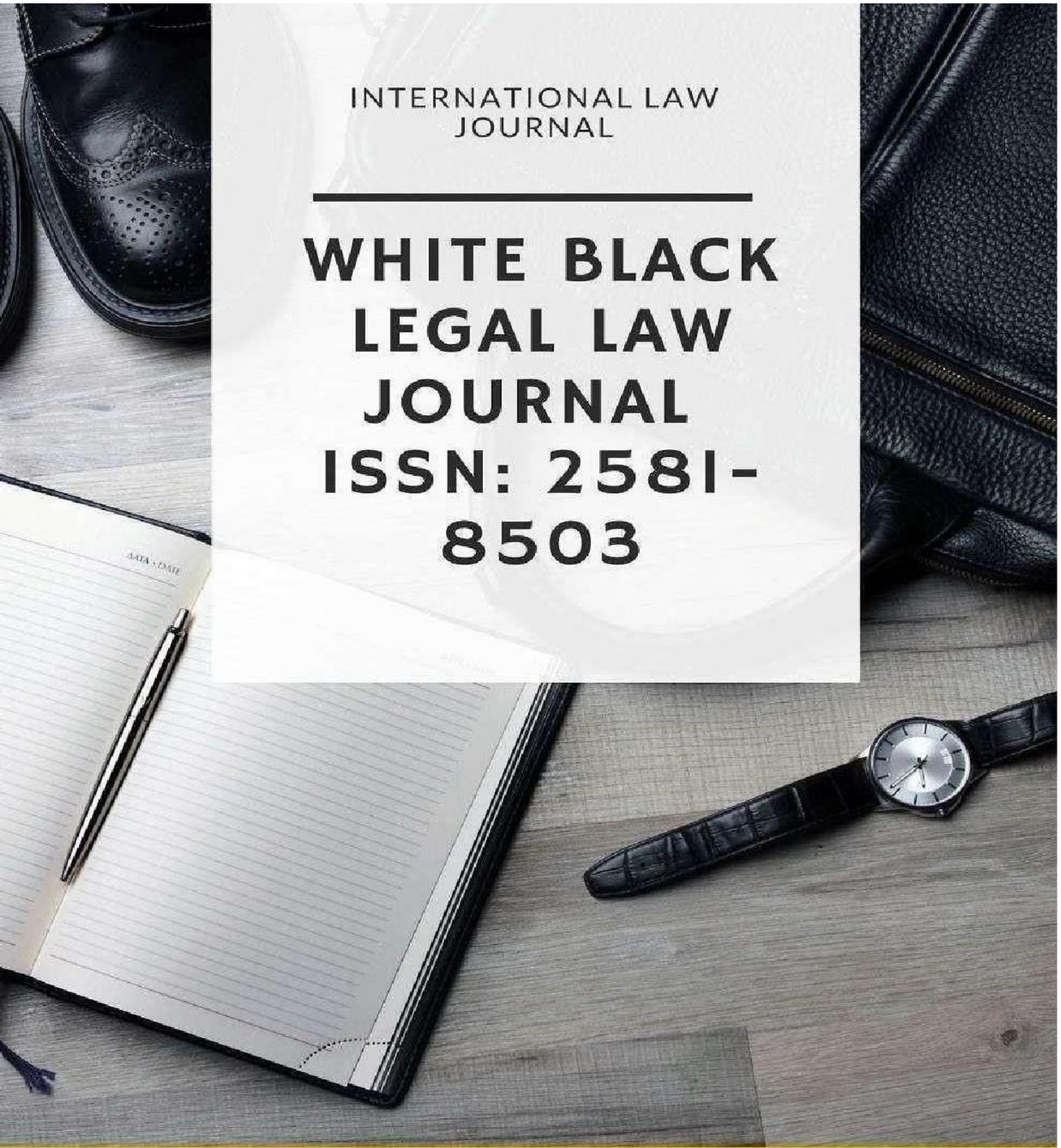

INTERNATIONAL LAW JOURNAL

WHITE BLACK LEGAL LAW JOURNAL

ISSN: 2581-8503

Peer - Reviewed & Refereed Journal

The Law Journal strives to provide a platform for discussion of International as well as National Developments in the Field of Law.

WWW.WHITEBLACKLEGAL.CO.IN

DISCLAIMER

The views, opinions, interpretations, and conclusions expressed in the articles are exclusively those of the respective authors. They do not represent or reflect the views of the Editorial Board, Editors, Reviewers, Advisors, Publisher, or Management of *White Black Legal*.

While reasonable efforts are made to ensure academic quality and accuracy through editorial and peer-review processes, *White Black Legal* makes no representations or warranties, express or implied, regarding the completeness, accuracy, reliability, or suitability of the content published. The journal shall not be liable for any errors, omissions, inaccuracies, or consequences arising from the use, interpretation, or reliance upon the information contained in this publication.

The content published in this journal is intended solely for academic and informational purposes and shall not be construed as legal advice, professional advice, or legal opinion. *White Black Legal* expressly disclaims all liability for any loss, damage, claim, or legal consequence arising directly or indirectly from the use of any material published herein.

ABOUT WHITE BLACK LEGAL

White Black Legal – The Law Journal is an open-access, peer-reviewed, and refereed legal journal established to provide a scholarly platform for the examination and discussion of contemporary legal issues. The journal is dedicated to encouraging rigorous legal research, critical analysis, and informed academic discourse across diverse fields of law.

The journal invites contributions from law students, researchers, academicians, legal practitioners, and policy scholars. By facilitating engagement between emerging scholars and experienced legal professionals, *White Black Legal* seeks to bridge theoretical legal research with practical, institutional, and societal perspectives.

In a rapidly evolving social, economic, and technological environment, the journal endeavours to examine the changing role of law and its impact on governance, justice systems, and society. *White Black Legal* remains committed to academic integrity, ethical research practices, and the dissemination of accessible legal scholarship to a global readership.

AIM & SCOPE

The aim of *White Black Legal – The Law Journal* is to promote excellence in legal research and to provide a credible academic forum for the analysis, discussion, and advancement of contemporary legal issues. The journal encourages original, analytical, and well-researched contributions that add substantive value to legal scholarship.

The journal publishes scholarly works examining doctrinal, theoretical, empirical, and interdisciplinary perspectives of law. Submissions are welcomed from academicians, legal professionals, researchers, scholars, and students who demonstrate intellectual rigour, analytical clarity, and relevance to current legal and policy developments.

The scope of the journal includes, but is not limited to:

- Constitutional and Administrative Law
- Criminal Law and Criminal Justice
- Corporate, Commercial, and Business Laws
- Intellectual Property and Technology Law
- International Law and Human Rights
- Environmental and Sustainable Development Law
- Cyber Law, Artificial Intelligence, and Emerging Technologies
- Family Law, Labour Law, and Social Justice Studies

The journal accepts original research articles, case comments, legislative and policy analyses, book reviews, and interdisciplinary studies addressing legal issues at national and international levels. All submissions are subject to a rigorous double-blind peer-review process to ensure academic quality, originality, and relevance.

Through its publications, *White Black Legal – The Law Journal* seeks to foster critical legal thinking and contribute to the development of law as an instrument of justice, governance, and social progress, while expressly disclaiming responsibility for the application or misuse of published content.

BEYOND REALITY: REGULATING AI-GENERATED DEEPFAKES IN INDIA TO SAFEGUARD PRIVACY, DEMOCRACY AND DIGITAL TRUST

AUTHORED BY - SHAIK NASEERUDDIN BASHA
B.A. LL.B. 5th Year
Jagarlamudi Chandramouli College of Law, Guntur, Andhra Pradesh

ABSTRACT

Artificial intelligence has fundamentally reshaped digital communication through the rise of highly realistic synthetic media, commonly known as AI-generated deepfakes. This technological shift has created a parallel digital reality where identities, voices, and actions can be convincingly fabricated, raising urgent serious for privacy, democratic integrity, financial security, and public trust. he rapid growth and easy accessibility of these tools have further intensified legal challenges, making existing cyber laws increasingly inadequate to address such emerging risks.

This study examines whether India's existing legal framework is sufficient to address harms arising from AI-generated deepfakes. It follows a doctrinal and comparative approach by analysing constitutional provisions, statutory laws, criminal legislation, data protection rules, evidentiary standards, intermediary obligations, judicial decisions, and relevant international frameworks. The study also considers regulatory developments in the European Union, China, the United States, and other global instruments to understand how different jurisdictions are responding to synthetic media governance.

The analysis indicates that India's legal framework is fragmented across multiple statutes originally designed to address traditional cyber offences. While these laws offer limited remedies, they do not clearly deal with AI-generated content or assign specific responsibility for its creation, manipulation or dissemination. This results in gaps in key areas such as content authentication, platform accountability, evidentiary reliability and AI-specific liability. Comparative developments also reflect a broader global shift towards preventive and technology-responsive regulatory approaches.

The analysis suggests that India must move beyond a fragmented legal response and adopt a dedicated, technology-specific framework for governing AI-generated deepfakes. It recommends the enactment of a Deepfake Regulation Act supported by clear institutional mechanisms, mandatory authentication standards, platform obligations and international cooperation to ensure legal certainty, digital safety and responsible development of artificial intelligence.

Keywords:
AI-Generated Deepfakes, Synthetic Media, Digital Identity, Privacy, Platform Accountability, Cyber Law, Artificial Intelligence Regulation, Comparative Legal Framework.

1. INTRODUCTION

Artificial intelligence has become an integral part of today's digital ecosystem, enabling the rapid creation and widespread dissemination of digital content across multiple sectors. One of its most significant developments is generative artificial intelligence, which enables the creation of highly realistic synthetic media, commonly known as deepfakes. While this technology offers useful applications in fields such as education, healthcare, entertainment, media and business, its increasing misuse for identity theft, financial fraud, cybercrime, non-consensual intimate content and political misinformation has emerged as a serious socio-legal concern. The increasing use of deepfakes has raised concerns relating to privacy, digital identity, democratic processes and public trust in online information.

Although several jurisdictions have introduced specialised regulatory measures to address AI-generated synthetic media, India continues to rely on existing constitutional, criminal, cyber and data protection laws to address such misuse. The absence of a dedicated statutory framework has created uncertainty in regulating emerging forms of AI-enabled harm. Against this backdrop, this study examines India's existing regulatory approach, analyses comparative international developments and suggests reforms for developing a comprehensive and technology-responsive framework governing AI-generated deepfakes.

1.1 Research Problem

The rapid growth of AI-generated deepfakes has exposed significant regulatory gaps in India's existing legal system. While current laws address offences such as cybercrime, identity theft, misinformation and privacy violations, they were not specifically designed to regulate synthetic media or clearly assign responsibility for its creation and dissemination. The absence of technology-specific definitions, platform obligations, evidentiary standards and coordinated enforcement mechanisms has reduced the overall effectiveness of the existing legal framework. This study therefore examines whether India's current regulatory approach is sufficient to address the legal challenges posed by AI-generated deepfakes.

1.2 Research Objectives

The study aims to examine India's existing legal framework governing AI-generated deepfakes and evaluate its effectiveness in addressing legal issues arising from synthetic media. It further analyses constitutional principles, statutory provisions, judicial developments and comparative international approaches to identify regulatory gaps and suggest practical legal reforms for strengthening AI governance in India.

1.3 Research Methodology

The study adopts a doctrinal research methodology based on qualitative legal analysis. Primary sources such as constitutional provisions, statutes, judicial decisions, government reports, policy documents and international instruments have been examined, along with secondary sources including books, journal articles and research papers. A comparative legal approach is also used to analyse regulatory models followed in the European Union, China, and the United States and to identify relevant best practices for India.

2. UNDERSTANDING AI-GENERATED DEEPFAKES

The emergence of generative artificial intelligence has significantly changed the way digital content is created and shared. One of its most important developments is AI-generated deepfakes, which enable the creation of highly realistic images, audio, videos and text that closely resemble real individuals. The term *deepfake*, derived from "deep learning" and "fake", refers to digitally created content produced using advanced artificial intelligence techniques. By analysing existing images, videos or voice samples, AI systems can swap faces, clone voices, synchronise lip movements or generate entirely new digital content that appears authentic.

Unlike conventional digital editing, deepfakes generate highly realistic synthetic content capable of closely replicating human identity and behaviour. These capabilities have created valuable opportunities in fields such as education, healthcare, entertainment, filmmaking, journalism, accessibility and digital communication. However, the increasing accessibility of AI-powered tools has also made the creation of manipulated content faster, cheaper and easier, significantly increasing the risk of misuse and harmful applications.

The dual-use nature of deepfake technology makes it one of the most significant governance challenges of the digital era. While it supports innovation and creative expression, it also enables identity manipulation, financial fraud, misinformation, digital impersonation and other forms of online abuse. This evolving intersection between artificial intelligence and digital communication highlights the need for a balanced regulatory approach that encourages innovation while safeguarding individual rights and the integrity of the digital ecosystem.

3. SOCIO-LEGAL CHALLENGES POSED BY AI-GENERATED DEEPFAKES

The increasing use of AI-generated deepfakes has expanded the scope of digital harm beyond traditional cyber offences. By enabling the creation of highly realistic synthetic content, deepfakes have changed the nature of privacy violations, financial fraud, political misinformation and digital communication. Their growing accessibility has raised serious concerns about the authenticity of online content, making synthetic media a significant socio-legal issue that requires careful legal and policy attention.

3.1 Privacy, Identity and Personality Rights

Deepfakes enable the unauthorised replication of an individual's face, voice and likeness, posing serious risks to privacy, identity and personality rights. Unlike traditional digital manipulation, synthetic media can create convincing content that falsely links individuals to events or statements. The Rashmika Mandanna deepfake incident (2023)[1] highlighted how AI-generated identity misuse can spread rapidly across digital platforms without consent, resulting in significant reputational harm. As digital identity becomes increasingly important, its protection has emerged as a key concern in the AI era.

3.2 Financial Fraud and Cyber-Enabled Crimes

AI-generated deepfakes have significantly transformed financial fraud by enabling highly

[1] Centre Issues Advisory to Social Media Platforms over Deepfakes after Rashmika Mandanna Video' The Indian Express (7 November 2023) https://indianexpress.com/article/business/centre-deepfake-advisory-to-social-media-platforms-9017283/

convincing digital impersonation. Fake investment advertisements, fabricated endorsements and voice-cloning scams increasingly exploit public trust to deceive consumers. The Bengaluru investment scam involving deepfake videos of Virat Kohli, Anant Ambani and Union Finance Minister Nirmala Sitharaman[2] illustrates how synthetic media can be used to facilitate large-scale organised financial fraud. Such incidents highlight the growing economic risks associated with AI-enabled deception and the increasing vulnerability of digital financial systems.

3.3 Non-Consensual Intimate Content and Gender-Based Harm

The creation of AI-generated intimate content without consent has emerged as one of the most harmful forms of digital abuse. Victims often face harassment, emotional distress, social isolation and long-term reputational damage, even when the content is entirely fabricated. The rapid spread of such synthetic media further aggravates these harms, making timely and effective redress difficult. This misuse of artificial intelligence shows how deepfakes can make online abuse more harmful, especially for vulnerable individuals.

3.4 Political Misinformation and Electoral Integrity

Deepfakes have emerged as powerful tools for manipulating political discourse by fabricating speeches, endorsements and campaign material capable of influencing public perception. Incidents involving a manipulated video of Amit Shah and AI-generated political content featuring Aamir Khan and Ranveer Singh[3] during the 2024 General Elections illustrate the growing misuse of synthetic media in the electoral process. The increasing circulation of such content poses risks to informed public debate and weakens public confidence in democratic communication.

3.5 Digital Evidence and Public Confidence

The growing use of deepfakes has made it increasingly difficult to distinguish authentic content from manipulated media. This uncertainty affects not only digital communication but also the reliability of electronic evidence used by investigative agencies, courts, journalists and the public. As fabricated content becomes more convincing, even genuine digital evidence may be

[2] 'Deepfake Videos Show Virat Kohli, Nirmala Sitharaman Promoting Stock Trading; Bengaluru Police Register Case' The Times of India (12 November 2024) https://timesofindia.indiatimes.com/city/bengaluru/deepfake-videos-show-virat-kohli-nirmala-sitharaman-promoting-stock-trading/articleshow/125141596.cms

[3] Reuters, 'Deepfakes of Bollywood Stars Spark Worries of AI Meddling in India Election' (22 April 2024) https://www.reuters.com/world/india/deepfakes-bollywood-stars-spark-worries-ai-meddling-india-election-2024-04-22/

questioned, thereby weakening confidence in digital records and creating broader challenges for the administration of justice.

4. LEGAL FRAMEWORK GOVERNING AI-GENERATED DEEPFAKES IN INDIA

The rapid advancement of generative artificial intelligence has challenged the ability of conventional legal frameworks to regulate synthetic media. AI-generated deepfakes have expanded the scope of traditional cyber offences by enabling identity theft, financial fraud, reputational harm, privacy violations, election misinformation and large-scale digital deception. Although India has not enacted dedicated legislation governing deepfakes, the existing legal framework draws support from constitutional guarantees, cyber laws, criminal legislation, data protection law, evidentiary rules, intermediary regulations and election laws. Collectively, these measures provide remedies for different forms of misuse; however, their effectiveness increasingly depends on adapting established legal principles to the evolving nature of artificial intelligence.

4.1 Constitutional Framework

The constitutional foundation for regulating AI-generated deepfakes is primarily derived from Articles 14, 19 and 21 of the Constitution of India. Together, these provisions aim to balance technological innovation with the protection of fundamental rights in an increasingly digital society.

Article 14 guarantees equality before the law and equal protection of the laws and its relevance extends to situations where AI systems enable discriminatory targeting, algorithmic bias or selective digital manipulation that can lead to unequal legal or social consequences. As automated technologies increasingly influence public life, the constitutional principle of equality reinforces the need for fairness, transparency and accountability in the design and deployment of AI systems.

Article 19(1)(a) protects legitimate forms of expression, including the responsible use of artificial intelligence for purposes such as research, education, journalism, artistic creativity and technological innovation. However, this freedom is subject to reasonable restrictions under Article 19(2), particularly where synthetic media is used to spread misinformation, damage reputations, disturb public order or influence democratic processes. The increasing use of deepfakes has therefore made the constitutional balance between free expression and harmful digital manipulation more complex than before.

Article 21 provides the strongest constitutional safeguard by protecting privacy, dignity, reputation, identity and personal autonomy. These protections become particularly relevant where synthetic media is used for facial cloning, voice replication, fabricated interviews or non-consensual intimate content. While the Constitution provides the foundational safeguard for these rights, effective protection ultimately depends on a modern statutory framework capable of addressing AI-driven digital harms.

The constitutional framework therefore lays down the guiding principles for regulating synthetic media, while the effective enforcement of these rights ultimately depends on supporting statutory legislation.

4.2 Information Technology Act, 2000

The Information Technology Act, 2000[4] remains India's primary legislation governing cyber offences involving electronic records and computer resources. Although enacted prior to the emergence of generative artificial intelligence, several of its provisions continue to serve as the principal statutory framework for addressing offences committed through deepfake technology.

Sections 66C and 66D are particularly important where synthetic media is used for identity theft, voice cloning, digital impersonation or AI-enabled financial fraud. These provisions allow prosecution for fraudulent personation using computer resources and remain useful in addressing several forms of digital deception. However, in practice, establishing criminal intent and identifying the actual offender becomes much more difficult when AI-generated identities or cloned voices hide the true source of the offence, creating challenges that go beyond traditional cybercrime investigation.

Sections 67 and 67A provide legal remedies against obscene and sexually explicit synthetic content, particularly deepfake pornography and other forms of image-based abuse. While these provisions offer meaningful protection against online exploitation, their scope is largely limited to obscene or sexually explicit material and provides limited support where manipulated content is used for political misinformation, corporate fraud, election interference or reputational harm without any sexual element.

Sections 69A and 79 strengthen online regulation by empowering the government to block unlawful content and requiring intermediaries to exercise due diligenceHowever, these

[4] Information Technology Act 2000 (Act 21 of 2000), ss 66C, 66D, 67 and 67A.

provisions largely operate after harmful content has already spread. Given how quickly deepfakes can spread across digital platforms, a purely notice-and-takedown approach offers limited protection against fast-moving AI-driven misinformation and coordinated digital manipulation.

Rather than becoming outdated, the Information Technology Act continues to serve as the foundation of India's cyber regulatory framework. However, the growing advancement of generative AI highlights the need for complementary legal measures that specifically address synthetic media, platform accountability and emerging forms of digital impersonation.

4.3 Digital Personal Data Protection Act, 2023

The Digital Personal Data Protection Act, 2023[5], particularly its provisions relating to lawful processing of personal data (Section 4), consent (Section 6) and the obligations of Data Fiduciaries (Section 8), strengthens India's privacy framework by regulating the lawful processing of personal data through consent, purpose limitation and accountability. Its relevance extends to deepfakes where facial images, voice recordings, biometric data or other personal information are collected and processed without an individual's knowledge to generate synthetic media.

However, regulating personal data is not the same as regulating digital identity. The Act primarily focuses on the collection and processing of personal information but offers limited clarity once AI systems generate synthetic faces or cloned voices that can exist independently of the original data source. More importantly, its emphasis on personal data raises questions about whether AI-generated likenesses, which may not always directly contain identifiable personal information, fall within its protective scope. As synthetic identities become more realistic, relying on privacy law alone may not be sufficient to address emerging forms of AI-enabled impersonation.

4.4 Bharatiya Nyaya Sanhita, 2023

The Bharatiya Nyaya Sanhita, 2023[6] provides the primary criminal law framework for prosecuting offences committed through deepfake technology. Provisions relating to cheating (Section 318), forgery (Section 336), defamation (Section 356) and obscene acts and songs (Section 294) of the Bharatiya Nyaya Sanhita, 2023 enable authorities to prosecute various

[5] Digital Personal Data Protection Act 2023 (Act 22 of 2023), ss 4, 6 and 8.
[6] Bharatiya Nyaya Sanhita 2023 (Act 45 of 2023), ss 294, 318, 336 and 356.

offences involving AI-generated deepfakes, including financial fraud, fabricated digital records, reputational harm, and non-consensual intimate content.

Despite their continuing relevance, these provisions were originally designed for traditional forms of criminal conduct rather than AI-driven deception. Deepfakes can be created without revealing the source, spread instantly across multiple platforms and altered through automated systems, which makes it much harder to identify the offender and prove intent compared to conventional offences. This clearly shows the need for stronger investigative methods that can deal with the technical and evidentiary challenges created by generative artificial intelligence.

4.5 Bharatiya Sakshya Adhiniyam, 2023

The Bharatiya Sakshya Adhiniyam, 2023[7] assumes particular significance as artificial intelligence increasingly challenges the authenticity of electronic evidence. Its provisions governing electronic or digital records (Section 61), special provisions relating to electronic records (Section 62), admissibility of electronic records (Section 63) and expert opinion on electronic evidence (Section 39(2)) provide the legal foundation for the use of digital evidence before courts. However, these provisions were enacted primarily to regulate conventional electronic evidence and do not specifically address AI-generated synthetic media or prescribe specialised standards for verifying deepfakes.

As AI-generated content becomes more realistic, courts may increasingly encounter disputes concerning the authenticity of digital photographs, videos and audio recordings. The absence of statutory procedures for AI-specific forensic examination, verification of the source of digital content and authenticity checks may make it more difficult to distinguish genuine electronic evidence from manipulated synthetic media. This creates a significant evidentiary challenge, particularly in criminal investigations, election-related disputes and civil proceedings where electronic evidence often plays a decisive role.

4.6 Copyright Act, 1957

The Copyright Act, 1957[8], particularly its provisions relating to copyright ownership (Section 14) and copyright infringement (Section 51), provides limited protection where deepfakes reproduce or substantially alter copyrighted photographs, films, videos or other original creative works without permission. However, most deepfakes primarily misuse an individual's

[7] Bharatiya Sakshya Adhiniyam 2023 (Act 47 of 2023), ss 39(2), 61, 62 and 63.
[8] Copyright Act 1957 (Act 14 of 1957), ss 14 and 51.

face, voice, likeness or personality rather than the copyright in a protected work.

This distinction highlights an important limitation within the existing framework. While copyright law effectively protects creative expression, it offers limited remedies where AI replicates an individual's identity without infringing any copyrighted work. Consequently, many deepfake-related disputes fall outside traditional copyright protection and increasingly raise concerns relating to personality rights, privacy and individual autonomy.

4.7 Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 and Election Laws

The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021[9], particularly the provisions relating to due diligence by intermediaries (Rule 3(1)(b)), removal of unlawful content (Rule 3(1)(d)) and grievance redressal (Rule 3(2)), strengthen intermediary accountability by requiring digital platforms to exercise due diligence and remove unlawful content upon receiving valid notice. While this framework supports content moderation, it remains largely reactive and offers limited protection against the rapid spread of AI-generated deepfakes. In practice, manipulated content can reach a wide audience before any regulatory action is taken, reducing the effectiveness of post-publication takedown mechanisms.

The challenge becomes even more significant during elections, where synthetic media can influence political discourse, mislead voters and weaken public confidence in democratic institutions. Although the Representation of the People Act, 1951[10] addresses corrupt practices (Section 123) and election offences (Part VII), it does not specifically regulate AI-generated campaign material, synthetic political advertisements or mandatory disclosure of manipulated content. As digital campaigning increasingly uses generative AI, stronger transparency requirements, platform accountability and election-specific safeguards will be essential to protect electoral integrity.

4.8 Critical Legal Assessment

India's existing legal framework shows that deepfake-related harms can be addressed through constitutional protections, cyber laws, criminal legislation, data protection rules, evidentiary standards, intermediary obligations and election laws. Taken together, these measures provide

---

[9] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021, rr 3(1)(b), 3(1)(d) and 3(2).

[10] Representation of the People Act 1951 (Act 43 of 1951), s 123, pt VII.

a workable legal foundation; however, they operate separately and were developed independently to address different legal concerns rather than the combined and evolving challenges posed by generative artificial intelligence.

Consequently, a single deepfake incident may simultaneously involve privacy violations, identity theft, cyber fraud, reputational harm, copyright issues and electoral manipulation, requiring authorities to rely on multiple statutes with different objectives and enforcement mechanisms. This fragmentation can lead to procedural complexity, inconsistent interpretation and regulatory uncertainty. As synthetic media continues to evolve in use and scale, India requires a comprehensive, technology-responsive regulatory framework that complements existing laws while strengthening legal certainty, institutional accountability, digital trust and resilience in democratic processes.

5. JUDICIAL FOUNDATIONS FOR REGULATING AI-GENERATED DEEPFAKES IN INDIA

Although Indian courts have not yet delivered a landmark Supreme Court judgment exclusively governing AI-generated deepfakes, constitutional and personality rights jurisprudence has gradually developed principles that can address many of the legal concerns arising from synthetic media. Judicial recognition of privacy, dignity, freedom of expression, reputation and identity provides the constitutional foundation on which any future regulatory framework for deepfakes must be built. These decisions therefore serve not merely as precedents but as guiding principles for interpreting existing laws in the age of generative artificial intelligence.

5.1 Justice K.S. Puttaswamy (Retd.) v. Union of India (2017)

The decision in *Justice K.S. Puttaswamy (Retd.) v. Union of India*[11] marked a significant shift in Indian privacy jurisprudence by recognising privacy as a fundamental right under Articles 14, 19 and 21 of the Constitution. The Supreme Court affirmed that informational privacy, personal autonomy, dignity and individual control over personal data form essential components of constitutional liberty.

These principles become even more significant in the context of AI-generated deepfakes. Unlike conventional privacy violations, synthetic media enables the unauthorised replication of an individual's face, voice, expressions and digital identity without any physical interaction or consent. Such misuse goes beyond the disclosure of personal information, it creates a

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

fabricated digital persona that can mislead society and cause lasting harm to dignity and reputation. In this context, *Puttaswamy* provides the strongest constitutional basis for treating deepfake misuse as a serious violation of informational privacy and digital identity.

5.2 Shreya Singhal v. Union of India (2015)

In *Shreya Singhal v. Union of India*[12], the Supreme Court reaffirmed that restrictions on online speech must comply with the safeguards under Article 19(2). By striking down Section 66A of the Information Technology Act, the Court emphasised that vague or disproportionate restrictions cannot be used to limit legitimate expression.

This principle is equally important for regulating deepfakes. An effective legal framework must clearly distinguish between malicious synthetic media intended to deceive, defraud or manipulate and legitimate uses of artificial intelligence in journalism, research, artistic creation, satire and public discourse. The core challenge is therefore not whether deepfakes should be regulated, but how to regulate them without disproportionately restricting freedom of expression. *Shreya Singhal* thus provides the constitutional standard for maintaining this balance.

5.3 R. Rajagopal v. State of Tamil Nadu (1994)

The judgment in R. Rajagopal v. State of Tamil Nadu[13] recognised that individuals have a protected sphere of privacy that cannot ordinarily be interfered with without consent. Although the case dealt with the unauthorised publication of personal information, the underlying principle extends beyond disclosure to the protection of personal identity itself. Deepfakes magnify this concern by replacing reality with fabrication. Instead of revealing private information, they manufacture false speeches, actions, images and events capable of fundamentally altering public perception of an individual. The constitutional injury therefore extends beyond privacy to include dignity, reputation and identity. Viewed in this light, Rajagopal provides an important jurisprudential foundation for recognising that digital identity deserves protection against technological manipulation as much as against unauthorised disclosure.

5.4 Anil Kapoor v. Simply Life India & Others (Delhi High Court, 2023)

[12] Shreya Singhal v Union of India (2015) 5 SCC 1.
[13] R Rajagopal v State of Tamil Nadu (1994) 6 SCC 632.

 ISSN: 2581-8503

The Delhi High Court's decision in *Anil Kapoor v. Simply Life India & Others*[14] represents one of India's most significant judicial responses to AI-enabled identity misuse. The Court restrained the unauthorised commercial exploitation of the actor's name, voice, image, likeness and other distinctive personality attributes, including AI-generated content that could mislead the public.

More importantly, the judgment recognises that personality rights continue to remain enforceable even as technology evolves. By extending protection against AI-enabled voice cloning, facial replication and digital impersonation, the Court acknowledged that personal identity itself has become increasingly vulnerable in the age of generative artificial intelligence. However, judicial protection through individual cases alone cannot offer a complete regulatory solution. As deepfake technology becomes more accessible, a comprehensive legislative framework will be necessary to ensure consistent protection beyond case-specific remedies.

5.5 Judicial Analysis

Taken together, these decisions show that Indian constitutional jurisprudence already protects the core values threatened by deepfakes, privacy, dignity, identity, reputation and freedom of expression. The key challenge is not the absence of constitutional principles, but the absence of a dedicated legal framework that can effectively apply those principles to AI-generated synthetic media. Future legislation, therefore, should build on these judicial foundations rather than replace them.

6. COMPARATIVE AND INTERNATIONAL LEGAL APPROACHES TO REGULATING AI-GENERATED DEEPFAKES

The rapid evolution of artificial intelligence has prompted both international organisations and national governments to develop legal frameworks to address the challenges posed by AI-generated synthetic media. While international instruments set broad standards for ethical AI governance, domestic jurisdictions have adopted specific regulatory models shaped by their constitutional values and policy priorities. Examining these developments offers useful guidance for building a balanced legal framework to address deepfake-related harms in India.

6.1 International Perspective

Although no universal treaty specifically regulates AI-generated deepfakes, several

[14] Anil Kapoor v Simply Life India and Others 2023 SCC OnLine Del 6914.

international instruments have laid down foundational principles for governing artificial intelligence. These frameworks consistently emphasise that technological development must remain aligned with human rights, democratic values and the rule of law.

The UNESCO Recommendation on the Ethics of Artificial Intelligence, 2021[15] is the first global normative framework dedicated to ethical AI governance. It encourages Member States to promote human dignity, privacy, equality, cultural diversity and human oversight, while ensuring that AI systems are developed and used in a transparent and responsible manner.

Similarly, the OECD Principles on Artificial Intelligence (2019)[16] promote responsible AI through fairness, reliability, explainability, risk management and responsible governance. Complementing these initiatives, the Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law (2024)[17] establishes the first legally binding international framework requiring AI systems to operate consistently with human rights, democratic principles and effective legal oversight.

Collectively, these instruments establish internationally recognised benchmarks that continue to shape national AI governance and the evolving regulation of synthetic media.

6.2 European Union - Artificial Intelligence Act, 2024

The European Union has adopted one of the world's first comprehensive legal frameworks on artificial intelligence through the Artificial Intelligence Act, 2024[18]. Unlike conventional cyber legislation, the Act follows a preventive and risk-based regulatory model by classifying AI systems based on their potential impact on fundamental rights, public safety and democratic values.

A defining feature of the Act is Article 5, which prohibits AI practices that pose unacceptable risks to individuals and society. This preventive approach aims to reduce harm before high-risk AI systems are deployed in public use.

Another significant safeguard appears in Article 50[19], which requires AI-generated images, audio, video and other synthetic content to be clearly disclosed. Mandatory labelling helps users distinguish authentic content from manipulated media, thereby strengthening digital authenticity and reducing deceptive uses of artificial intelligence.

[15] https://unesdoc.unesco.org/ark:/48223/pf0000381137
[16] https://www.oecd.org/en/topics/sub-issues/ai-principles.html
[17] https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence
[18] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L 1689 https://eur-lex.europa.eu/eli/reg/2024/1689/oj
[19] Regulation (EU) 2024/1689, arts 5 and 50 https://eur-lex.europa.eu/eli/reg/2024/1689/oj

The Act also imposes extensive compliance obligations relating to risk assessment, technical documentation, cybersecurity, human oversight, post-market monitoring, and incident reporting. Supervision by the European AI Office, the European Commission, and national competent authorities further strengthens regulatory compliance through a structured governance framework.

6.3 China - Deep Synthesis Internet Information Service Provisions, 2023

China became the first country to introduce dedicated regulations governing AI-generated deepfakes through the Deep Synthesis Internet Information Service Provisions, 2023[20]. The Regulations, particularly the provisions relating to mandatory labelling of synthetic content (Article 9), identity verification (Article 14) and consent for the use of another person's face or voice (Article 16), establish a specialised legal framework for regulating AI-generated synthetic media

Mandatory disclosure of AI-generated content enhances transparency by enabling users to distinguish authentic material from manipulated media, thereby reducing deception and strengthening public confidence in digital information.

Similarly, identity verification and consent requirements enhance accountability by discouraging the unauthorised use of another person's face or voice and facilitating regulatory oversight. Together, these measures provide a comprehensive regulatory model that addresses the creation, distribution and misuse of deepfakes.

6.4 United States - State-Level Deepfake Regulation

The United States has developed a sector-specific regulatory approach to AI-generated deepfakes through targeted legislation, policy initiatives and technical governance standards. This multi-layered framework focuses on protecting elections, privacy, digital identity and victims of AI-enabled abuse, while also supporting continued technological innovation.

California Assembly Bill 730[21] prohibits materially deceptive deepfake videos intended to influence elections, while Assembly Bill 602[22] provides civil remedies to victims of AI-generated non-consensual intimate content. Similarly, Texas Senate Bill 751[23] criminalises

[20] Provisions on the Administration of Deep Synthesis of Internet Information Services (promulgated 25 November 2022, effective 10 January 2023), arts 9, 14 and 16 https://www.chinalawtranslate.com/en/deep-synthesis/
[21] California AB 730 https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB730
[22] California AB 602 https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201920200AB602
[23] Texas SB 751 https://capitol.texas.gov/BillLookup/History.aspx?LegSess=86R&Bill=SB751

deceptive synthetic media used to interfere with electoral processes and Virginia Code Section 18.2-386.2[24] extends criminal liability to AI-generated intimate images.

The regulatory framework is further supported by the Blueprint for an AI Bill of Rights (2022)[25] and the NIST AI Risk Management Framework[26], which promote safe, trustworthy and rights-based AI development through voluntary governance standards. Together, these legislative measures and policy initiatives reflect an adaptive regulatory model capable of responding to the rapidly evolving nature of artificial intelligence technologies.

6.5 Comparative Analysis

The comparative study shows that jurisdictions have adopted different regulatory approaches to address AI-generated deepfakes while sharing a common objective of protecting individuals and democratic institutions. International instruments set out broad governance principles, whereas domestic legal systems translate these principles into enforceable rules suited to their respective legal frameworks.

The European Union prioritises preventive regulation through risk-based AI governance, China focuses on technology-specific oversight supported by labelling, traceability and consent requirements, while the United States adopts targeted legislative protections complemented by policy initiatives and technical standards. Although their regulatory strategies differ, each framework highlights the importance of combining preventive safeguards with effective legal mechanisms capable of addressing emerging AI-enabled harms.

These comparative developments provide useful guidance for India in designing a balanced regulatory framework that protects privacy, digital identity, electoral integrity and public confidence while also supporting the responsible growth of artificial intelligence.

7. Critical Analysis

The foregoing analysis indicates that India's existing regulatory response remains largely consequence-driven rather than technology-oriented. The current legal regime does address offences arising from deepfake misuse, such as identity theft, cyber fraud, defamation, privacy violations and digital impersonation. However, it does not establish a dedicated governance framework for regulating the creation, dissemination and accountability of AI-generated

[24] Virginia Code § 18.2-386.2 https://law.lis.virginia.gov/vacode/title18.2/chapter8/section18.2-386.2/
[25] White House Office of Science and Technology Policy, Blueprint for an AI Bill of Rights (October 2022) https://bidenwhitehouse.archives.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf
[26] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1, January 2023) https://www.nist.gov/itl/ai-risk-management-framework

synthetic media. As a result, enforcement often relies on applying statutory provisions that were enacted well before the emergence of generative artificial intelligence.

A further concern is the absence of a coordinated regulatory framework governing artificial intelligence. Existing statutes operate independently, leading to fragmented compliance requirements, overlapping enforcement responsibilities and inconsistent legal remedies. The lack of uniform standards for AI-generated content authentication, platform obligations, digital identity protection and evidentiary verification creates uncertainty for regulators, intermediaries, courts and victims alike. This fragmentation reduces regulatory efficiency and limits the ability of existing laws to respond effectively to rapidly evolving artificial intelligence technologies.

The comparative study shows that leading jurisdictions have gradually moved towards preventive AI governance through technology-specific obligations, regulatory oversight, transparency standards and institutional accountability. Indian constitutional jurisprudence has also strengthened the protection of privacy, dignity and freedom of expression, thereby providing a strong normative foundation for future AI regulation. However, constitutional principles and judicial interpretation alone cannot replace comprehensive legislative intervention in a field marked by continuous technological change.

Viewed collectively, these developments show that the key challenge before India is no longer merely recognising the harms caused by deepfakes, but developing a comprehensive regulatory framework capable of preventing such harms before they occur. A modern governance approach must therefore move beyond fragmented statutory responses and adopt technology-specific regulation that integrates legal certainty, institutional oversight, platform accountability and effective enforcement, while still preserving innovation and constitutional freedoms.

8. RECOMMENDATIONS FOR STRENGTHENING INDIA'S LEGAL FRAMEWORK ON AI-GENERATED DEEPFAKES

8.1 Enact a Dedicated Deepfake Regulation Act

The proposed legal framework should introduce a dedicated Deepfake Regulation Act rather than relying on fragmented provisions spread across multiple statutes. The legislation should clearly define deepfakes, classify high-risk applications, prescribe graded civil and criminal liability, provide victim compensation, regulate political and commercial use and establish effective enforcement mechanisms. A comprehensive statute of this nature would enhance legal certainty while ensuring a uniform regulatory approach to AI-generated synthetic media.

8.2 Introduce Mandatory AI Content Authentication and Watermarking Standards

The regulatory framework should mandate secure digital watermarking and content authentication for AI-generated images, audio and videos based on internationally recognised standards such as the Coalition for Content Provenance and Authenticity (C2PA). These obligations should apply to both AI developers and digital platforms to strengthen content authenticity, enable source verification and reduce the misuse of synthetic media.

8.3 Strengthen Platform Accountability and Safe Harbour Compliance

The existing intermediary framework should impose stronger obligations on digital platforms to detect, report and remove malicious deepfakes. Verified unlawful synthetic content should be taken down within a prescribed 24-48-hour timeframe, supported by specialised grievance officers, AI-based detection systems, transparency reporting and proportionate penalties for non-compliance. Such measures would strengthen intermediary accountability without restricting legitimate online expression.

8.4 Statutorily Recognise Digital Identity and Personality Rights

The legal framework should explicitly recognise an individual's face, voice, likeness and other biometric attributes as protected digital personality rights. Building on Article 21 and the principles laid down in *Anil Kapoor v. Simply Life India*, unauthorised AI-enabled replication or commercial exploitation of these attributes should attract effective civil remedies and criminal sanctions, thereby ensuring stronger protection of personal autonomy and digital identity.

8.5 Strengthen Electoral Safeguards Against AI-Generated Misinformation

Election laws should explicitly regulate AI-generated political content to safeguard electoral integrity. Mandatory disclosure of AI-generated campaign material, restrictions on deceptive synthetic media during election periods, quick content removal mechanisms and enhanced regulatory powers for the Election Commission of India would strengthen electoral transparency while preserving democratic fairness.

8.6 Develop AI Forensic and Evidentiary Standards

The regulatory framework should establish specialised AI forensic standards for detecting, preserving and authenticating AI-generated evidence. Standardised forensic protocols, certified expert examination, digital chain-of-custody procedures and dedicated forensic laboratories would enhance the reliability of electronic evidence and strengthen the investigation and adjudication of deepfake-related offences.

8.7 Establish an Independent Artificial Intelligence Regulatory Authority

A specialised statutory authority should be constituted to regulate high-risk AI applications, formulate technical standards, conduct compliance audits, investigate violations, coordinate with sectoral regulators and advise the government on emerging AI risks. Such a dedicated regulatory institution would ensure consistent implementation while also supporting responsible AI governance and technological innovation.

8.8 Strengthen International Cooperation and Periodic Regulatory Review

The regulatory framework should strengthen international cooperation through mutual legal assistance, cross-border information sharing, technical collaboration and participation in global AI governance initiatives. It should also incorporate a periodic legislative review mechanism to ensure that India's legal framework remains responsive to technological advancements and emerging forms of AI-generated synthetic media.

9. CONCLUSION

The rapid advancement of artificial intelligence has fundamentally reshaped the creation and dissemination of digital content, making AI-generated deepfakes one of the most complex legal challenges of the digital age. While synthetic media presents significant opportunities for innovation, its misuse has raised serious concerns relating to privacy, identity, electoral integrity, cybercrime and the authenticity of digital information. These developments show that regulating deepfakes is no longer a future policy consideration but an immediate legal necessity.

The study shows that India's constitutional principles, cyber laws, criminal legislation, data protection framework and judicial precedents together provide an important legal foundation for addressing deepfake-related offences. However, these protections remain spread across multiple statutes and were not originally designed to govern AI-generated synthetic media as a distinct technological phenomenon. In contrast, emerging international approaches increasingly emphasise preventive regulation, clearly defined institutional roles and specialised safeguards to address evolving AI-related risks. These developments highlight the need for a dedicated statutory framework that ensures legal certainty, improves institutional coordination and strengthens accountability.

Ultimately, the challenge before India is not whether artificial intelligence should be regulated, but how it can be governed without undermining innovation or constitutional freedoms. A dedicated Deepfake Regulation Act, supported by clear regulatory standards, specialised institutional oversight and effective enforcement mechanisms, would enable India to respond effectively to the evolving realities of synthetic media. As digital content becomes increasingly difficult to distinguish from reality, the strength of the legal system will be measured not by its ability to react to technological misuse, but by its capacity to respond to it and protect the values on which a democratic digital society depends.