DEEPFAKES, MISINFORMATION AND CRIMINAL LIABILITY UNDER INDIAN LAW

Anushka Keshri*

ABSTRACT

The emergence of deepfake technology has significantly transformed the digital landscape by enabling the creation of highly realistic yet fabricated audio, video and visual content through artificial intelligence. While such technology possesses legitimate applications in entertainment, education and media production, its increasing misuse for misinformation, impersonation, political manipulation, financial fraud and non-consensual explicit content has raised serious legal and ethical concerns across jurisdictions. In India, the rapid circulation of deepfake content through social media platforms and digital communication channels has exposed substantial gaps within the existing legal framework governing cybercrime, privacy and digital accountability. This article critically examines the concept of deepfakes and analyses the extent to which current Indian laws address the criminal and civil liabilities arising from their misuse. It evaluates the applicability of the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023, data protection principles and constitutional jurisprudence relating to privacy and freedom of speech. The article further explores the challenges associated with identifying perpetrators, regulating intermediary platforms and balancing technological innovation with individual rights in the digital era. Through a comparative analysis of international regulatory approaches, including the European Union and the United States, the article argues that India's existing legal mechanisms remain fragmented and insufficient to effectively combat the evolving threat posed by AI-generated misinformation. It concludes by suggesting the need for a comprehensive and specialised legal framework that ensures accountability, victim protection and responsible technological governance while safeguarding constitutional freedoms.

*BBA LLB, SECOND YEAR, BENNETT UNIVERSITY, GREATER NOIDA.

Keywords: Deepfake Technology, Information Technology Act, 2000, Bharatiya Nyaya Sanhita, 2023, European Union, AI-Generated Misinformation.

INTRODUCTION

The rapid advancement of artificial intelligence has fundamentally altered the manner in which information is created, disseminated and consumed in the digital era. Among the most concerning developments is the rise of "deepfake" technology — synthetic media capable of generating highly realistic yet manipulated audio, video and image content. By employing deep learning techniques and generative adversarial networks (GANs), deepfakes enable the fabrication of human expressions, speech patterns and visual appearances with remarkable accuracy, often making it difficult for ordinary individuals to distinguish fabricated content from authentic material.

The Indian digital ecosystem presents a particularly vulnerable environment for the spread of deepfake content. India possesses one of the world's largest internet user bases, with millions relying upon digital platforms as primary sources of information and communication. In such an environment, manipulated media possesses the potential to disrupt electoral processes, incite communal tensions, damage personal reputations and facilitate sophisticated forms of cybercrime.

The legal implications arising from deepfake technology extend across several branches of law, including cyber law, criminal law, constitutional law and data protection jurisprudence. However, India currently lacks a dedicated statutory framework specifically designed to regulate deepfake technology and AI-generated misinformation, resulting in reliance upon fragmented provisions dispersed across the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023 and intermediary liability regulations.

The regulation of deepfakes also presents significant constitutional tensions. On one hand, excessive state regulation of synthetic media may raise issues relating to freedom of speech and expression under Article 19(1)(a) of the Constitution of India. On the other hand, the unrestricted circulation of malicious deepfake content threatens the right to privacy recognised in Justice K.S. Puttaswamy v. Union of India as well as the broader principles of dignity, autonomy and informational self-determination.

Against this backdrop, this article critically analyses the criminal liability and regulatory challenges associated with deepfake technology under Indian law. It examines the existing statutory framework governing cyber offences, misinformation and privacy violations while assessing the limitations of current legal mechanisms. The article ultimately argues that the evolving nature of synthetic media necessitates a comprehensive and specialised legal framework capable of balancing innovation, accountability and constitutional protections in the digital age.

RESEARCH METHODOLOGY

This study adopts a doctrinal and analytical method of legal research to examine the challenges posed by deepfake technology within the Indian legal framework. The research is primarily based upon secondary sources, including statutes, judicial decisions, legal commentaries, academic journals, policy papers and government reports relating to artificial intelligence, cyber law, privacy rights and digital governance.

The study analyses the applicability of the Information Technology Act, 2000, the Bharatiya Nyaya Sanhita, 2023 and constitutional jurisprudence concerning privacy and freedom of speech. A comparative analysis of international approaches adopted by the European Union, the United States and China further assists in identifying global regulatory trends and potential reforms suitable for the Indian context. The scope of this research is limited to the legal and regulatory implications of deepfake technology in relation to criminal liability, privacy concerns, intermediary accountability and constitutional protections under Indian law.

RESEARCH OBJECTIVES

1. To examine the concept, evolution and functioning of deepfake technology and synthetic media.
2. To analyse constitutional and legal concerns associated with deepfakes, particularly in relation to privacy, dignity and freedom of speech.
3. To evaluate the adequacy of existing Indian laws in addressing offences arising from deepfake-related misuse.
4. To examine issues relating to criminal liability, intermediary accountability and digital evidence in deepfake cases.
5. To undertake a comparative analysis of international regulatory approaches governing AI-generated misinformation.

6. To suggest legal and institutional reforms for strengthening India's regulatory framework concerning deepfake technology.

RESEARCH QUESTIONS

1. What are deepfakes, and how do they impact digital communication and public trust?
2. Whether the existing Indian legal framework is sufficient to regulate deepfake-related offences and AI-generated misinformation?
3. How do deepfakes affect constitutional rights relating to privacy, dignity and freedom of speech?
4. Who should bear criminal liability for the creation and dissemination of harmful deepfake content?
5. What regulatory approaches have been adopted internationally to govern deepfake technology?
6. Does India require specialised legislation specifically addressing deepfake-related harms and synthetic media manipulation?

EXISTING INDIAN LEGAL FRAMEWORK GOVERNING DEEPFAKES

The regulation of AI-generated manipulated media in India presently relies upon a fragmented combination of cyber laws, criminal provisions, constitutional protections and intermediary regulations. Since no dedicated legislation specifically addressing deepfakes currently exists, courts and investigative authorities must interpret traditional legal provisions to address harms arising from synthetic media. A single deepfake incident may involve identity theft, cyber impersonation, defamation, privacy violations, obscenity, misinformation and financial fraud, requiring reliance upon several overlapping statutory provisions rather than a unified framework.

Information Technology Act 2000: The Information Technology Act 2000 constitutes the primary legislation governing cyber-related offences and electronic communication in India. Although enacted before modern AI technologies, several provisions may apply to deepfake-related misconduct.

Section 66C: Identity Theft: Section 66C criminalises fraudulent or dishonest use of electronic signatures, passwords and unique identification features belonging to another person. When manipulated content falsely represents another individual for deceptive

purposes, it may constitute identity theft. However, the provision was not drafted with AI-generated biometric replication in mind, raising interpretational ambiguities as to whether facial likeness and voice cloning constitute "unique identification features" under the statute.

Section 66D: Cheating by Personation: Section 66D penalises cheating by personation using computer resources or communication devices. This provision applies in cases involving AI-generated impersonation for financial fraud, phishing scams or deceptive communication. However, it primarily addresses fraudulent intent and financial deception rather than broader harms such as misinformation, reputational damage or political manipulation.

Section 67: Obscene Electronic Content: Section 67 criminalises the publication and transmission of obscene material in electronic form, becoming relevant in cases involving non-consensual explicit deepfake content. However, it primarily addresses obscenity rather than consent-based identity manipulation and does not fully capture broader privacy and dignity violations arising from synthetic media exploitation.

Section 69A: Blocking Powers and Intermediary Obligations: Section 69A empowers the government to block public access to electronic information on grounds of sovereignty, security and public order, potentially restricting harmful deepfake content. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules further impose due diligence obligations upon digital platforms. However, concerns regarding transparency, procedural safeguards and arbitrary censorship remain significant, as noted by the Supreme Court in Shreya Singhal v. Union of India.

Bharatiya Nyaya Sanhita, 2023: The Bharatiya Nyaya Sanhita (BNS), 2023, replacing the Indian Penal Code, contains several provisions potentially applicable to deepfake offences. Provisions relating to defamation may be invoked where AI-generated videos falsely depict individuals engaging in criminal or immoral conduct. Provisions concerning forgery and fabrication of false electronic records may apply where deepfakes involve deceptive digital representations intended to mislead viewers. Cheating, fraud and impersonation provisions may address financial scams conducted through AI-generated voice or visual identities. Offences relating to public mischief may apply where deepfake misinformation provokes communal tension or social disorder.

Nevertheless, traditional offences within the BNS were designed primarily around direct human conduct. Establishing authorship and intent in algorithmically generated content, therefore, presents significant interpretational and evidential challenges.

Digital Personal Data Protection: The Digital Personal Data Protection framework becomes relevant because deepfakes involve the manipulation of identifiable personal data, including photographs, facial structures and voice samples, without informed consent. The principles of lawful processing, informed consent and purpose limitation are particularly important where AI systems are trained using publicly available digital content. However, existing data protection mechanisms remain insufficient because AI models utilise massive cross-jurisdictional datasets, individuals are often unaware of data extraction, and enforcement mechanisms remain underdeveloped.

Limitations of Existing Indian Laws: The Indian legal framework suffers from several critical structural limitations in addressing deepfake-related harms:

Absence of Deepfake-Specific Legislation: India lacks a comprehensive statutory framework specifically regulating synthetic media, and existing provisions fail to address modern technological realities adequately.

Fragmented Legal Approach: Deepfake harms are dispersed across cyber law, criminal law, privacy law and intermediary regulation, creating uncertainty regarding enforcement, liability and procedural consistency.

Attribution Difficulties: Determining responsibility is extremely difficult where anonymous creators, automated AI systems, foreign platforms and cross-border dissemination are involved.

Technological and Investigative Challenges: Law enforcement agencies frequently lack advanced digital forensic infrastructure, AI detection capabilities and rapid response mechanisms.

Jurisdictional Complexities: Deepfake content may be created in one country, hosted in another and consumed globally, creating substantial procedural challenges for national enforcement agencies.

Constitutional Constraints: Any regulatory response must remain compatible with freedom of speech, artistic expression and technological innovation under Article 19(1)(a), making broad criminalisation constitutionally problematic.

CRIMINAL LIABILITY AND ACCOUNTABILITY IN DEEPFAKE CASES

Liability of Deepfake Creators: Primary criminal responsibility in most deepfake cases rests upon the individual directly involved in creating manipulated content. Where deepfake content is created for fraudulent, defamatory or exploitative purposes, criminal liability may arise under provisions relating to impersonation, identity theft, obscenity, defamation and cyber fraud. The principle of mens rea plays a central role — proving that the accused knowingly created deceptive or harmful content is essential for prosecution. However, the increasing automation of AI tools complicates traditional concepts of intent, as software can generate manipulated media with minimal human decision-making. Identifying creators is further complicated by anonymous dissemination through encrypted platforms and fake accounts operating across jurisdictions.

Liability of Distributors and Forwarders: Criminal liability may extend to individuals who knowingly distribute or amplify harmful deepfake content. A fabricated video shared by a small number of individuals may rapidly reach millions, causing irreversible reputational or political damage. Liability may depend upon knowledge of falsity, malicious intention, degree of participation and resulting harm. However, excessive liability for ordinary users may produce chilling effects on online communication. Many individuals unknowingly share manipulated content, and criminal law must carefully distinguish between deliberate dissemination and innocent forwarding.

Intermediary Liability and Platform Responsibility: Intermediary platforms occupy a central position in the deepfake ecosystem because they facilitate large-scale dissemination of digital content. Under the Information Technology Act and the Intermediary Guidelines, platforms are generally protected from direct liability provided they comply with due diligence obligations and remove unlawful material upon receiving actual knowledge or governmental direction. Deepfake regulation presents substantial challenges for intermediaries because automated moderation systems remain imperfect, encrypted communications limit content monitoring, and distinguishing malicious misinformation from satire or legitimate commentary is technologically difficult. The constitutional principles in Shreya Singhal v. Union of India

caution against imposing excessive intermediary obligations that may encourage arbitrary censorship.

Liability of AI Developers and Technology Providers: The growing sophistication of synthetic media tools has generated important questions regarding the liability of AI developers and technology providers. While AI systems do not possess legal personality or criminal intent, concerns arise where developers knowingly create tools facilitating harmful misuse without incorporating reasonable safeguards. Questions regarding negligent design, failure to implement safety protocols and reckless technological deployment may become increasingly relevant. However, imposing broad criminal liability upon technology developers may discourage innovation and hinder legitimate AI research. Future legal frameworks may need to establish clearer standards regarding due diligence obligations for synthetic media developers.

Evidentiary and Jurisdictional Challenges: Deepfake technology fundamentally undermines traditional evidentiary assumptions regarding the reliability of audio and visual recordings. Digital forensic examination, analysing metadata, pixel inconsistencies, voice irregularities and algorithmic signatures, has become essential in deepfake investigations. However, forensic technologies continue struggling against rapidly evolving AI systems, creating a continuous technological arms race. Jurisdictional complexity further complicates enforcement: deepfake content may be created in one jurisdiction, hosted on foreign servers and disseminated globally within minutes. Law enforcement agencies frequently face shortages of specialised AI expertise, insufficient forensic infrastructure and delayed procedural responses, allowing harmful content to spread extensively before legal remedies become effective.

REGULATORY CHALLENGES AND NEED FOR REFORM

Rapid Evolution of AI Technology: One of the most significant regulatory challenges associated with deepfakes is the extraordinary speed at which artificial intelligence continues to evolve. Traditional legislative processes often struggle to adapt quickly enough to technological innovation. The growing accessibility of AI tools further intensifies this challenge, as deepfake generation applications are now widely available through mobile software and open-source platforms. Any effective legal framework must therefore remain technologically adaptive and flexible rather than narrowly drafted around presently existing capabilities.

Balancing Regulation with Constitutional Freedoms: The regulation of deepfake content must remain consistent with the constitutional guarantee of freedom of speech and expression under Article 19(1)(a). AI-generated media possesses legitimate applications in entertainment, satire, filmmaking and education; consequently, excessive regulation may impose disproportionate restrictions upon creative expression and digital communication. However, Article 19(2) permits reasonable restrictions where harmful deepfake content spreads misinformation, defames individuals or disrupts public order. The constitutional challenge lies in distinguishing legitimate synthetic expression from malicious digital deception. As affirmed in Shreya Singhal v. Union of India, restrictions upon online speech must remain narrowly tailored and constitutionally proportionate.

International Comparative Perspective: Several jurisdictions have begun formulating specialised regulatory frameworks addressing synthetic media and AI-generated deception. The European Union's AI Act adopts a risk-based approach requiring transparency and disclosure obligations for AI-generated content. Various United States state legislatures have enacted targeted statutes addressing non-consensual deepfake content and election-related misinformation, reflecting a free-speech-conscious model. China has adopted stricter state-centric regulations mandating disclosure of synthetic media and platform compliance obligations. These comparative developments demonstrate growing global recognition of the dangers posed by deepfake technology and offer valuable insights for Indian legislative reform.

PROPOSED REFORMS AND RECOMMENDATIONS

A comprehensive deepfake regulatory framework for India should incorporate the following elements:

Mandatory Disclosure and Watermarking: AI-generated synthetic content should contain mandatory disclosure labels or digital watermarking mechanisms enabling users to identify manipulated media. Such transparency requirements may significantly reduce misinformation risks while preserving legitimate creative applications of artificial intelligence.

Specialised Deepfake Legislation: A dedicated statutory framework should include a clear definition of deepfakes and synthetic media, graduated criminal penalties calibrated to the severity of harm, specific provisions addressing non-consensual explicit content, election-related misinformation and financial impersonation fraud, and procedural safeguards protecting satire, parody and legitimate artistic expression.

Strengthened Platform Due Diligence: Intermediary platforms should implement rapid grievance redressal systems, victim reporting mechanisms, AI-based content detection tools and content authenticity verification systems. Such obligations must remain proportionate and constitutionally compliant, avoiding excessive censorship obligations.

Cyber Forensic Infrastructure: India should establish specialised AI forensic laboratories and cyber investigation units trained in deepfake detection and digital evidence authentication, addressing the current institutional gap in investigative capacity.

Data Protection Safeguards: Stronger safeguards should regulate biometric data collection, facial recognition systems and AI training datasets to protect against unauthorised identity manipulation and ensure consent-based usage of personal information.

International Cooperation: India should actively participate in international discussions concerning AI governance, digital authenticity standards and cross-border cyber enforcement cooperation to address the transnational dimensions of deepfake dissemination effectively.

Public Awareness and Digital Literacy: Public education initiatives are essential for combating misinformation and improving digital literacy. Citizens must be educated regarding the identification of manipulated media, verification techniques and the risks associated with synthetic misinformation.

CONCLUSION

The rapid advancement of artificial intelligence has fundamentally transformed the nature of digital communication and information dissemination within contemporary society. Deepfake technology has emerged as one of the most concerning consequences of this evolution, enabling the creation of highly realistic fabricated content capable of misleading individuals, damaging reputations and undermining public trust. While synthetic media possesses legitimate applications in entertainment, education and innovation, its increasing misuse for misinformation, impersonation, cyber fraud and non-consensual exploitation has generated serious legal and constitutional concerns.

This article has demonstrated that deepfake technology creates multifaceted legal challenges affecting privacy rights, informational autonomy, dignity, democratic integrity and digital authenticity. The analysis of constitutional principles reveals that deepfake regulation exists at

the intersection of competing rights. On one hand, manipulated media may violate the right to privacy and dignity under Justice K.S. Puttaswamy v. Union of India. On the other hand, excessive governmental regulation may interfere with freedom of speech and expression under Article 19(1)(a). The constitutional challenge lies in formulating a balanced legal response capable of preventing harmful digital deception without suppressing legitimate expression.

The existing statutory provisions under the Information Technology Act, 2000 and the Bharatiya Nyaya Sanhita, 2023, while partially applicable, remain fragmented, technologically outdated and inadequate when applied to modern AI-generated synthetic media. The absence of a dedicated statutory framework creates significant uncertainty regarding criminal liability, intermediary responsibility, evidentiary standards and enforcement procedures. Determining accountability becomes particularly complex where manipulated content is created anonymously, disseminated through global digital networks and amplified by intermediary platforms operating across jurisdictions.

Comparative international experience — from the European Union's transparency-focused AI Act to targeted United States state-level legislation and China's disclosure-based model — demonstrates growing global recognition of the dangers posed by synthetic media, while also highlighting that no single regulatory model is entirely sufficient given the rapidly evolving and transnational nature of deepfake technology.

India, therefore, requires a comprehensive and specialised legal framework specifically addressing deepfake-related harms. Such regulation must include clear statutory definitions, graduated criminal penalties for malicious misuse, mandatory disclosure obligations, strengthened intermediary due diligence standards, robust biometric data protection safeguards, expedited victim redressal mechanisms and enhanced cyber forensic capabilities. At the same time, future legislation must remain constitutionally proportionate, adopting a balanced approach grounded in transparency, accountability and procedural fairness so as not to suppress free expression, satire, artistic creativity or legitimate technological development.

Ultimately, deepfake technology represents a defining challenge of the artificial intelligence era because it directly affects the relationship between truth, identity and digital trust within modern society. As synthetic media continues evolving, the legal system must proactively adapt in order to safeguard constitutional rights, democratic stability and informational

integrity. The future of digital governance will depend upon the ability of legal institutions to balance innovation with accountability in an increasingly AI-driven world.

REFERENCES

1. Constitution of India, 1950.
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 Artificial Intelligence Act.
7. General Data Protection Regulation (GDPR), European Union.
8. Justice K.S. Puttaswamy v. Union of India.
9. Shreya Singhal v. Union of India.
10. Anuradha Bhasin v. Union of India.
11. Talat Fatema, Cyber Crimes (Eastern Book Company).
12. Justice Yatindra Singh, Cyber Laws.
13. Chris Reed and John Angel, Computer Law.
14. Stuart Russell and Peter Norvig, Artificial Intelligence: A Modern Approach.
15. Richard Susskind, Online Courts and the Future of Justice.
16. Chesney, Robert and Citron, Danielle, "Deepfakes and the New Disinformation War."
17. Hao Li, "Deepfake Technology: Challenges and Legal Responses."
18. Danielle Citron, "The Fight Against Digital Impersonation."
19. Articles relating to AI governance, misinformation and cyber regulation published in legal journals and technology law reviews.
20. NITI Aayog Reports on Artificial Intelligence.
21. OECD Principles on Artificial Intelligence.
22. UNESCO Recommendations on the Ethics of Artificial Intelligence.
23. European Commission Reports on AI Regulation.
24. Government of India reports relating to cybersecurity and digital governance.
25. Indian Kanoon — www.indiankanoon.org
26. SCC Online — www.scconline.com
27. Manupatra — www.manupatra.com

28. Ministry of Electronics and Information Technology (MeitY) — www.meity.gov.in
29. European Commission AI Policy Portal — digital-strategy.ec.europa.eu
30. OECD AI Policy Observatory — oecd.ai