INTERNATIONAL JOURNAL FOR LEGAL RESEARCH AND ANALYSIS

Open Access, Refereed Journal Multi Disciplinary
Peer Reviewed

www.ijlra.com

ISSN: 2582-6433

DISCLAIMER

The views, opinions, interpretations, and conclusions expressed in the articles published in this journal are solely those of the respective authors. They do not necessarily reflect the views of the Editorial Board, Editors, Reviewers, Advisors, or the Publisher of IJLRA.

Although every reasonable effort has been made to ensure the accuracy, authenticity, and proper citation of the content published in this journal, neither the Editorial Board nor IJLRA shall be held liable or responsible, in any manner whatsoever, for any loss, damage, or consequence arising from the use, reliance upon, or interpretation of the information contained in this publication.

The content published herein is intended solely for academic and informational purposes and shall not be construed as legal advice or professional opinion.

ISSN: 2582-6433

ABOUT US

The *International Journal for Legal Research & Analysis (IJLRA)* (ISSN: 2582-6433) is a peer-reviewed, academic, online journal published on a monthly basis. The journal aims to provide a comprehensive and interactive platform for the publication of original and high-quality legal research.

IJLRA publishes Short Articles, Long Articles, Research Papers, Case Comments, Book Reviews, Essays, and interdisciplinary studies in the field of law and allied disciplines. The journal seeks to promote critical analysis and informed discourse on contemporary legal, social, and policy issues.

The primary objective of IJLRA is to enhance academic engagement and scholarly dialogue among law students, researchers, academicians, legal professionals, and members of the Bar and Bench. The journal endeavours to establish itself as a credible and widely cited academic publication through the publication of original, well-researched, and analytically sound contributions.

IJLRA welcomes submissions from all branches of law, provided the work is original, unpublished, and submitted in accordance with the prescribed submission guidelines. All manuscripts are subject to a rigorous peer-review process to ensure academic quality, originality, and relevance.

Through its publications, the *International Journal for Legal Research & Analysis* aspires to contribute meaningfully to legal scholarship and the development of law as an instrument of justice and social progress.

ISSN: 2582-6433

PUBLICATION ETHICS, COPYRIGHT & AUTHOR RESPONSIBILITY STATEMENT

The *International Journal for Legal Research and Analysis (IJLRA)* is committed to upholding the highest standards of publication ethics and academic integrity. All manuscripts submitted to the journal must be original, unpublished, and free from plagiarism, data fabrication, falsification, or any form of unethical research or publication practice. Authors are solely responsible for the accuracy, originality, legality, and ethical compliance of their work and must ensure that all sources are properly cited and that necessary permissions for any third-party copyrighted material have been duly obtained prior to submission. Copyright in all published articles vests with IJLRA, unless otherwise expressly stated, and authors grant the journal the irrevocable right to publish, reproduce, distribute, and archive their work in print and electronic formats. The views and opinions expressed in the articles are those of the authors alone and do not reflect the views of the Editors, Editorial Board, Reviewers, or Publisher. IJLRA shall not be liable for any loss, damage, claim, or legal consequence arising from the use, reliance upon, or interpretation of the content published. By submitting a manuscript, the author(s) agree to fully indemnify and hold harmless the journal, its Editor-in-Chief, Editors, Editorial Board, Reviewers, Advisors, Publisher, and Management against any claims, liabilities, or legal proceedings arising out of plagiarism, copyright infringement, defamation, breach of confidentiality, or violation of third-party rights. The journal reserves the absolute right to reject, withdraw, retract, or remove any manuscript or published article in case of ethical or legal violations, without incurring any liability.

ISSN: 2582-6433

THE DEEPFAKE DILEMMA: PRIVACY, REPUTATION, AND REGULATION IN THE AGE OF ARTIFICIAL INTELLIGENCE

AUTHORED BY - DR. GUNJAN SHARMA[1]
& MS. NEHA SHREE BHATNAGAR[2]

Abstract

The problem of "deep fakes," created using artificial intelligence, has emerged as a significant menace to privacy, dignity, reputation and confidence in democracy in India. The Sumsub State of Deepfakes 2023 report suggests there has been a surge in deepfake fraud globally and that India has been among the nations seriously impacted using AI to generate identities for fraudulent purposes. In recent years, incidents involving the creation of deepfake videos of public figures and actors, such as the widely publicized deepfake of actor Rashmika Mandanna in 2023, have been used for sexual harassment, impersonation, misinformation, and reputational damage.

This article explores the issue of misuse of deepfake in India from the perspective of privacy and legal regulation. It examines the constitutional framework of privacy that was established in K.S. Puttaswamy v. Union of India in which the Supreme Court declared privacy as a fundamental right guaranteed by Article 21 of the Constitution. It also considers legal provisions regarding offences under the Bharatiya Nyaya Sanhita, 2023, the Information Technology Act, 2000 and the Digital Personal Data Protection Act, 2023 relating to identity theft, cheating by personation, obscenity, defamation, cyberstalking, non-consensual circulation of intimate content etc.

The article states that current Indian laws offer piecemeal solutions, and are inadequate for the rapid, anonymous and viral misuse of AI. It also considers judicial principles in the context of cases like Shreya Singhal v. Union of India, Subramanian Swamy v. Union of India and K.S. Puttaswamy, which involve a balance between free speech, reputation, privacy and State regulation.

[1] Assistant Professor, School of Law, Bennett University, Greater Noida.
[2] Assistant Professor, School of Law, Bennett University, Greater Noida.

ISSN: 2582-6433

The article concludes with the need for a clearer legal definition of a deepfake in India, consent to the use of biometric data, labels on synthetic media, quick takedown, platform responsibility, compensation for deepfake victims and digital literacy. Any regulation should ensure privacy and dignity without stifling proper innovation and expression.

Key words- Deepfakes, Privacy, Artificial Intelligence and Cyber Law.

1. Introduction

The power of artificial intelligence (AI) has revolutionized digital content creation, modification, and dissemination. The most notable development in this area is deepfakes. Deepfakes are synthetic audio, video, or image content produced or altered using AI to generate the most realistic possible deepfakes of a person, from the combination of deep learning and fake. Deepfakes may be used to modify facial expressions, clone voices, and make up events that never took place, depending on the advanced machine-learning techniques that are employed, such as Generative Adversarial Networks (GANs) and other generative AI models.[3] The use of deepfake technology in entertainment, education, advertising, accessibility and digital innovation is legitimate, but it is increasingly a concern worldwide when it is misused. With the rise in the accessibility of AI tools, anyone can easily produce believable fake content. Consequently, deepfakes are being employed in non-consensual intimate imagery, identity theft, financial fraud, cyber harassment, political propaganda and misinformation campaigns, and reputational attacks. Deepfakes challenge the credibility of digital media and Internet communication. In India, the issue is especially critical as the country undergoes digitalisation at a fast pace, and social media is used by a significant portion of the population, making manipulated content more accessible and influential. In 2023, ***the Rashmika Mandanna***, deepfake scandal brought attention to the risks of AI-generated media, particularly of women, to many people. Concerns also have been raised about AI-driven political misinformation, content designed to influence elections, and voice-cloning scams to impersonate family members, government officials, executives of businesses and others to facilitate financial gain.[4] These cases illustrate that Deepfakes are not just a technological novelty anymore, but very real threats of privacy, personal security and democracy. The proliferation of deepfakes poses

[3] Bobby Chesney and Danielle Citron, "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security" 107 *California Law Review* 1753 (2019).
[4] Press Information Bureau, Government of India, "Government Issues Advisory to Social Media Platforms to Comply with IT Rules Following Deepfake Concerns" (2023), available at: https://pib.gov.in (visited on June 23, 2026).

ISSN: 2582-6433

a significant legal dilemma because traditional laws are not equipped to handle AI-generated synthetic content. In India, the right to privacy, dignity, reputation and information autonomy are constitutionally protected and are infringed upon by deepfakes. On the other hand, any regulatory measure should be taken in a way that does not exceed the constitutional bounds of freedom of speech and expression. What is unclear is who is liable, how it will be enforced, protection for victims, and platform responsibility, due to the lack of a specific legal framework regarding deepfakes. While provisions in the Information Technology Act, 2000, the Digital Personal Data Protection Act, 2023 and the Bharatiya Nyaya Sanhita, 2023 can be applied when a deepfake is being used for illegal purposes, these laws offer incomplete remedies and lack a holistic approach to the specific issues arising from synthetic media. AI's speed, anonymity, scalability, and borderless production make the legal responses to the challenge of traditional content production ever less effective. In this context, the problem of deepfakes has become one of the biggest challenges in the fields of technology, privacy and law.

2. Understanding Deepfakes and Their Societal Impact

2.1 Meaning and Evolution of Deepfakes

The rapid advancement of Artificial Intelligence (AI) has revolutionized the way information is created, shared, and consumed. Among the most significant developments in this field is the emergence of **deepfake technology**, which has transformed digital content creation while simultaneously creating unprecedented legal and ethical challenges. The term deepfake is a combination of the words deep learning and fake, referring to synthetic media generated or manipulated using artificial intelligence to create realistic but fabricated audio, video, images, or text.[5]

The most common methods for creating deepfakes are using Generative Adversarial Networks (GANs), deep neural networks, and voice-cloning software. These systems look at large quantities of images, video and audio recordings and learn an individual's facial expressions, speech patterns, gestures and mannerisms. The trained AI will be able to produce content that is very realistic and looks authentic to the average viewer. This means that deepfakes can portray people doing things, saying things or being in places they never were.

Deepfake technology has its roots in early 2010s advancements in computer vision and AI. At first, these technologies were born for good reason, such as facial recognition systems, film production, image enhancement and virtual reality. As time passed, the open-source software

[5] Bobby Chesney and Danielle Keats Citron, "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security" 107 *California Law Review* 1753 (2019).

ISSN: 2582-6433

and easy-to-use AI platforms reduced the technical hurdles of generating synthetic media even more. Thanks to readily accessible apps, it is now possible to create convincing deepfake content in just minutes with little technical expertise.

This has been amplified by the advent of Generative Artificial Intelligence (Generative AI). Tools like chatbots, image generators like ChatGPT, Dall-E and Stable Diffusion, Midjourney, and other powerful voice-synthesis solutions have illustrated how technology is increasingly capable of creating human-like material. These innovations have improved creativity and productivity but have also sparked questions about the authenticity of creation, misinformation and digital manipulation. As a result, the potential use of deepfakes has become one of the most debated applications in current society.

2.2 Legitimate Uses and Misuse

"Deepfakes" do not always have to be bad. As with most new technologies, its effect is determined by its use. Synthetic media has many valuable uses in the proper context, across many fields.

In the film industry, deepfake technology is utilized for visual enhancements, dubbing in various languages, removing aging effects from actors, and creating historical figures. AI-generated content is being used more in film studio productions to cut down on production costs and enhance viewers' experiences. Synthetic media is used at schools to develop interactive learning materials and virtual simulations that facilitate students' learning on complex subjects.[6]

Deepfake technologies have also been used to enhance the healthcare sector. Voice-synthesis systems are used to assist those who have speech difficulties to communicate better. The avatars created by AI are employed in mental health counselling, customer service, and access support for individuals with disabilities. Virtual assistants and AI-powered voice actors are used by businesses to enhance communication and customer engagement.[7]

The potential application of deep fake technology has benefits, but also significant concerns. The use of deepfakes for financial scams, misinformation, or to create intimate images and videos without consent, impersonation of people, and manipulating public opinion is becoming more common. AI voices used to trick recipients into sending funds or revealing personal

[6] Daniel Leiker et al., "Generative AI for Learning: Investigating the Potential of Synthetic Learning Videos" *arXiv* (Apr. 7, 2023), available at: http://arxiv.org/abs/2304.03784 (visited on June 23, 2026).

[7] Erik Hermann, Gizem Yalcin Williams and Stefano Puntoni, "Deploying Artificial Intelligence in Services to Aid Vulnerable Consumers" 52(6) *Journal of the Academy of Marketing Science* 1431 (2024), available at: https://doi.org/10.1007/s11747-023-00986-8 (visited on June 23, 2026).

ISSN: 2582-6433

details. Criminals use AI voices to make the recipient feel like they are being deceived and transfer money or give away personal information.[8] This kind of use shows how the use of a technology created for the purpose of innovation can be turned into a means of exploitation and deception.

2.3 Examples, Incidents and Statistical trends

With the rising popularity of deepfakes, the issue has become serious. The Sumsub State of Deepfakes 2023 Report, which covered new trends in AI-generated identity fraud and synthetic document manipulation, found India among the top countries seeing a surge in such frauds. [9]

The Rashmika Mandanna deepfake controversy in 2023 is one of the most talked-about incidents in India. A fake video featuring the actress was shared on social media, which was created using artificial intelligence. A fake video that used artificial intelligence to create the actress was spread across social media platforms. The case highlighted the potential misuse of AI-generated media in violation of privacy, damaging reputations, and harassment, especially of women. [10]

Voice cloning fraud is also a concern that is increasing. AI has been used in some countries, such as India, to impersonate family members, company executives, and government officials, and the fraudsters have used these voice recordings to gain access to victims. The impersonated voice is deemed to be a legitimate call, and the victims are fooled into wiring funds or divulging sensitive details. As voice-cloning technology becomes more widely available and cheap, cybersecurity experts have predicted these scams will likely rise, too.

Another application of deep fake technology is the spreading of political misinformation. In recent times, false information about political leaders and candidates has been spread via manipulated videos and audio recordings in election campaigns around the world. The report by the World Economic Forum (WEF) Global Risks Report 2024 highlighted one of the biggest threats of the short term – AI generated misinformation and disinformation – as it will harm democratic institutions by undermining public trust in information systems. [11]

The research also shows that there is a strong gender aspect on deepfake misuses. In a study by

[8] Marc Schmitt and Ivan Fléchais, "Digital Deception: Generative Artificial Intelligence in Social Engineering and Phishing" 57 *Artificial Intelligence Review* (2024), available at: https://doi.org/10.1007/s10462-024-10973-2 (visited on June 23, 2026).
[9] Sumsub, *The State of Deepfakes 2023: The Battle for Truth* (2023), available at: https://sumsub.com (visited on June 23, 2026).
[10] Press Information Bureau, Government of India, "Government Issues Advisory to Social Media Platforms on Deepfakes" (2023), available at: https://pib.gov.in (visited on June 23, 2026).
[11] World Economic Forum, *Global Risks Report 2024* (2024), available at: https://www.weforum.org (visited on June 23, 2026).

ISSN: 2582-6433

Deeptrace Labs, roughly 96% of online deep fake videos were pornographic in nature with the vast majority being of sexually explicit origin targeting women without their consent.[5] This type of content often inflicts serious emotional, psychological, reputation and harm upon the victim and exposes the intersection between deep fake technology and gender-based digital violence.[12]

2.4 Societal, Democratic, and Legal Concerns

The impact of deepfakes on society goes well beyond the individual. The issue of a threat to privacy is one of the greatest concerns. Deepfakes are frequently based on the unauthorized access and alteration of personal information, such as photos, videos, and voice recordings. These practices have a detrimental impact on people's capacity to manage their own personal data and digital identity.

Besides posing a threat to human dignity and reputation, deepfakes also have negative implications for human rights. The use of fabricated content showing people in situations that are compromising or offensive can irreparably harm personal or professional relationships. Deepfakes can be easily shared on digital platforms, and correcting and removing them is extremely difficult, unlike the traditional types of defamation.

One of their biggest worries is how they affect the democratic conversation. Democracies depend on citizens' access to information and their trust in information. The potential for deepfakes to mislead public discourse, influence public opinion, and confuse about actual events is a cause for concern. Even a brief video deepfake can sway the public opinion during an election before it is proven to be fake.

Deepfakes also cause problems for the administration of justice. Photographs, videos, and audio recordings have been seen as valuable evidence traditionally. But with the sophisticated nature of synthetic media, courts can find it hard to establish authenticity of digital evidence. This development may lead to a lack of trust in the judicial system and make it more difficult to bring criminal charges.

The societal impact of deepfakes therefore extends beyond privacy violations and reputational harm. It affects democratic governance, cybersecurity, law enforcement, and public confidence in digital communication. These challenges underscore the urgent need for an effective legal and regulatory framework capable of addressing the risks posed by AI-generated synthetic media.

[12] Deeptrace Labs, *The State of Deepfakes: Landscape, Threats and Impact* (2019), available at: https://deeptracelabs.com (visited on June 23, 2026).

ISSN: 2582-6433

3. Constitutional and Legal Dimensions of Deepfakes in India

The emergence of deepfake technology has created significant constitutional and legal challenges in India. Deepfakes affect fundamental rights such as privacy, dignity, reputation, and freedom of speech while simultaneously exposing gaps in the existing legal framework. Since India does not yet have a dedicated law regulating AI-generated synthetic media, deepfake-related harms are addressed through a combination of constitutional principles, cyber laws, criminal laws, data protection legislation, and intermediary liability rules. However, the rapid advancement of artificial intelligence has exposed the limitations of these traditional legal mechanisms.

3.1 Right to Privacy and Informational Autonomy

The constitutional foundation for addressing deepfake-related harms lies in the landmark judgment of K.S. Puttaswamy v. Union of India, in which a nine-judge bench of the Supreme Court recognized the right to privacy as a fundamental right under Article 21 of the Constitution.[13] The Court held that privacy encompasses personal autonomy, dignity, bodily integrity, and informational self-determination.

One of the most important aspects of the judgment is the recognition of **informational privacy**, which grants individuals control over the collection, use, and dissemination of their personal information. Deepfakes directly threaten this right because they frequently rely on the unauthorized use of photographs, videos, voice recordings, and biometric data. AI systems can scrape publicly available content from social media platforms and use it to generate realistic but fabricated content without the individual's consent.

The unauthorized use of facial recognition data, voice samples, and digital identities undermines an individual's ability to control their online presence. Voice-cloning technologies, for instance, can replicate a person's speech patterns with remarkable accuracy using only a few seconds of recorded audio. Such cloned voices may be employed for fraud, impersonation, misinformation, or reputational attacks. Consequently, deepfakes represent a serious intrusion into personal privacy and autonomy in the digital age.

3.2 Right to Dignity, Reputation, and Bodily Autonomy

Deepfakes also pose a significant threat to the constitutional right to dignity. The Supreme Court has repeatedly interpreted Article 21 to include the right to live with dignity and self-

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

ISSN: 2582-6433

respect. Deepfake technology is frequently used to create fabricated content that humiliates, degrades, or misrepresents individuals, particularly women.

The most alarming examples involve non-consensual intimate deepfakes, where a person's face is digitally inserted into sexually explicit content. Such content often causes severe emotional distress, social stigma, professional damage, and psychological trauma. Even though the depicted acts never occurred, the consequences for victims are often devastating.

The right to dignity is closely connected to the protection of reputation. In Subramanian Swamy v. Union of India,[14] the Supreme Court recognized reputation as an integral component of Article 21.[3] Deepfakes that falsely portray individuals engaging in immoral, criminal, or offensive conduct can therefore violate both constitutional rights and legal protections against defamation.

3.3 Freedom of Speech and Expression

While deepfakes can cause serious harm, any attempt to regulate them must be balanced against the constitutional guarantee of freedom of speech and expression under Article 19(1)(a). Artificial intelligence and synthetic media can be used for legitimate purposes such as satire, parody, artistic expression, education, journalism, and political commentary.

In Shreya Singhal v. Union of India, the Supreme Court emphasized the importance of protecting online speech and struck down Section 66A of the Information Technology Act, 2000 as unconstitutional. [4] The Court held that restrictions on online expression must be reasonable, narrowly tailored, and consistent with Article 19(2).[15]

This is especially important when it comes to deepfakes. Not all deepfakes are negative or illegal. Some may be not only satire but also creative expression. Thus, it is important for regulators to be able to separate out the malicious from the legitimate use of synthetic media. Too general an approach might lead to censorship, suppression of innovation, or restrictions on constitutionally protected speech.

3.4 Information Technology Act, 2000

The Information Technology Act, 2000 was enacted prior to the advent of artificial intelligence but contains several provisions that may be invoked in cases of misuse of deepfakes.

Section 66C criminalizes identity theft and restricts the use of someone's electronic signature, password or unique identifying characteristic, which is relevant to deepfakes.

[14] *Subramanian Swamy v. Union of India,* (2016) 7 SCC 221.
[15] *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.

ISSN: 2582-6433

Cheating by impersonation via computer resources or a communication device may be caught by this provision. [16] Voice cloning scams and synthetic impersonations to secure funds or confidential information may be considered cheating by impersonation by computer resources/communications devices.

These provisions are relevant to non-consensual sexual deepfakes and intimate images created with AI.

The section 72 and punishable disclosure of personal information without permission may present a potential liability in respect of unauthorized collection and misuse for creating the Deepfakes.[17]

In addition, Section 69A gives the Government the power to issue blocking orders in response to online content that poses a specific threat to public order, national security or individual rights. [18]Blocking orders could be issued under Section 69A regarding deepfake content that poses a particular threat to national security, public order or individual rights.

However, the Information Technology Act was not created with the express intent of regulating synthetic media created with AI. This means that numerous harms associated with deepfakes are not covered by the scope of this law.

3.5 Digital Personal Data Protection Act, 2023

The Digital Personal Data Protection Act 2023 (DPDP Act) is one of India's main data protection laws and is particularly significant for deepfakes, given that their generation frequently involves the gathering and utilization of personal data.

Personal data is often used without the free, informed, specific, and unambiguous consent of the data principal by deepfake creators for purposes of personal identification. One of the common issues deepfake creators face is the lack of free, informed, specific, and unambiguous consent to process personal data via photos, videos, voice data, and other forms of personal identification.

Data Fiduciaries, as defined in Section 2(i) of the Act, could include social media platforms, AI developers and technology companies that are engaged in collecting or processing personal information.

[16] Information Technology Act, 2000, s. 66C.
[17] Id., s. 72.
[18] Id., s. 69A.

ISSN: 2582-6433

3.6 Bharatiya Nyaya Sanhita, 2023

The Bharatiya Nyaya Sanhita, 2023 (BNS) contains several provisions in the criminal law that could apply to criminal acts involving deepfakes.

Section 318 covers cheating and could be relevant when deepfakes are used to trick someone for monetary or personal enrichment. [19]

Sections 335 and 336[20] about forgery could apply if fake digital material is created with a motive to deceive or cause damage.

Section 356 of the Act criminalizes defamation and could be used in cases where deepfakes have undermined a person's reputation by depicting them in a harmful or offensive way. [21]

Depending on the circumstances of the allegations, there may be additional provisions that are relevant, such as those in relation to intimidation, harassment, extortion and offences relating to cyberspace.

3.7 Intermediaries are especially crucial in the fight against deepfakes since social media networks serve as the main channel for the propagation of synthetic content.

The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, make due diligence requirements on intermediaries. Platforms must make reasonable efforts to ensure that users do not host, display, upload, modify, publish, transmit, store, update or share illegal information (Rule 3). [22]

Intermediaries are required to remove content when they are provided with actual knowledge via an order from court or notification from government, if such content is to comply with legal requirements. Platforms must also implement grievance redressal processes, grievance officers to receive user complaints.

The government had also cautioned social media platforms to take immediate steps to prevent the spread of AI-generated abusive content and to abide by existing laws.[23] The government stated that deepfakes of imitated, obscene, misinformation, or privacy violations could lead to legal repercussions under current laws.

But there are still some concerns about intermediary liability and the liability of AI companies

[19] Bharatiya Nyaya Sanhita, 2023, s. 318.
[20] Id., ss. 335–336.
[21] Id., s. 356.
[22] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, r. 3.
[23] Vinita Singh and Ritu Gautam, *Cyber Crime, Security and Regulation in India* 147 (2022), available at: https://doi.org/10.55662/book.2022ccrs.005 (visited on June 23, 2026).

ISSN: 2582-6433

for creating generative models. As deepfakes become more prevalent, policymakers are urging the adoption of measures that can mitigate the negative aspects of their use, including watermarking, content labelling, disclosures of AI-generated content and fast takedown procedures.

3.8 The need for a comprehensive regulatory framework

The constitutional and legal aspects of deepfakes show that the existing regulation in India is not consolidated. Current laws on privacy, data protection, cybercrime, defamation, obscenity and intermediary liability provide some relief, but fall short of being adequate protections for the specific issue of AI-generated synthetic media. Deepfake dissemination is speedy, anonymous, scalable and cross-border, calling for a more coherent and specialised regulatory approach.

The proposed new legislation should define a clear legal definition of a deepfake, set out consent expectations for the use of biometric data, require transparency by the AI developers and platforms and offer robust remedies for victims. Any regulation should respect privacy, dignity, reputation and democratic integrity, respect freedom of expression and technological innovation.

4. Judicial Developments

Until now, Indian courts have not had a specific or extensive case law on deepfakes. But the current constitutional and cyber-law jurisprudence offers a valuable starting point for the development of a legal framework for synthetic media by AI. Privacy, dignity, reputation and the right to responsible online speech have already been identified as important legal values by the courts. It is possible to apply these rules to any dispute involving deepfakes.

4.1 K.S. Puttaswamy v. Union of India

The judgment K.S. Puttaswamy v. Union of India is highly relevant to deepfakes, as it safeguards informational privacy, bodily autonomy and control over personal data.[24] The typical deepfake is the unauthorized use of facial images, videos, voice samples and biometric data. This kind of action is in direct violation of a person's right to control his/her identity and digital persona. Hence the constitutional basis for regarding malicious deepfakes as privacy and autonomy violations is provided by Puttaswamy.

[24] *Supra note* 8.

ISSN: 2582-6433

4.2 Shreya Singhal v. Union of India

The significance of the judgment in Shreya Singhal v. Union of India is that it establishes that, online speech is protected under the Constitution, and that restrictions must be clear, reasonable and constitutionally valid. This principle should be considered when regulating deepfakes. Harmful deepfakes, such as those that are used in fraud, harassment, sexual exploitation or misinformation must be limited, but lawful deepfakes for satire, parody, artistic expression, or political criticism should not be banned. Keeping such laws from being too broad and overly restrictive is therefore essential for any future deepfake law.[25]

4.3 Subramanian Swamy v. Union of India

In Subramanian Swamy v. Union of India, the Supreme Court has held that criminal defamation is valid, and that reputation is an important part of the right to life as per Article 21.[3] Deepfakes have the potential to cause significant reputational damage by falsely attributing statements, actions, or appearances to a person that never actually happened. This decision paves the way for legal action against deepfake defamation, particularly when perpetrated using AI tools that inflict harm on personal dignity, professional reputation, or public image.[26]

4.4 Aveek Sarkar v. State of West Bengal

In Aveek Sarkar v. State of West Bengal, the Supreme Court discussed the legal test for obscenity and emphasized that content must be judged according to contemporary community standards. [27] This case is relevant to deepfakes involving obscene or sexually explicit content, particularly non-consensual intimate deepfakes.

5. Comparative Regulatory Approaches

5.1 European Union

The EU Artificial Intelligence Act is one of the most organised approaches taken by the European Union. It’s based on risk and demands transparency with AI-generated and manipulated content. Generally, deepfakes must be identified as 'artificial or manipulated' unless for legitimate purposes such as satire, art and security. This model is useful in India as it's a balance between regulation and freedom of expression.

[25] *Supra note* 10.
[26] *Supra note* 9.
[27] *Aveek Sarkar v. State of West Bengal*, (2014) 4 SCC 257.

ISSN: 2582-6433

5.2 United States

There is no single federal law about deepfakes, although there are laws in several states about certain harms. There are numerous state laws concerning election-related deepfakes, particularly demagogic videos spread prior to elections. There are also civil and/or criminal penalties in some states for non-consensual sexual Deepfakes. The U.S. practice is issue-oriented and does not encompass everything, but rather sexy privacy, elections, and fraud.

5.3 China

China has set very tough standards for synthetic media. It has regulations that mandate that content generated by AI be properly identified and not used in the production of misleading news, or fraudulent or violent materials. It's expected that platforms will also be checking users and tracking synthetic content. Platform responsibility and the State control of misinformation are key features of China's model.

5.4 United Kingdom

The UK is having to react to deepfakes by enacting changes in online safety and criminal law. There are calls for the UK to respond to deepfakes harms by making changes to online safety and criminal law. The Online Safety Act, 2023 sets out obligations for online platforms on how to minimise harmful content. The UK have also progressed to the criminalisation of sharing intimate deep-fake images without consent. This would have applicability in India regarding victim protection, harms online and platform accountability.

5.5 Australia

The Australian approach to addressing harms from deepfakes largely focuses on online safety policies and frameworks. The Online Safety Act, 2021 also gives powers to the safety Commissioner to act on image-based abuse, cyber abuse and harmful online content. If a non-consensual sexual deepfake is a component of image-based abuse it could be within any of the image-based abuse frameworks. Australia's model is helpful in establishing a specialized digital safety regulator with takedown powers.[28]

5.6 Singapore

In Singapore, there is a measure called Protection from Online Falsehoods and Manipulation

[28] Melanie Burton et al., "Regulating Image-Based Abuse: An Examination of Australia's Reporting and Removal Scheme" 2 *Journal of Online Trust and Safety* (2025).

ISSN: 2582-6433

Act, 2019 (POFMA) that attempts to tackle misinformation involving deepfakes. The law provides for correction directions, stop communication directions, takedown orders for false online content. While it does not apply to deepfakes exclusively, it can also be used to combat AI-generated misinformation. Singapore's emphasis on the use of speedy correction and removal mechanisms.[29]

5.7 Lessons for India

There are several lessons for India to take from these jurisdictions. To begin with, as the EU is doing, India should define deepfakes and mandate disclosure or labelling of synthetic media. Secondly, India, like the U.S. and UK, should develop robust countermeasures to non-consensual sexual deepfakes and election manipulation. Thirdly, India should also put in place some obligations for platforms about detecting, removing and redressing user complaints, like China and Australia. Last but not the least, India must be a balance act ensuring privacy, dignity, respect, and trust in democracy while not stifling satire, parody, innovation, and legitimate expression.

6. Recommendations

India needs a specific and balanced law on deepfakes. The law should clearly define deepfakes and classify harmful forms such as sexual deepfakes, financial fraud deepfakes, election-related deepfakes, and defamatory deepfakes. Clear definitions are necessary to avoid vague restrictions on online speech, as emphasized in *Shreya Singhal v. Union of India*.

Consent should be mandatory before using a person's face, voice, image, or body likeness for AI-generated content. This follows the privacy principle recognized in *K.S. Puttaswamy v. Union of India*, where informational privacy was held to be part of Article 21.

India should introduce mandatory labelling or watermarking of AI-generated content. This will help users identify synthetic media and reduce misinformation. Similar transparency duties are found in the EU AI Act and China's deep synthesis rules.

Platforms should be made accountable through quick takedown systems, grievance redressal, detection tools, and cooperation with law enforcement. The IT Rules, 2021 already impose due

[29] Adrian Ang and Adrian Ujin Yap, "Generative AI in Singapore" in *Oxford University Press eBooks* (2025), available at: https://doi.org/10.1093/oxfordhb/9780198940272.013.0042 (visited on June 23, 2026).

ISSN: 2582-6433

diligence duties, but deepfake-specific obligations should be added. [30]

Victims should receive fast remedies such as compensation, injunctions, anonymity, legal aid, and identity protection. Since reputation is protected under Article 21, as held in *Subramanian Swamy v. Union of India*, victims of defamatory deepfakes need effective civil and criminal remedies.

Stronger criminal penalties should apply to sexual deepfakes, election manipulation, financial fraud, repeated harassment, and deepfakes targeting children. Deeptrace Labs reported that about 96% of online deepfake videos were pornographic, mostly targeting women, while Sumsub reported a major rise in deepfake fraud globally.[31]

Finally, digital literacy campaigns are necessary. Citizens should be trained to identify AI-generated misinformation, verify suspicious content, and report harmful deepfakes. This is important because AI-generated misinformation has been identified as a major global risk.[32]

7. Conclusion

Deepfakes have emerged as one of the most complex legal and technological challenges of the artificial intelligence era. By using AI to create highly realistic but false images, videos, and audio, deepfakes blur the line between truth and fabrication. Their misuse can cause serious harm to privacy, dignity, reputation, personal autonomy, cybersecurity, and democratic trust. The danger is intensified by the fact that deepfake tools are increasingly accessible, inexpensive, and capable of producing convincing content within a short time.

In the Indian context, the issue is particularly serious because of the rapid growth of social media, digital communication, and online political discourse. Incidents involving celebrity deepfakes, voice-cloning frauds, and AI-generated misinformation show that deepfakes are no longer a distant technological concern. They have become a real threat to individuals, institutions, and democratic processes. Women, public figures, children, and ordinary internet users are especially vulnerable to reputational harm, sexual exploitation, identity misuse, and online harassment.

[30] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, r. 3.
[31] Deeptrace Labs, *The State of Deepfakes: Landscape, Threats and Impact* (2019), available at: https://deeptracelabs.com; Sumsub, *The State of Deepfakes 2023* (2023), available at: https://sumsub.com.
[32] World Economic Forum, *Global Risks Report 2024* (2024), available at: https://www.weforum.org.

ISSN: 2582-6433

The existing legal framework in India provides only partial protection. Provisions under the Information Technology Act, 2000, the Digital Personal Data Protection Act, 2023, the Bharatiya Nyaya Sanhita, 2023, and the Information Technology Rules, 2021 may be used in certain cases involving identity theft, cheating, obscenity, defamation, privacy violations, and intermediary liability. However, these laws were not designed specifically for AI-generated synthetic media. As a result, they do not fully address issues such as consent-based use of biometric data, AI-generated impersonation, rapid circulation of harmful content, platform accountability, and victim compensation.

Therefore, India requires a comprehensive and balanced legal framework for deepfakes. Such a framework should define deepfakes, classify harmful uses, require consent for use of facial and voice data, mandate labelling or watermarking of synthetic content, impose duties on platforms, and provide quick remedies for victims. Stronger penalties should apply to sexual deepfakes, financial fraud, election manipulation, and repeated harassment. At the same time, lawful uses such as satire, parody, education, research, and artistic expression must remain protected.

Ultimately, the regulation of deepfakes should not aim to suppress artificial intelligence but to ensure its responsible use. A rights-based and technology-sensitive law can protect individual dignity and democratic integrity while allowing innovation to flourish. India's response must be timely, clear, and victim-centred so that the legal system remains capable of addressing the harms of synthetic media in the digital age.