The Indian Journal for Research in Law and Management

Open Access Law Journal – Copyright © 2025

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

AI-GENERATED DEEPFAKES AND LEGAL ACCOUNTABILITY: ADEQUACY OF CYBER LAW FRAMEWORKS IN INDIA

~ Utkarsh Yadav

Introduction

Artificial Intelligence (AI) is one of the most disruptive technologies of the 21st century, affecting, challenging, and changing a myriad of areas, whether it be personal connection to the world, business, or governance. Among many forms of application is one of the most insidious aspects of AI, the use of deepfakes. Deepfakes uses AI methods, especially deep learning and generative adversarial networks (GANs) to manipulate audio-visual content in a manner that creates fake material that appears real. Although deepfake technology has legitimate uses in cinema, education, and accessibility, its potential for abuse should be taken seriously.

The risks of deepfakes are not theoretical; they are real social and political harms. Indeed, deepfakes were used for the purposes of misinformation, disinformation, influencing election outcomes, reputational harm, and non-consensual pornographic material (which overwhelmingly targets women) in a variety of contexts, around the world. This is especially pressing in India where the expansion of digital access is growing quickly. With both an active and polarized political discourse and historically entrenched gender bias in many forms, the use of deepfake could carry significant and harmful risks. The recent coordination of morphed videos during elections, and deepfake materials used for harassing women demonstrate the immediacy of the risks and harms.[1]

This paper aims to analyze whether or not India's current cyber law framework, based on the Information Technology Act and its subsidiary regulations from 2000, is adequate to face these challenges. Then, by comparing India’s legal position with international best practice, it argues that there are some remedies available, but in a piecemeal and outdated fashion. The central argument is that India's cyber threats, particularly deepfakes, require a broader, technology-sensitive legal framework to effectively respond to evolving threats.

Deepfake Technology and it's Social Impact

Deepfakes are a product of artificial intelligence and machine learning, which involve using algorithms to imitate and manipulate human features, voices, or actions in ways that can be almost imperceptible. The technique most frequently used involves Generative Adversarial Networks (GANs), which use two neural networks (one that makes the fake content and one that assesses it) to compete with each other until it is impossible to detect the generated content from real content. While this is certainly a sophisticated technological accomplishment, the credible creation and dissemination of deepfakes has entered into an easier and simpler period with the availability of open-source software and mobile applications.

The wrongful use of deepfake technology can be categorized broadly into three categories. Political: Deepfakes have been used, and have the potential use to spread political propaganda, manipulate election outcomes, and falsely attribute a speech to a political leader. Examples from around the world include the fabricated videos of U.S. President Barack Obama and Ukraine’s President Volodymyr Zelenskyy, each sparking public debate about their implications for democratic fragility. India, where elections are heatedly competitive, can easily be impacted negatively from the use of a single manipulated video concession to elicit extreme polarization over a messaging that undermines trust in an institution.

Third, misinformation and damage to reputation: whether fake news stories about the pandemic or fake financial solicitations, deepfakes allows for scams to flourish, eroding the public's trust in digital media as a whole.

And if there are specific points I can make with respect to deepfakes, the broad social ramifications are serious.[2] Deepfakes undermine trust in media and organizations, present uncertainty and ambiguity about truth and falsehood, and invade privacy parameters. They may create chilling effect toward free expression. Indeed, if people fear their image can be misused or manipulated, they may become less willing to speak freely or act without regard for their image. In this sense, deepfakes are not simply a technological issue, but a societal and democratic problem of the highest order.

The Indian Legal Framework on Deepfakes

India's legal response to deepfakes draws upon existing statutes, including primarily the

Information Technology Act, 2000 (the IT Act) and the Indian Penal Code (IPC). Both statutes were written prior to the global rise of AI and contain no express mention of synthetic media.

The IT Act, under section 66D, punishes a person who is said to have committed "cheating by impersonation" with the use of computer resources that could apply in cases of fraud or deception using deepfakes. Section 66E includes the breach of privacy offenses whereby a person is said to commit breach of privacy by capturing someone's image without their consent or by the transmission of one's image without their consent, all of which can be applicable to non-consensual deepfake porn. Sections 67 and 67A criminalize the publishing or transmitting of obscene and sexually explicit content, which could be useful in a prosecution for morphed video content that circulates online. In all cases, deepfakes are not mentioned and the interpretation and application of these provisions do not provide much clarity regarding technical usages like deep fakes, which hinders the prosecution of offenders.

Violations of the same nature are also derivable under the Bharatiya Nyaya Sanhita, 2023 (BNS) - provisions with the same effect as found in the IPC. For example, defamation, previously provided under Section 500 of the IPC, has a corresponding provision in Section 354 of the BNS, which would apply in the case where the deepfake was created or shared with the intention to harm someone's reputation. Likewise, the offence of insulting the modesty of a woman was previously found under Section 509 of the IPC, however it can be found in Section 79 of the BNS, and it would extend to deepfakes that were made to humiliate or degrade a woman. Furthermore, the unlawful act of voyeurism under Section 354C of the IPC has been continued into Section 77 of the BNS, which provides for unlawful capturing or distributing the image of a woman who is engaging in a private act without her consent, which would include manipulated or pornographic deepfake. The law does provide for some potential avenues of accountability; however, the practical hurdles are still significant especially with respect to proving that images or videos were created or manipulated, determining what jurisdiction to pursue given that they are most often shared internationally, and the ability to actually act upon such enforcement in an environment where such content spreads like wildfire over digital channels.

The constitutional regime creates another layer. Article 19(1)(a) guarantees freedom of speech and expression, while Article 19(2) provides reasonable restrictions on freedom of speech and expression to protect public order, morality and reputation. Deepfakes walk a fine

line; the absence of suitable restrictions could chill free expression, while unbridled misuse could infringe the democratic and individual dignity of others. In addition, Article 21 provides the right to privacy, which builds from the important judgment Justice K.S. Puttaswamy v. Union of India (2017), where the right to privacy was recognized as a fundamental right in India. This means non-consensual deepfakes infringe the sanctity of the Constitution.

Judicial pronouncements uncover yet another layer of complexity. In Shreya Singhal v. Union of India (2015)[3], the Supreme Court interpreted the impact of Section 66A of the IT Act and found it overly vague and therefore unconstitutional as it curtailed free speech. Since it has been declared unconstitutional; Indian law has lacked a remedy for online harms, such as misinformation, including deepfake public discourse. The Puttaswamy judgment[4] expanded the understanding of privacy, as the right to dignity is also a recognized aspect of privacy, and provides citizens with a constitutional lens, but there is no direct statutory approach and there are challenges in providing evidentiary authority and jurisdiction to pursue complaints.

So while the current context provides overlapping legal options for addressing some aspects of deepfake technology, and incoherent legal remedies, there are many issues, including an absence of the recognition of synthetic media and legal tools, evidentiary hurdles, and jurisdictional issues that inhibit public interest objectives of lawmaking. The complexity of evolving technology and harm underscores the need for bespoke legislation aimed at addressing deepfake technology.

Difficulties in Legal Accountability

The regulation of deepfake content faces numerous interrelated challenges which extend beyond compliance with the law to include various technical challenges which exist on their own level. Firstly, at the technical level, detection is a significant issue in and of itself. Artificial intelligence is evolving rapidly, enabling deepfakes to become so realistic that conventional forensic identification tools may not even recognize them. The task of authentication is further complicated when infringing content is widely propagated, disseminated without accompanying metadata, or manipulated in a way that traverses platforms. Similar challenges arise with the attribution or tracing of the creator or distributor of the deepfake, especially in instances where creators hide their identity using identity masking or anonymous communication technologies.

Secondly, on the legislative side, the existing provisions in the Information Technology Act and

IPC relating to this content were enacted without any consideration for synthetic media, and need new legislation to bring them into the current century. There are sections dealing with impersonation, obscenity, or defamation that could apply, but would not appropriately account for the new harms associated with AI-generated content. Financial constraints also complicate this situation. Many deepfakes originate outside of India's geographical territory, but are then transmitted into India, which raises issues regarding enforcement across borders and working with foreign corporate platforms. Additionally, the lack of a specific statutory reference to "deepfakes" or "synthetic media" means that laws will be ineffective to enforce, with courts and enforcement agencies pressed to make strained interpretations instead.[5]

There are ethical dilemmas, especially when it comes to managing regulation and adhering to constitutional guarantees of free expression. As an example, allowing excessive censorship as part of the approach to banning deepfakes is going to contribute to a culture that chills legitimate speech, parody and artistic freedom. On the other hand, under-regulation of deepfakes poses risks to distinct personal privacy, political integrity and gender justice.

Enforcement challenges have perhaps been the most visible. There is a general lack of competent forensics cyber cells in India, trained to counter AI-led crime. Case law reported in courts often takes too long for victims to see a resolution, particularly in viral cases, instances of bullying and harassment online.

Additionally, low digital literacy among Indian citizens makes them even more vulnerable, as victims may not even know there are ways to redress, benefit from preventive measures, or even know who to complain to in a situation where the deepfake victim does not receive validation; given the low barriers to accountability for perpetrators, a victim will always be investigating under an assumption that their victimhood can move onto the next 'social media level'.

In conclusion, legal accountability for deepfakes is complex across technological, legal, ethical and institutional issues that do not appear achievable without urgent reform. If India does not accelerate these reforms, and continues to reactively respond to new technologies like deepfakes, it will also have to apply the same reactive response to the dangers of deepfakes in the future.

Recommendations for Reform

Addressing the threat of deepfakes requires a multi-pronged reform agenda that combines legislative, regulatory, technological, and institutional measures.

On the legislative front, India must either amend the Information Technology Act, 2000 or enact a dedicated "Deepfake Regulation Act." Such a framework should explicitly criminalize non-consensual deepfake pornography, recognizing its gendered nature and devastating impact on victims. Provisions mandating disclosure or watermarking of AI-generated and synthetic content can also help distinguish genuine media from manipulated versions. Clear statutory language will reduce interpretive ambiguity and ensure accountability.

Regulatory reforms must also target intermediaries. The Intermediary Guidelines and Digital Media Ethics Code, 2021, provide a foundation for imposing due diligence obligations on platforms. These should be expanded to specifically address deepfakes, requiring rapid takedowns, stronger verification mechanisms, and proactive monitoring of manipulated content. Accountability must extend to social media platforms that host or circulate deepfakes, ensuring they play an active role in mitigation.

Technological solutions are equally vital. Investment in AI-driven detection tools, in partnership with global technology companies and research institutions, is essential to stay ahead of increasingly sophisticated fabrications.

Alongside this, India must strengthen digital literacy campaigns to educate citizens on recognizing, reporting, and safeguarding against deepfakes. Public awareness can significantly reduce susceptibility to misinformation and exploitation.

Judicial and institutional reforms are needed to bridge enforcement gaps. Fast-track courts for cybercrimes could provide timely redress in cases where viral circulation causes immediate harm. A dedicated National Deepfake Task Force, comprising legal experts, technologists, and enforcement agencies, could serve as a central authority for monitoring, policy-making, and inter-agency coordination.

Comparatively, India can draw lessons from global models. The European Union's Digital Services Act emphasises platform accountability, while China has introduced binding obligations on watermarking and disclosure of synthetic content. Adapting these measures within India's democratic and constitutional framework can strike a balance between combating misuse and preserving freedom of expression.

In essence, effective reform must go beyond piecemeal fixes and build a comprehensive, forward-looking ecosystem capable of tackling the evolving challenge of deepfakes.

Conclusion

In conclusion, concepts such as deepfakes threaten individual dignity, social trust, and the stability of democratic institutions. The risks around deepfakes, as evidenced in the political, pornographic, and misinformation contexts, have demonstrated the lawful deficiencies of India's digital social and democratic systems. However, India's approach to deepfake regulation, which is largely dependent on the IT Act, 2000, and the limited provisions in the IPC, is incapable of effectively responding to the potential privacy and reputational threats associated with synthetic media. India needs a meaningful response that both criminalizes harmful deepfakes, burdens platforms, provides meaningful investment in detecting technologies, as well as ensures the emancipation of victims while preserving freedom of expression. Without substantial reforms, deepfakes will not only continued impact individuals, but will also threaten institutional credibility and the public conversation.