
DEEPAKES AND DEMOCRACY: CONTROLLING AI-GENERATED MISINFORMATION IN INDIA

Mouneka Selvam, B.A., LL.B., The Central Law College, Salem

ABSTRACT

The accelerated advancement of Artificial Intelligence has transformed the creation and circulation of digital content. Among its most far-reaching implications are deepfakes, which adopt artificial intelligence to create or influence audio, video, pictures, or other digital contents so credibly that fabricated events may appear authentic. While synthetic media can advance a legitimate objective in education, entertainment, access and political disclosure, its misuse creates significant coercion to privacy, human dignity, standing, freedom of speech and, most significantly, democratic decision-making. India brought a framework through, the IT Act, 2000, some electoral guidelines brought by the Election Commission of India, the IT rules, 2021, and related criminal and data privacy laws. The 2026 amendments to the Information Technology Rules determine an important advancement by explicitly regulating through a structured framework generated information and requiring clear disclosure and technological origin mechanisms. However, obstacles remain concerning execution, technological detection, intermediary accountability and safeguarding of legitimate political expression. This essay submits a rights-based regulatory framework that addresses deceptive and abusive uses of AI-generated media while safeguarding satire, critique, artistic expression and legitimate political discourse.

Keywords: AI Media, Technology Regulation, Deepfakes, Information Technology.

I. INTRODUCTION

Democracy depends upon the capacity of citizens to acquire information, evaluate conflicting political claims and make independent choices. The reliability of information is particularly substantial during elections because voters exercise constitutional power on the basis of political communication. The development of machine-generated artificial intelligence has intricately this process by making it progressively easy to create realistic but completely fabricated audio-visual material.

Deepfakes are specifically sophisticated form of synthetic media in which AI is used to create or manipulate a person's visual representation, voice, conduct or statements. The technology may be used lawfully for entertainment, education, creative representation or accessibility. The legal difficulty emerges when synthetic content is Portrayed as authentic and is disseminated with an objective or effect of misleading the public.

The democratic importance of deepfakes does not lie simply in the probability that citizens may believe false information. A fabricated video may misrepresent a political candidate making a provocative statement, announcing a false policy or expressing views that the candidate never expressed.¹

The regulatory framework has since developed further. The IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 were amended in February 2026 to precisely address. “Synthetically generated information.” the amendment needs intermediaries promoting synthetic content to apply technical measures against unlawful synthetic information and, for other synthetic information, demand prominent labelling and technical source-verification mechanisms where technically feasible. This development raises an important question: **How can India regulate misleading deepfakes without Manipulating false information governance into a mechanism for restraining legitimate political speech?**²

This essay affirms that Indian legal framework should implement a rights-based and risk-based approach. The objective should not focus on prohibiting artificial intelligence or synthetic media as such, but to regulate fraud and harmful uses while safeguarding legitimate political

¹UNESCO, Artificial Intelligence, Freedom of Expression and Elections: Safeguarding Democratic Processes (2025).

²Election Commission of India, Advisory on use of Artificial Intelligence (AI)-based tools and deepfakes during elections (Jan. 16, 2025).

evaluation, satire, parody, creative portrayal and democratic discourse.

II. DEEPFAKES AND THEIR DEMOCRATIC IMPACT³

The most significant threat of deepfakes is their capacity to blur the distinction between reality and fabrication. While conventional misinformation may depend on the acceptance of textual claims, deepfakes can produce highly realistic audio and visual material that appears to substantiate false events. This issue is specifically serious in political communication. A fabricated video showing a candidate making a provocative statement may be circulated shortly before an election. Accordingly, the legal problem is not merely whether the content is misleading but whether the pace, extent and technological credibility of the content can intervene with informed democratic choice. If citizens become conscious of realistic audio and video can easily be constructed, genuine recordings may also be rejected as “deepfakes.” Thus, synthetic media can undermine both the reliability of false information and the reliability of truthful information.⁴The ECI’s advisory expressly acknowledged in January 2025 that AI-generated images, videos and audio have the prospective to impact voters’ opinion and public trust.

III. FREEDOM OF SPEECH AND THE REGULATION OF DEEPFAKES

The Article 19(1)(a) of constitution guarantees freedom of speech and expression. Political expression particular significance in at democratic society, as it enables individual took question, criticize and express opinions regarding government, political party and public authorities. The supreme court decision in *Shreya Singhal v. Union of India* is particularly important in this context. The court invalid section 66a of information Technology act 2000 on the ground that its vague and excessively broad wording imposed and unconstitutional restriction on online speech. The code distinguished between discussing advocacy and incitement and emphasized that restrictions on speech must fall within the grounds specified in article 19 (2). these principles are highly relevant to the regulation of deepfake. A law that simply prohibits “false”, “offence” or “misleading” online content without Clear legal standards could produce the same chilling -effect concerns that influenced the court in *Shreya Singhal*. Accordingly, deepfake regulation should interesting distinguish between legitimate

³Constitution of India, arts. 19(1)(a), 19(2).

⁴OECD, Facts not Fakes: Tackling Disinformation, Strengthening Information Integrity (OECD Publishing 2024), <https://doi.org/10.1787/d909ff7a-en>.

synthetic expression and deceptive synthetic representation. A clearly labelled parody of a political leader should ordinarily be treated differently from a labelled AI- generator video falsely presented as a genuine statement by that leader. The former communication an idea through satire the latter attempts to manufacture evidence of an event that never occurred. The supreme court's recognition of internet- based expression as constitutionally protected future strengthens the need for proportionality. In *Anuradha Bhasin v. Union of India*, the court recruitment freedom of speech and expression through the inner internet receive constitutional protection under article 19(1)(a). Therefore, regulation of online synthetic media must satisfy constitutional standards rather than operate merely as an administrative mechanism.^{5 6 7}

IV. PRIVACY, DIGNITY AND REPUTATION

In *Justice K.S. Puttaswamy (Retd.) v. Union of India*, a nine-judge Bench acknowledged as a constitutionally protected right and linked privacy with autonomy, dignity and individual liberty. ⁹ the unlawful creation and dissemination of synthetic representations may intervene with these interests. An individual's face, voice or resemblance can be appropriated and used to create content that the person never created. The harm becomes particularly serious where synthetic media is sexually unambiguous, humiliating or defamatory.⁸

Deepfake technology creates the application of this doctrine more intricate, as fabricated content may not simply disclose private information; it can also generate entirely fictitious incident or situation involving the individual. The Information Technology Act, 2000 contains stipulations that may become relevant depending upon the nature of the misuse. Section 66C examines identity theft, Section 66D examines cheating by personation using a computer resource, and Section 66E examines violation of privacy. Sections 67, 67A and 67B may also extend to specific obscene, sexually explicit or child sexual-abuse material communicated electronically.

The legal framework therefore provides numerous possible remedies. Moreover, these provisions were not initially designed specifically for generative artificial intelligence. Applying these provisions to increasingly advanced synthetic media may rise questions regarding identification of responsibility, purpose, and jurisdiction and the establishment of

⁵Constitution of India, art. 19(1)(a).

⁶Constitution of India, art. 19(2).

⁷Information Technology Act, 2000, No. 21 of 2000, §§ 66C, 66D, 66E, 67, 67A, 67B, India Code.

⁸Information Technology Act, 2000, §§ 69A, 79, India Code.

evidence.^{9 10}

V. INDIA'S DEVELOPING REGULATORY FRAMEWORK¹¹

India has not historically relied upon a single “deepfake law.” Instead, the issue has been addressed through overlapping provisions of information technology, criminal law, privacy law and election regulation.

The Information Technology Act, 2000 remains the principal statutory foundation for intermediary regulation and several forms of cyber misconduct. Section 69A permits the Government to direct blocking of public access to information in specified circumstances, while Section 79 provides the statutory framework for intermediary safe harbor subject to prescribed conditions.

In *Shreya Singhal*, the Supreme Court upheld Section 69A and the associated blocking framework, while reading down the intermediary liability framework so that intermediaries would not become private adjudicators of unlawful speech merely upon receiving every private complaint. This remains relevant to contemporary debates concerning platform responsibility for synthetic media. The Information Technology Rules, 2021 provide additional due-diligence obligations for intermediaries. The February 2026 amendments mark a substantial development because they expressly introduce obligations concerning “synthetically generated information” (“SGI”).¹²¹³

Under the amended framework, where an intermediary facilitates the creation or dissemination of SGI, it must deploy reasonable and appropriate technical measures to prevent synthetic information that violates applicable law. The rules specifically refer to unlawful material including non-consensual intimate imagery, privacy-invasive content and false electronic records. They also address synthetic content that falsely depicts or portrays a natural person or real-world event in a manner likely to deceive by misrepresenting identity, voice, conduct,

⁹Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, rr. 3(2)(b), 3(3), 4(1A), as amended Feb. 10, 2026.

¹⁰Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, r. 3(3), as amended Feb. 10, 2026.

¹¹Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, r. 3(3), as amended Feb. 10, 2026.

¹²Election Commission of India, Guidance/Advisory to political parties on responsible use of social media and AI-generated/deepfake content (May 6, 2024).

¹³Election Commission of India, Advisory on use of Artificial Intelligence (AI)-based tools and deepfakes during elections (Jan. 16, 2025).

action, statement or the occurrence of an event.

The amendments further require synthetic information that does not fall within the prohibited categories to be prominently labelled as synthetically generated and, to the extent technically feasible, embedded with permanent metadata or other technical provenance mechanisms. The rules also restrict the modification, suppression or removal of such labels and metadata. These provisions represent a significant shift from a purely reactive approach towards a model based upon transparency, provenance and technological accountability.

VI. ELECTIONS AND DEEPPAKES

Elections comprise a particularly sensitive framework because misinformation can influence democratic choice at a time when the scope for rectification is limited. The ECI's May 2024 directions required political parties to abstain from using deepfakes and AI-generated misrepresented content that could circulate fake information, misinformation or disinformation. The ECI's January 2025 advisory went additionally in recommending prominent labels such as "AI-Generated," "Digitally Enhanced" or "Synthetic Content" and disclosure in electoral materials where synthetic content is used.¹⁴

The framework is significant because it does not forbid all AI-generated political content. Rather, it seeks to preserve voter awareness by securing that citizens know when political communication has been significantly generated or altered using AI.¹⁵

Political parties may legally use technology to communicate with voters, translate speeches, generate availability tools or create campaign material. A blanket prohibition would impose unwarranted restrictions on political expression. The emphasis should therefore remain on deception and concealment, particularly where synthetic content is presented as authentic evidence.

Nevertheless, election-related regulation must also include procedural safeguards. The ECI, government authorities and digital platforms should not have unrestricted discretion to decide what constitutes political fabrication. Definitions should be clear, removal decisions should be

¹⁴Election Commission of India, Advisory on use of Artificial Intelligence (AI)-based tools and deepfakes during elections (Jan. 16, 2025).

¹⁵Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, r. 4(1A), as amended Feb. 10, 2026.

reasoned, and mechanisms for review and appeal should be available.

VII. INTERMEDIARY LIABILITY AND PLATFORM ACCOUNTABILITY

Social-media platforms are central to the dissemination of deepfakes because their recommendation and sharing systems can amplify content at extraordinary speed. The 2026 IT Rules impose specific responsibilities on intermediaries that facilitate synthetic content. For significant social-media intermediaries, the rules require users to declare whether information is synthetically generated and require technical measures to verify such declarations. Where synthetic content is identified, it must be clearly and prominently labelled. This represents an important recognition that the platform environment itself contributes to the scale of synthetic-media harms.¹⁶

Nevertheless, imposing unlimited liability upon platforms would be problematic. Automated detection systems are imperfect. False positives could result in genuine political speech being labelled or removed, while sophisticated deepfakes may escape detection.

A sensible approach is therefore one of **graduated platform responsibility**. The law should also distinguish between ordinary users, content creators, AI developers and large platforms. Their technical capacities and roles in the creation and dissemination of synthetic content are not identical.

VIII. THE CHALLENGE OF DETECTION AND PROVENANCE¹⁷

Technology cannot provide a complete answer to the legal problem. Deepfake detection systems may become outdated as generative models improve. A system capable of identifying one generation method may fail when a different technique is used.¹⁸¹⁹

For this reason, the 2026 emphasis on provenance is significant. Rather than relying entirely on after-the-fact detection, provenance mechanisms can help establish information about the

¹⁶Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, r. 4(1A), as amended Feb. 10, 2026.

¹⁷K.S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1; Anuradha Bhasin v. Union of India, (2020) 3 SCC 637.

¹⁸National Institute of Standards and Technology (NIST), Examining Digital Media: Terminology for Digital Media (2025).

¹⁹Haiying Guan, James Horan & Andrew Zhang, Guardians of Forensic Evidence: Evaluating Analytic Systems Against AI-Generated Deepfakes, NIST Forensics NIST 2024 (2025).

origin or synthetic nature of digital material.

However, metadata and watermarking also have limitations. Content may be copied, compressed or transferred between platforms, potentially affecting technical markers. Consequently, provenance should supplement rather than replace legal remedies, platform governance and public awareness.²⁰

The long-term solution must therefore be technological as well as legal. Courts, investigators, platforms and users require better tools for authenticating digital evidence. At the same time, individuals must be educated about the possibility that apparently authentic audio-visual material may be manipulated.²¹

IX. THE RISK OF OVER-REGULATION

The most significant constitutional danger in regulating deepfakes is not under-regulation alone; it is over-regulation. A broad legal power to remove “false” online information could easily be used against political criticism, investigative journalism, satire or dissent. In a democracy, governments cannot be the final arbiters of every contested political claim.

The Supreme Court's jurisprudence on free speech requires particular caution. *Shreya Singhal* demonstrates the constitutional problem with vague restrictions, while *Subramanian Swamy v. Union of India* illustrates the Court's recognition that reputation is a legitimate constitutional interest capable of supporting restrictions within Article 19(2).²²

Deepfake regulation should therefore adopt the least restrictive mechanism capable of addressing the harm. For example, labelling may be preferable to removal where synthetic content is lawful but potentially misleading. Removal may be justified where content is unlawful, impersonating, privacy-invasive or demonstrably deceptive in a manner causing serious harm. Criminal sanctions should be reserved for serious misconduct satisfying the requirements of criminal law. Such a graduated approach would protect both democratic integrity and constitutional liberty.

²⁰National Institute of Standards and Technology (NIST), GenAI: Deepfakes 2026, AI Challenges and Forensics (2026).

²¹UNESCO, Artificial Intelligence and Disinformation: Examining Challenges and Solutions (2019).

²²*Subramanian Swamy v. Union of India*, (2016) 7 SCC 221.

X. RIGHTS-BASED FRAMEWORK FOR INDIA²³

India should build upon the existing framework through five principles.

1. Clear definitions

The law should distinguish between synthetic media, manipulated media, satire, parody, artistic expression and materially deceptive synthetic representations. Vague terms should be avoided because they create uncertainty and chilling effects.

2. Mandatory disclosure in high-risk contexts²⁴

Political campaign material containing substantial synthetic content should carry prominent and understandable disclosures. Citizens should not be required to possess technical knowledge to determine whether a political video is authentic.

3. Rapid remedies for victims

Individuals whose likeness, voice or identity is misused should have accessible procedures for reporting harmful synthetic content. The current IT Rules' specific attention to impersonation and artificially morphed images provides a useful foundation.

4. Independent and transparent review

Where content is removed or restricted, users should have access to an effective review mechanism. Decisions affecting political expression should receive heightened procedural scrutiny.

XI. CONCLUSION

India's regulatory response has evolved considerably. The Information Technology Act, constitutional jurisprudence, intermediary rules and election-related directions already provide several legal mechanisms. The February 2026 amendments to the IT Rules are particularly significant because they expressly address synthetically generated information, deceptive

²³Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, r. 3(2)(b), as amended Feb. 10, 2026.

²⁴Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, rr. 3(3), 4(1A), as amended Feb. 10, 2026.

representations, prominent labelling and technical provenance. Consequently, the legal response must distinguish between synthetic expression and deceptive synthetic representation. The appropriate objective is therefore not to eliminate deepfakes but to minimize their capacity to deceive and cause unlawful harm. Transparency, labelling, provenance, intermediary accountability, effective victim remedies and digital literacy should operate together within a constitutionally proportionate framework. The central democratic challenge is ultimately one of trust. India's future regulatory framework should therefore seek to preserve both the freedom to speak and the public's ability to know when speech has been artificially manufactured.