.

The EU AI Act[9] imposes obligations relating to:

- Transparency in AI-generated content
- Disclosure requirements for synthetic media
- Risk-based AI regulation, and
- Accountability for high-risk AI systems.

The Digital Services Act further strengthens platform accountability mechanisms.

The European model attempts to balance innovation with privacy and democratic safeguards.

12.2 United States

The United States follows a decentralized and speech-protective approach.

Several states have enacted laws targeting:

- Election-related deepfakes
- Non-consensual explicit deepfakes, and
- AI impersonation fraud.

However, the First Amendment significantly limits broad restrictions on synthetic media.

The U.S. approach therefore emphasizes targeted regulation rather than comprehensive federal control.

12.3 China

[9] European Union AI Act.

China has adopted one of the strictest deepfake regulatory models.

Chinese regulations require:

- Mandatory labelling of synthetic media
- Verification obligations for service providers
- User identity authentication, and
- Prevention of misinformation.

While effective in controlling harmful content, critics argue that China's approach risks excessive state surveillance and censorship.

13. Role of Social Media Platforms and Intermediaries

Social media platforms play a central role in the circulation and amplification of deepfake content. Since manipulated videos and AI-generated media can spread rapidly through online platforms, intermediaries have become important actors in the regulation process?

Platforms such as video-sharing websites, messaging applications, and social networking services influence how information reaches users. Deepfake content often gains visibility because of recommendation algorithms, reposting systems, and viral engagement patterns. As a result, questions regarding platform accountability have become increasingly significant.

Under the Information Technology Rules, 2021, intermediaries are required to exercise due diligence and remove unlawful content upon receiving valid notice. However, the practical implementation of these obligations remains challenging.

One major difficulty is identifying manipulated media accurately. Automated detection systems are not always reliable and may incorrectly remove legitimate content. Satirical videos, edited documentaries, or artistic works may sometimes be flagged as harmful content even when no malicious intention exists.

Another issue relates to the speed at which deepfake content spreads online. In many cases, harmful videos go viral before fact-checkers or platforms can respond effectively. By the time such content is removed, the reputational or political damage may already have occurred.

At the same time, imposing excessive liability on intermediaries may encourage over-censorship. Platforms may begin removing content aggressively in order to avoid legal consequences. Such an approach may affect free speech, journalism, and online political discussion.

Therefore, the role of intermediaries should be balanced carefully. Platforms should adopt transparency measures, improve detection technologies, cooperate with fact-checking agencies, and provide efficient grievance redressal systems. However, regulatory obligations should remain proportionate and should not convert private platforms into excessive censorship authorities.

Digital literacy also plays an important role. Users should be educated regarding AI-generated misinformation and methods for identifying manipulated media. A combination of platform responsibility, public awareness, and legal safeguards may provide a more effective solution than strict censorship alone.

14. Need for a Dedicated Deepfake Regulatory Framework in India

The existing legal framework in India provides only limited protection and does not fully address the practical and technological challenges created by generative AI systems.

A dedicated deepfake law should incorporate the following features

14.1 Statutory Definition of Deepfakes

The law should clearly define:

- Deepfakes
- Synthetic media
- AI-generated impersonation; and
- Manipulated digital content.

The definition should distinguish harmful malicious content from legitimate parody, satire, and artistic expression.

14.2 Consent-Based Protection Framework

The unauthorized use of a person’s:

- Face
- Voice
- Image
- Biometric identifiers; or
- Behavioural attributes
- Should require explicit consent.

Special protections should be introduced for minors and vulnerable individuals.

14.3 Criminal and Civil Liability

Separate offences should be created for:

- Non-consensual explicit deepfakes
- Election manipulation
- Financial fraud
- Identity theft, and
- Deepfake-enabled cyber harassment.

Victims should also have access to:

- Compensation
- Injunctions
- Expedited takedown mechanisms, and
- Rehabilitation support.

14.4 Platform Accountability and Safe Harbour Reform

Social media intermediaries should be required to:

- Implement AI-content detection systems
- Provide transparent complaint mechanisms
- Label synthetic media where feasible, and
- Respond promptly to verified complaints.

However, intermediary liability should remain proportionate to avoid excessive censorship.

14.5 Transparency and Watermarking Requirements

AI-generated media should include:

- Digital watermarking
- Metadata indicators, and
- Authenticity disclosures.

Transparency measures may reduce misinformation and improve digital trust.

14.6 Independent AI Regulatory Authority

India may consider establishing an independent AI regulatory body responsible for:

- AI governance standards
- Ethical AI guidelines
- Auditing high-risk AI systems
- Coordinating with technology companies; and
- Protecting constitutional rights.

15. Ethical Concerns and Societal Impact of Deepfakes

Apart from legal concerns, deepfake technology also raises important ethical and social issues. The misuse of synthetic media affects not only individual victims but also public trust and democratic culture.

One major ethical concern relates to consent. Deepfake technology allows a person's face, voice, or identity to be replicated without permission. In many situations, individuals may not even be aware that their digital likeness is being used in manipulated content. This creates serious concerns regarding autonomy and personal dignity.

The issue becomes more severe in cases involving non-consensual explicit content. Women are disproportionately targeted through AI-generated sexualized media, which may lead to emotional distress, reputational damage, and online harassment. Such misuse reflects broader concerns regarding gender-based cyber violence and digital exploitation.

Another ethical concern involves misinformation and erosion of trust. In democratic societies, citizens rely upon audio-visual content for news, political communication, and public debate. If manipulated media becomes increasingly difficult to identify, public trust in digital information may weaken significantly.

Deepfakes also create problems in the context of criminal justice and evidence law. Audio or video recordings have traditionally been treated as strong forms of evidence. However, the rise of synthetic media may create uncertainty regarding authenticity. Courts and investigative agencies may face challenges in verifying digital evidence in future cases.

From a societal perspective, the misuse of deepfake technology may contribute to fear, confusion, and polarization. Viral misinformation often spreads faster than corrective information. In emotionally sensitive situations, manipulated content may disturb public order or increase social tensions.

At the same time, it is important not to treat all AI-generated content as inherently harmful. Ethical regulation requires a balanced understanding of both the risks and benefits associated with generative AI. Educational tools, accessibility technologies, language translation systems, and creative media applications also rely upon similar AI techniques.

Therefore, ethical governance should focus on responsible innovation rather than blanket prohibition. Transparency, accountability, informed consent, and public awareness should form the foundation of any long-term regulatory strategy.

16. Balancing Innovation, Privacy and Freedom of Speech

The regulation of deepfakes requires a nuanced constitutional approach. In my view, the debate should not be limited to whether deepfakes should be banned or freely allowed. The real issue is how the law can distinguish between harmful misuse and legitimate innovation.

There is no doubt that deepfakes can cause serious harm, especially in cases involving privacy violations, misinformation, and non-consensual explicit content. At the same time, AI technology also has important social and economic benefits. Therefore, a balanced approach becomes necessary.

Instead of imposing broad censorship, Indian law should focus more on regulating harmful conduct, ensuring accountability, and protecting constitutional rights. In my opinion, technology itself should not be treated as the problem because the actual issue lies in its misuse.

An excessively restrictive framework may-

- Suppress innovation

- Discourage AI research
- Restrict artistic freedom; and
- Encourage over-censorship

Conversely, inadequate regulation may-

- Enable cybercrime
- Harm privacy and dignity
- Facilitate misinformation and
- Undermine democratic processes.

Therefore, India must adopt a proportional and rights-based framework grounded in constitutional morality.

The following principles should guide regulation:

(a**) Principle of Proportionality**

Restrictions on synthetic media must satisfy constitutional proportionality standards.

(**b) Harm-Based Regulation**

Regulation should target malicious and harmful uses rather than the technology itself.

(**c) Transparency and Accountability**

AI developers and platforms must remain accountable for harmful synthetic content.

(d) Judicial Oversight

Content removal and enforcement mechanisms should remain subject to procedural safeguards and judicial review.

(e) Technological Neutrality

The law should remain adaptable to future technological developments. Since AI systems evolve rapidly, a rigid legal framework may become outdated within a short period of time.

17. Conclusion

Deepfake technology represents one of the most significant regulatory challenges of the digital age. While synthetic media offers transformative possibilities in entertainment, communication, accessibility, and innovation, it simultaneously threatens privacy, dignity, democratic integrity, and public trust.

India's current legal framework—comprising the Information Technology Act, the Bharatiya Nyaya Sanhita, the Digital Personal Data Protection Act, and intermediary regulations—provides only fragmented and indirect protection against deepfake-related harms. The absence of a dedicated statutory framework has created significant legal uncertainty and enforcement challenges.

At the constitutional level, deepfake regulation requires careful balancing between the right to privacy under Article 21 and the freedom of speech and expression under Article 19(1) (a). Any regulatory model that prioritizes censorship over constitutional freedoms risks undermining democratic values and digital innovation.

India therefore requires a comprehensive and technologically adaptive legal framework specifically addressing synthetic media and generative AI. Such a framework must combine:

- Consent-based protections
- Criminal and civil remedies
- Platform accountability
- Transparency obligations
- Procedural safeguards and
- Constitutional proportionality.

In the end, the purpose of regulation should not be to stop technological growth. Rather, the law should ensure that innovation develops in a balanced manner that respects privacy, dignity, democratic values, and constitutional freedoms. A practical and rights-oriented framework will help India deal with the risks associated with deepfakes while still encouraging responsible technological development.

18. References

❖ **Books and Journal Articles**

1. Chesney, Robert & Citron, Danielle, *"Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security,"* 107 California Law Review (2019).

2. Citron, Danielle, *The Fight for Privacy: Protecting Dignity, Identity, and Love in the Digital Age* (W. W. Norton & Company, 2022).
3. . Kietzmann, Jan et al., *"Deepfakes: Trick or Treat*?" 64 Business Horizons 135 (2020).
4. Westerlund, Mika, *"The Emergence of Deepfake Technology: A Review,*" 9 Technology Innovation Management Review 39 (2019).
5. Vaccari, Cristian & Chadwick, Andrew, "*Deepfakes and Disinformation,"* 49 Social Media + Society (2020).
6. . Agarwal, Shweta*, "Regulating AI-Generated Deepfakes in India,*" Indian Journal of Law and Technology (2024).
7. Paris, Britt & Donovan, Joan, *"Deepfakes and Cheap Fakes,*" Data & Society Research Institute (2019).

❖ **Statutes and Legislations**

1. *Constitution of India.*
2. *Information Technology Act,* 2000.
3. *Bharatiya Nyaya Sanhita,* 2023.
4. *Digital Personal Data Protection Act,* 2023.
5. *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules,* 2021.
6. *European Union AI Act*.
7. *Digital Services Act (European Union).*

❖ **Cases**

1. *Justice K. S. Puttaswamy v. Union of India,* (2017) 10 SCC 1.
2. *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.
3. *Subramanian Swamy v. Union of India*, (2016) 7 SCC 221.

❖ **Web Source**

1. UNESCO, *"Guidance for Regulating Digital Platforms."*
2. World Economic Forum, *"The Global Risks of Deepfake Technology."*
3. European Commission, *"Artificial Intelligence Act."*
4. arXiv *Research Papers on Deepfake Detection and Governance.*
5. *Ministry of Electronics and Information Technology (MeitY) Advisories on AI and Deepfake.*