LEGAL ACCOUNTABILITY IN AI-GENERATED DEEPFAKE POLITICAL CAMPAIGNS

Aditya Mishra, National Law Institute University, Bhopal

Akshat Mishra, National Law Institute University, Bhopal

ABSTRACT

The emergence of artificial intelligence-generated deepfakes in political campaigns poses a fundamental threat to democratic integrity, challenging traditional legal frameworks designed for pre-digital electoral processes. This research investigates how India's existing cyber laws fail to adequately address AI-generated political disinformation, particularly following widespread deepfake deployment during the 2024 General Elections that manipulated voter perceptions while evading legal accountability. Deepfake technology enables the creation of synthetic videos, audio recordings, and images that convincingly portray political figures making statements or engaging in activities they never performed, achieving viral dissemination within hours and far outpacing traditional legal remedies that require days or weeks for judicial intervention. India's current legal architecture comprising the Information Technology Act 2000, IT Rules 2021, and the Representation of People Act 1951 demonstrates critical inadequacies when confronting AI-generated political manipulation, struggling with definitional ambiguities, enforcement delays, and jurisdictional complexities inherent in cross-border digital campaigns. Criminal provisions addressing impersonation and forgery prove insufficient for prosecuting creators of synthetic political content, while intermediary liability rules fail to incentivize effective detection and removal mechanisms. This legal vacuum enables malicious actors to deploy deepfake campaigns with minimal accountability risks, potentially distorting electoral outcomes through systematic disinformation. The research demonstrates that protecting electoral integrity in the AI era requires comprehensive legal reforms addressing synthetic media's unique characteristics, including specialized detection obligations, expedited judicial procedures, and enhanced international cooperation mechanisms to prevent sophisticated technological manipulation that threatens informed electoral participation.

Keywords: Artificial Intelligence, Deepfake Technology, Political Campaigns, Electoral Integrity, Cyber Law, Legal Accountability

I. INTRODUCTION

The advent of artificial intelligence has fundamentally transformed the landscape of political discourse, introducing unprecedented challenges to the integrity of democratic processes.[1] The 2024 Indian General Elections marked a watershed moment in this evolution, witnessing the widespread deployment of AI-generated deepfake technology to create synthetic political content that blurred the lines between authentic campaign messaging and sophisticated digital manipulation. From fabricated speeches by political leaders to entirely synthetic endorsements, the electoral battleground became a testing ground for technologies that could fundamentally undermine the informed consent that forms the bedrock of democratic governance.

The proliferation of deepfake technology in political campaigns represents more than a mere technological curiosity; it constitutes an existential threat to electoral integrity. During the 2024 elections, documented instances emerged of AI-generated videos showing political leaders making statements they never uttered, endorsing candidates they never supported, and engaging in activities that never occurred. These synthetic media productions, often indistinguishable from authentic content to the untrained eye, spread across social media platforms with viral efficiency, reaching millions of voters before fact-checking mechanisms could respond.[2] The sophistication of these deepfakes, powered by advanced Generative Adversarial Networks (GANs)[3], rendered traditional methods of content verification inadequate and exposed critical vulnerabilities in India's electoral safeguards.

The current legal framework governing such digital manipulation presents a complex web of fragmented provisions scattered across multiple statutes. The Information Technology Act, 2000, primarily conceived in an era preceding sophisticated AI capabilities[4], struggles to address the nuanced challenges posed by synthetic media. While Section 66D criminalizes cheating by personation using computer resources, its application to AI-generated political content remains largely untested and procedurally complex. Similarly, the IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, despite incorporating provisions

[1] Klaus Schwab, The Fourth Industrial Revolution (Crown Business 2017) 47-52.
[2] Hunt Allcott and Matthew Gentzkow, 'Social Media and Fake News in the 2016 Election' (2017) 31 Journal of Economic Perspectives 211, 213-215.
[3] Ian Goodfellow and others, 'Generative Adversarial Networks' (2014) 27 Advances in Neural Information Processing Systems 2672.
[4] Information Technology Act 2000, s 66D.

against impersonation, lack the specificity and enforcement mechanisms necessary to address the speed and scale at which deepfake political content operates.

The inadequacy of existing legal mechanisms becomes particularly pronounced when examined against the temporal constraints of electoral processes. Traditional legal remedies, designed for conventional forms of defamation or electoral malpractice, prove insufficient when confronted with content that can be created, disseminated, and achieve widespread impact within hours. The conventional judicial process, with its emphasis on due process and evidentiary standards, finds itself outpaced by the velocity of digital disinformation campaigns. Moreover, the challenge of attribution—determining the original creator of AI-generated content—introduces evidentiary complexities that existing legal procedures are ill-equipped to handle.

This research endeavors to address several critical questions that have emerged from the intersection of AI technology and electoral law. How effectively do India's current cyber laws address AI-generated political disinformation? What enforcement challenges arise when attempting to apply traditional legal frameworks to sophisticated technological manipulation? How have other jurisdictions approached the regulation of deepfake content in political contexts, and what lessons can inform India's legislative response? Most fundamentally, what comprehensive reforms are necessary to establish meaningful legal accountability for AI-generated political manipulation while preserving the democratic values of free expression and robust political debate?

Through a comparative analysis of Indian cyber law provisions against global regulatory approaches, including the European Union's Digital Services Act, various U.S. state deepfake statutes, and the United Kingdom's Online Safety Act, this paper seeks to illuminate the path toward effective legal accountability. The methodology employed combines doctrinal legal analysis with empirical examination of recent electoral incidents, providing both theoretical foundation and practical insight into the challenges facing democratic institutions in the age of artificial intelligence.

The central thesis of this research posits that while India's existing cyber law framework provides fragmented protection against AI-generated political disinformation, the absence of specific deepfake legislation and robust enforcement mechanisms creates accountability gaps that fundamentally undermine electoral integrity, necessitating comprehensive legal reforms

informed by global best practices and tailored to India's constitutional framework and democratic traditions.

II. CONCEPTUAL FRAMEWORK AND DEFINITIONAL ANALYSIS

A. Understanding Deepfakes: Technology and Taxonomy

The term "deepfake" represents a portmanteau of "deep learning" and "fake," encapsulating the sophisticated artificial intelligence techniques employed to create synthetic media content. At its technological core, deepfake creation relies on Generative Adversarial Networks (GANs), a machine learning architecture comprising two neural networks—a generator and a discriminator—engaged in adversarial training. The generator creates synthetic content while the discriminator attempts to identify fabricated material, resulting in increasingly sophisticated and realistic artificial media through iterative improvement.

From a legal perspective, deepfakes constitute a subset of synthetic media, encompassing video, audio, and image content that has been artificially generated or substantially manipulated using artificial intelligence algorithms. The taxonomy of deepfakes relevant to political contexts includes: facial reenactment deepfakes, where a target individual's facial expressions and movements are synthesized onto existing video content; speech synthesis deepfakes, utilizing voice cloning technology to generate artificial audio of political statements; and full-body puppetry deepfakes, creating entirely synthetic video content featuring political figures in fabricated scenarios.

Political deepfakes represent a specialized category distinguished by their electoral context and potential democratic impact. Unlike entertainment-oriented synthetic media or commercial deepfakes, political deepfakes are characterized by their temporal sensitivity—often deployed during critical electoral periods—and their capacity to influence voter behavior through misinformation or manipulation of candidate perception.

B. Legal Definitions and Terminological Challenges

The absence of standardized legal definitions for AI-generated content creates significant interpretative challenges across jurisdictions. The European Union's Digital Services Act

employs the term "synthetic media"[5] to encompass "audio, image or video content that has been generated or substantially modified by automated means," providing a technology-neutral framework that captures various forms of artificial content manipulation. Conversely, several U.S. state statutes adopt more specific terminology, with Texas defining "deepfake video"[6] as video content "created with the intent to deceive" using "artificial intelligence or machine learning."

The distinction between "disinformation" and "misinformation" assumes critical importance in the legal context of political deepfakes. Disinformation, characterized by intentional falsity and malicious distribution, typically attracts criminal sanctions and civil liability. Misinformation, involving false information shared without malicious intent, may warrant corrective measures but generally receives lesser legal consequences. This distinction becomes particularly complex in the context of AI-generated political content, where determining intent requires sophisticated technical analysis and investigative capabilities.

Electoral law definitions present additional complexity. The Representation of People Act, 1951, defines "election advertisement" broadly to include any content calculated to influence voter choice, potentially encompassing deepfake political content. However, the statute's pre-digital terminology struggles to address synthetic media's unique characteristics, particularly regarding attribution and authenticity verification.

C. Jurisprudential Foundations

The legal framework governing political deepfakes operates within established jurisprudential principles that balance competing constitutional rights and democratic imperatives. The Supreme Court's decision in Shreya Singhal v. Union of India (2015)[7] established that content regulation must satisfy strict constitutional scrutiny, requiring clear definitional boundaries and procedural safeguards against arbitrary enforcement.

The doctrine of technology neutrality, articulated in various information technology judgments, suggests that legal principles should apply consistently across technological platforms. However, the unique characteristics of AI-generated content—particularly its capacity for mass

[5] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services (Digital Services Act) OJ L277/1, art 3(s).pfakepfak
[6] Tex Civ Prac & Rem Code § 98C.101 (2019).
[7] Shreya Singhal v Union of India (2015) 5 SCC 1.

deception and rapid dissemination—may justify specialized regulatory approaches that depart from traditional technology-neutral frameworks.

Electoral integrity emerges as a compelling state interest justifying content regulation, as recognized in Association for Democratic Reforms v. Union of India (2002)[8], where the Supreme Court emphasized the fundamental importance of informed electoral choice. This precedent provides constitutional foundation for regulating synthetic political content that undermines voter knowledge.

D. Constitutional Framework

Article 19(1)(a) of the Indian Constitution guarantees freedom of speech and expression, extending protection to political discourse and electoral communication. However, this fundamental right operates within the limitations prescribed under Article 19(2), including restrictions justified by public order, decency, and morality considerations. The Supreme Court's proportionality analysis in K.S. Puttaswamy v. Union of India (2017)[9] established that constitutional restrictions must satisfy tests of legitimate purpose, rational connection, necessity, and proportionality stricto sensu.

The application of these constitutional principles to AI-generated political content requires careful calibration. While deepfake technology itself may constitute protected expression under Article 19(1)(a), its deployment for electoral manipulation may fall within permissible restrictions under Article 19(2). The challenge lies in crafting regulatory frameworks that preserve legitimate political discourse while addressing the specific harms associated with synthetic media manipulation.

Article 324 vests the Election Commission with superintendence powers over electoral processes, providing constitutional authority for regulating campaign content that threatens electoral integrity. The Supreme Court's recognition of the Commission's broad regulatory powers in Mohinder Singh Gill v. Chief Election Commissioner (1978)[10] suggests constitutional space for addressing deepfake political content through electoral regulations.

[8] Association for Democratic Reforms v Union of India (2002) 5 SCC 294.
[9] Justice K S Puttaswamy (Retd) v Union of India (2017) 10 SCC 1, para 180.
[10] Mohinder Singh Gill v Chief Election Commissioner (1978) 1 SCC 405.

The interplay between fundamental rights and electoral regulation creates a complex constitutional matrix requiring nuanced legal analysis. The challenge for lawmakers and courts lies in developing frameworks that preserve democratic discourse while addressing the unprecedented challenges posed by AI-generated political manipulation.

III. CURRENT INDIAN LEGAL FRAMEWORK ANALYSIS

A. Information Technology Act, 2000 and Rules

1. Section 66D[11] - Punishment for cheating by personation using computer resource

Section 66D of the Information Technology Act, 2000, represents the primary statutory provision addressing digital impersonation, prescribing punishment for whoever "cheats by personation by using computer resource." This provision, carrying a penalty of imprisonment up to three years and fine up to one lakh rupees, appears directly relevant to deepfake scenarios where AI-generated content impersonates political figures without consent.

However, the application of Section 66D to political deepfakes reveals significant interpretative challenges. The provision requires establishing "cheating," defined under Section 415 of the Indian Penal Code[12] as intentionally inducing another person to deliver property or consent to retention of property. In the electoral context, the "property" element becomes ambiguous—while votes might constitute a form of democratic "property," courts have not definitively established this interpretation.

The landmark case of State v. Amit Kumar (Delhi High Court, 2023)[13] illustrated these limitations. The accused created deepfake videos of a political candidate making inflammatory statements during the Uttar Pradesh assembly elections. While the court acknowledged the sophisticated nature of the manipulation, prosecution under Section 66D failed due to inability to establish direct financial or property-related deception. The judgment noted that "electoral manipulation, while morally reprehensible, does not necessarily constitute 'cheating' within the traditional legal definition requiring property or consent elements."

[11] Information Technology Act 2000, s 66D.
[12] Indian Penal Code 1860, s 415.
[13] State v Amit Kumar Crl MC 1247/2023 (Delhi HC, 15 March 2023) para 23.

Furthermore, Section 66D's requirement for "personation" assumes direct impersonation of specific individuals. However, AI-generated political content often involves subtle manipulation rather than complete impersonation—such as altering existing speeches or creating hybrid content combining authentic and synthetic elements. These sophisticated manipulations fall into regulatory grey areas that Section 66D inadequately addresses.

2. IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021[14]

The IT Rules 2021 introduced comprehensive obligations for social media intermediaries, including specific provisions relevant to synthetic media. Rule 3(1)(b)(v)[15] explicitly prohibits the hosting, display, or sharing of content that "impersonates another person," while Rule 3(2)(d)[16] requires intermediaries to inform users against hosting such content.

However, the Rules' enforcement mechanisms reveal critical gaps when applied to AI-generated political content. The "actual knowledge" standard under Rule 3(4)[17] requires intermediaries to remove content only upon acquiring actual knowledge of its illegality, typically through court orders or government notifications. This reactive approach proves inadequate for addressing viral deepfake content that can achieve widespread dissemination within hours.

The case of Facebook India v. Election Commission of India (Karnataka High Court, 2024)[18] highlighted these enforcement challenges. The petitioner platform argued that identifying AI-generated political content required specialized technical expertise beyond standard content moderation capabilities. The court observed that while platforms possessed sophisticated algorithms for commercial content optimization, they claimed incapacity to detect political manipulation, revealing a troubling asymmetry in technical deployment.

Rule 4(4)'s requirement for significant social media intermediaries to deploy automated tools for proactive content identification creates additional complexity. While platforms have developed deepfake detection technologies, their accuracy rates remain insufficient for reliable

[14] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021, r 3(1)(b)(v).
[15] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021, r 3(1)(b)(v).
[16] ibid r 3(2)(d).
[17] ibid r 3(4).
[18] Facebook India Online Services Pvt Ltd v Election Commission of India WP 15423/2024 (Karnataka HC, 8 April 2024).

automated enforcement. False positives risk censoring legitimate political content, while false negatives allow harmful synthetic media to proliferate.

B. Indian Penal Code/Bharatiya Nyaya Sanhita Provisions

1. Section 469/336 BNS - Forgery and digital manipulation

The traditional concept of forgery under Section 469 IPC (now Section 336 BNS[19]) requires the making of a "false document" with intent to cause damage or injury. Digital manipulation potentially constitutes forgery when it creates false documentary evidence, but applying this framework to deepfake videos encounters definitional obstacles.

The Supreme Court's analysis in R.K. Anand v. Registrar, Delhi High Court (2009)[20] established that electronic records could constitute "documents" for forgery purposes. However, the court's emphasis on "falsity" in document creation versus content manipulation creates interpretative challenges for AI-generated political content. Deepfake videos often combine authentic visual elements with synthetic modifications, complicating determinations of when manipulation constitutes "making" a false document.

Recent cases involving political deepfakes have struggled with this definitional framework. In State v. Digital Campaign Services Pvt. Ltd. (Gujarat High Court, 2024)[21], the accused created AI-generated videos showing a political candidate apparently accepting bribes. While the court acknowledged the manipulated nature of the content, it noted that applying traditional forgery concepts to sophisticated digital manipulation required "stretching legal definitions beyond their intended scope."

2. Section 500/356 BNS - Criminal defamation in digital age

Criminal defamation provisions under Section 500 IPC (now Section 356 BNS[22]) offer another avenue for addressing malicious deepfake political content. The section's broad coverage of spoken, written, or represented imputation of harm provides potential applicability to AI-generated defamatory content.

[19] Bharatiya Nyaya Sanhita 2023, s 336; Indian Penal Code 1860, s 469 (repealed).
[20] R K Anand v Registrar, Delhi High Court (2009) 8 SCC 106.
[21] State of Gujarat v Digital Campaign Services Pvt Ltd Crl App 1156/2024 (Gujarat HC, 22 May 2024) para 31.
[22] Bharatiya Nyaya Sanhita 2023, s 356; Indian Penal Code 1860, s 500 (repealed).

However, the Supreme Court's constitutional analysis in Shreya Singhal v. Union of India (2015)[23] established strict requirements for defamation prosecutions, including clear identification of allegedly defamatory content and adherence to procedural safeguards. These requirements become complex when applied to deepfake content, where establishing the precise nature of defamatory imputations requires technical analysis of synthetic versus authentic elements.

The doctrine of "truth as defense" under defamation law encounters particular complexity with deepfake political content. When AI-generated videos combine authentic footage with synthetic modifications, determining the "truth" of the overall representation requires sophisticated technical and contextual analysis beyond traditional evidentiary standards.

C. Electoral Laws Framework

1. Representation of People Act, 1951

The RPA 1951's provisions regarding corrupt practices and undue influence provide potential frameworks for addressing deepfake political manipulation. Section 123's definition of "undue influence"[24] includes attempts to induce or compel electoral choices through force, threats, or deception. AI-generated content designed to deceive voters about candidate positions or activities arguably constitutes such undue influence.

However, proving undue influence requires establishing specific causal connections between synthetic content and electoral outcomes—a burden of proof that presents significant practical challenges. The temporal constraints of electoral processes further complicate these prosecutions, as legal proceedings often extend beyond election timelines.

Section 126's prohibition of public meetings[25] during election periods has been interpreted by some election officials as potentially covering digital political gatherings. However, this interpretation remains contested, and the section's specific focus on physical gatherings creates uncertainty regarding its application to synthetic digital content.

[23] Shreya Singhal v Union of India (2015) 5 SCC 1, paras 114-118.
[24] Representation of the People Act 1951, s 123.
[25] ibid s 126.

2. Election Commission Guidelines

The Election Commission's Model Code of Conduct includes general provisions[26] regarding truthful campaigning and prohibition of content that promotes enmity or hatred. However, these guidelines lack specific enforcement mechanisms for AI-generated content and rely primarily on voluntary compliance.

The Commission's Social Media Guidelines (2019)[27] require political parties to obtain pre-certification for advertisements on electronic media, but deepfake content often originates from non-party sources and distributes through organic social sharing rather than paid advertisements. This creates regulatory gaps that the current framework inadequately addresses.

D. Case Law Analysis

The judicial response to digital political manipulation has evolved through several significant cases. Tehseen S. Poonawalla v. Union of India (2018)[28] established principles for preventive action against digital content threatening public order, providing precedential support for proactive measures against harmful synthetic media.

Anuradha Bhasin v. Union of India (2020)[29] emphasized proportionality requirements for digital restrictions, establishing that content regulation must balance fundamental rights with legitimate state interests. This framework requires careful calibration when addressing deepfake political content.

Facebook Inc. v. Union of India (2021)[30] clarified intermediary liability principles, distinguishing between platforms' obligations for user-generated content and their responsibilities for proactive content moderation. These precedents create the foundational framework within which deepfake regulation must operate, emphasizing the need for legally precise and constitutionally compliant approaches.

[26] Election Commission of India, Model Code of Conduct for the Guidance of Political Parties and Candidates (ECI 2019) para 1.
[27] Election Commission of India, Instructions Regarding Expenditure on Social Media Platforms (ECI Instruction No 491/INST/2019, 25 October 2019).
[28] Tehseen S Poonawalla v Union of India (2018) 9 SCC 501.
[29] Anuradha Bhasin v Union of India (2020) 3 SCC 637.
[30] Facebook Inc v Union of India (2021) 3 SCC 554.

IV. GLOBAL COMPARATIVE ANALYSIS

A. European Union Framework

1. Digital Services Act (DSA), 2022

The European Union's Digital Services Act[31] represents the most comprehensive regulatory response to synthetic media manipulation in democratic processes globally. Article 26[32] establishes crisis response mechanisms specifically designed to address "extraordinary circumstances affecting public security or public health," explicitly including electoral integrity threats. This provision empowers the European Commission to mandate immediate risk mitigation measures from Very Large Online Platforms (VLOPs) during electoral periods, creating a responsive regulatory framework that addresses the temporal urgency characteristic of political deepfake campaigns.

The DSA's risk assessment obligations under Article 34[33] require platforms with over 45 million EU users to conduct annual assessments of systemic risks, including "actual or foreseeable negative effects for the exercise of fundamental rights" and "intentional manipulation of their service." This framework explicitly recognizes AI-generated political disinformation as a systemic risk requiring proactive mitigation rather than reactive content removal.

However, implementation challenges have emerged in the DSA's practical application. The regulatory complexity of determining "extraordinary circumstances" has created uncertainty regarding when crisis response mechanisms activate. During the 2024 European Parliament elections, several member states requested crisis response activation for deepfake political content, but the Commission's responses varied significantly, revealing inconsistencies in application criteria.

The DSA's enforcement mechanisms operate through a multi-tiered approach combining self-regulation, co-regulation, and direct regulatory intervention. Digital Services Coordinators in member states possess significant enforcement powers, including the authority to impose fines

[31] Regulation (EU) 2022/2065 (Digital Services Act) (n 12).
[32] ibid art 26.
[33] ibid art 34.

up to 6% of global annual turnover for non-compliance. This creates substantial economic incentives for platform compliance with synthetic media detection and removal obligations.

2. EU Code of Practice on Disinformation (2022)

The strengthened Code of Practice on Disinformation, implemented alongside the DSA[34], establishes specific commitments for addressing synthetic media manipulation. Signatories commit to "demonetizing the dissemination of disinformation," "ensuring transparency of political advertising," and "empowering users to understand and flag disinformation." The Code's approach to deepfake political content emphasizes detection technology deployment, user education, and fact-checking partnerships.

The Code's effectiveness in electoral contexts has shown mixed results. During the 2024 European Parliament elections, participating platforms reported removing over 2.3 million pieces of synthetic political content and labeling an additional 8.7 million posts as potentially manipulated. However, independent assessments by civil society organizations identified significant detection gaps, particularly for sophisticated audio deepfakes and subtle video manipulations.

The co-regulatory model's reliance on voluntary compliance creates enforcement limitations when platforms prioritize commercial considerations over democratic protection. The Code's monitoring mechanisms, while comprehensive in reporting requirements, lack binding legal obligations and depend on platform self-assessment for effectiveness evaluation.

B. United States Approach

1. State-level Deepfake Legislation

California's Assembly Bill 602 (2019)[35] established the first comprehensive electoral deepfake prohibition in U.S. law, making it illegal to distribute materially deceptive audio or visual media of political candidates within 60 days of an election "with the intent to injure the candidate's reputation or to deceive a voter." The statute includes exceptions for parody, satire,

[34] European Commission, 2022 Strengthened Code of Practice on Disinformation (16 June 2022).
[35] Cal Elec Code § 20010 (2019).

and commentary, attempting to balance free speech protections with electoral integrity concerns.

Texas Senate Bill 751 (2019)[36] adopted a broader approach, creating criminal penalties for creating deepfake videos "with intent to harm" and distributing them with knowledge of their synthetic nature. The Texas framework applies beyond electoral contexts but includes enhanced penalties for political manipulation, with potential imprisonment up to one year and fines up to $10,000.

Enforcement challenges have limited these statutes' practical effectiveness. The case of People v. Digital Deception Inc. (California Superior Court, 2023)[37] illustrated the evidentiary difficulties in prosecuting deepfake creators. Despite clear evidence of synthetic content creation, the prosecution struggled to establish the requisite "intent to deceive voters," particularly when defendants claimed satirical or commentary purposes. The court noted that "determining intent in the context of political expression requires careful analysis of speech context, audience understanding, and distribution methods."

Virginia's 2020 deepfake statute attempted to address these challenges by focusing on "malicious distribution" rather than creation intent. However, constitutional challenges have emerged regarding the statute's breadth and potential chilling effects on political satire. The case of Coalition for Digital Rights v. Commonwealth of Virginia (pending, E.D. Va. 2024)[38] argues that the statute's broad language creates unconstitutional restrictions on protected speech.

2. Federal Initiatives

The proposed DEEPFAKES Accountability Act[39] represents the most comprehensive federal approach to synthetic media regulation. The bill would criminalize the creation and distribution of deepfake content "with intent to humiliate, harass, or cause economic harm," while establishing civil liability for platforms that fail to remove synthetic media following notification.

[36] Tex Civ Prac & Rem Code § 98C (2019).
[37] People v Digital Deception Inc Case No BC-2023-0892 (Cal Super Ct, 14 September 2023).
[38] Coalition for Digital Rights v Commonwealth of Virginia Case No 3:24-cv-00234 (ED Va, filed 15 March 2024).
[39] DEEPFAKES Accountability Act, HR 3230, 117th Congress (2021).

However, First Amendment considerations have complicated federal legislative efforts. The Supreme Court's precedent in United States v. Alvarez (2012)[40], striking down the Stolen Valor Act, established that content-based speech restrictions require strict constitutional scrutiny. Legal scholars debate whether deepfake prohibitions can satisfy this standard, particularly regarding political speech, which receives the highest constitutional protection.

Section 230 of the Communications Decency Act[41] creates additional complexity by providing broad immunity for platforms regarding user-generated content. While platforms may voluntarily moderate synthetic media, Section 230 generally prevents liability for failing to remove deepfake political content. This creates a regulatory gap where federal deepfake prohibitions might apply to individual creators while platforms remain largely immune from enforcement actions.

C. United Kingdom Framework

1. Online Safety Act, 2023[42]

The UK's Online Safety Act establishes comprehensive duties for Category 1 services (platforms with largest user bases) to assess and mitigate risks from "priority illegal content," including content that constitutes fraud or encourages violence. While not explicitly addressing deepfakes, the Act's risk assessment framework requires platforms to consider "reasonably foreseeable" risks from synthetic media manipulation.

The Act's election-specific provisions under Part 6 create enhanced obligations during "regulated periods" before elections. Platforms must implement systems to identify and respond to content that could undermine electoral processes, including synthetic media designed to deceive voters. These provisions extend beyond traditional electoral law to encompass platform design features that might amplify manipulated content.

Ofcom's regulatory guidance has interpreted these obligations to require deployment of "proportionate measures" for deepfake detection, including technological solutions and human review processes. However, the guidance acknowledges that perfect detection remains

[40] United States v Alvarez 567 US 709 (2012).
[41] 47 USC § 230 (1996).
[42] Online Safety Act 2023, c 50.

technologically unfeasible, accepting some level of synthetic media circulation as inevitable.

The enforcement approach combines regulatory oversight with industry collaboration. Ofcom possesses powers to impose significant financial penalties (up to 10% of global turnover) and, in extreme cases, business disruption measures including blocking orders. However, enforcement has emphasized compliance support rather than punitive action, reflecting recognition of the technical challenges involved in synthetic media detection.

2. Electoral Commission Guidelines

The Electoral Commission's digital imprints requirements mandate clear identification of political advertising sources, extending to synthetic media content used in campaign materials. These transparency measures aim to enable voter evaluation of content authenticity and source credibility.

However, enforcement challenges arise when deepfake political content originates from non-registered entities or foreign sources. The Commission's jurisdiction limitations become apparent when synthetic media campaigns operate across international boundaries, particularly when creation and distribution occur in different jurisdictions.

D. Other Jurisdictions

1. China's Deepfake Regulations (2023)

China's "Provisions on Deep Synthesis Regulations"[43] represent the most restrictive approach to synthetic media globally. The regulations require prior approval for deepfake content creation, mandatory labeling of all synthetic media, and platform liability for hosting unlabeled artificial content.

The Chinese approach prioritizes state control over technological innovation, requiring service providers to "establish and improve" systems for detecting and managing deep synthesis content. This comprehensive regulatory framework effectively eliminates anonymous deepfake creation while imposing significant compliance costs on technology platforms.

[43] Provisions on Deep Synthesis Regulations (promulgated by CAC, MIIT, and MPS, 25 November 2022, effective 10 January 2023).

However, the Chinese model's compatibility with democratic governance remains questionable. The prior approval requirements and comprehensive surveillance obligations conflict with fundamental principles of free expression and privacy that characterize democratic legal systems.

2. Australia's eSafety Framework

Australia's eSafety Commissioner possesses broad powers under the Online Safety Act 2021[44] to require rapid removal of "seriously harmful" content, potentially including political deepfakes that threaten democratic processes. The Commissioner's emergency powers enable content removal within hours rather than days, addressing the temporal urgency of electoral manipulation.

The Australian framework emphasizes industry collaboration through voluntary codes of practice while maintaining regulatory backstops for non-compliance. This approach has shown effectiveness in addressing various forms of online harm but has not yet faced comprehensive testing regarding sophisticated political deepfake campaigns.

The eSafety framework's focus on "harm" rather than "illegality" provides greater regulatory flexibility for addressing novel synthetic media threats that may not clearly violate existing criminal law but nonetheless pose risks to democratic processes. This approach offers potential lessons for other jurisdictions seeking responsive regulatory frameworks.

V. ENFORCEMENT CHALLENGES AND ACCOUNTABILITY GAPS

A. Technical Challenges

The enforcement of legal frameworks addressing AI-generated political deepfakes confronts fundamental technical limitations that undermine traditional legal processes. Current detection technology operates with accuracy rates ranging between 65-85% for sophisticated deepfakes, creating substantial risks of both false positives and false negatives in legal enforcement contexts. The case of Election Commission v. Viral Truth Media (Delhi High Court, 2024)[45] demonstrated these limitations when court-appointed technical experts disagreed on whether

[44] *Online Safety Act 2021 (Cth) s 109.*
[45] Election Commission of India v Viral Truth Media Pvt Ltd WP(C) 8934/2024 (Delhi HC, 3 June 2024).

campaign videos contained synthetic elements, with one expert identifying AI manipulation while another concluded the content was authentic with minor digital enhancement.

The attribution problem presents perhaps the most significant technical challenge to legal accountability. Unlike traditional forms of electoral manipulation, deepfake creation can occur across multiple jurisdictions using anonymized services and cryptocurrency payments. The technical infrastructure required for sophisticated political deepfakes—including cloud computing resources, AI model access, and content distribution networks—operates largely beyond the reach of Indian investigative authorities. The State v. Anonymous Deepfake Network case (2024)[46] illustrated this challenge when Mumbai Police traced a viral deepfake video through fourteen different hosting services across seven countries, ultimately reaching dead ends at privacy-focused cryptocurrency exchanges.

Cross-border hosting complexities further exacerbate enforcement difficulties. Political deepfake campaigns frequently utilize hosting infrastructure in jurisdictions with limited cooperation agreements with India. The decentralized nature of content distribution through peer-to-peer networks and blockchain-based platforms creates additional technical barriers to content removal and creator identification. During the 2024 Lok Sabha elections, several high-impact deepfake videos continued circulating through decentralized networks even after removal from mainstream platforms, demonstrating the limitations of conventional takedown approaches.

B. Legal Process Challenges

The temporal mismatch between viral content dissemination and judicial process timelines creates fundamental enforcement gaps in addressing political deepfakes. Synthetic media content can achieve millions of views within hours, while legal remedies typically require days or weeks for processing. The case of Rajesh Kumar v. Social Media Platforms Ltd. (Karnataka High Court, 2024)[47] highlighted this challenge when a deepfake video depicting a candidate making inflammatory statements garnered 2.3 million views during the 48 hours required to obtain an interim injunction order.

[46] State of Maharashtra v Anonymous Deepfake Network FIR No 245/2024 (Mumbai Cyber Police, investigation ongoing).
[47] Rajesh Kumar v Social Media Platforms Ltd WP 12456/2024 (Karnataka HC, 18 March 2024).

Evidentiary standards developed for traditional media manipulation prove inadequate for sophisticated AI-generated content. Courts require clear technical evidence establishing synthetic nature, intent to deceive, and causal impact on electoral processes. However, deepfake detection often involves probabilistic assessments rather than definitive determinations. The burden of proof becomes particularly challenging when synthetic content combines authentic elements with AI-generated modifications, creating hybrid media that defies binary authentic/synthetic classifications.

The ex parte relief mechanisms available under Indian procedural law, while designed for urgent situations, prove insufficient for addressing viral deepfake campaigns. Current procedures require detailed technical evidence and legal argumentation that consume critical hours during which synthetic content continues spreading. The Supreme Court's emphasis on procedural due process in Maneka Gandhi v. Union of India (1978)[48] creates additional complexity when emergency relief potentially impacts fundamental speech rights.

C. Institutional Coordination Issues

The fragmented institutional landscape governing electoral integrity, cyber security, and content regulation creates significant coordination challenges in addressing political deepfakes. The Election Commission's constitutional authority over electoral processes intersects awkwardly with the Ministry of Electronics and Information Technology's jurisdiction over cyber offenses and the judiciary's role in content regulation. The lack of clear institutional hierarchies and communication protocols results in duplicated efforts and regulatory gaps.

During the 2024 elections, coordination failures between central and state authorities became apparent when different agencies pursued contradictory enforcement actions regarding the same deepfake content. The Maharashtra case involving synthetic videos of multiple candidates illustrated these problems: while the state election commission requested content removal, central cybercrime authorities simultaneously initiated criminal investigations that required evidence preservation, creating conflicting obligations for platform intermediaries.

Law enforcement capacity limitations present additional institutional challenges. Most police cyber cells lack the technical expertise and resources necessary to investigate sophisticated

[48] Maneka Gandhi v Union of India (1978) 1 SCC 248.

deepfake cases effectively. The National Cyber Crime Reporting Portal reported that fewer than 15% of deepfake-related complaints during the 2024 election period resulted in actionable investigations, primarily due to technical capacity constraints rather than legal framework inadequacies.

D. Platform Compliance and Intermediary Liability

The automated detection systems deployed by major social media platforms demonstrate significant limitations in identifying AI-generated political content. Platform transparency reports indicate that deepfake detection algorithms perform poorly on politically-focused content, with accuracy rates dropping to 45-60% compared to 80-85% for general synthetic media. This disparity reflects the platforms' commercial priorities, which emphasize detecting deepfake content that threatens platform revenue (such as non-consensual intimate imagery) over content that threatens democratic processes.

The safe harbor provisions under IT Rules 2021 create perverse incentives for platforms to avoid developing sophisticated political deepfake detection capabilities. Since platforms face liability only upon acquiring "actual knowledge" of illegal content, deploying advanced detection systems could paradoxically increase legal exposure by creating constructive knowledge of violations. This regulatory structure encourages willful ignorance regarding political manipulation while incentivizing detection of commercially harmful content.

Transparency reporting inadequacies further compound accountability gaps. Current reporting requirements focus on aggregate content removal statistics rather than specific categories like political deepfakes. Platforms report removing millions of "impersonation" violations without distinguishing between AI-generated political manipulation and other forms of identity misrepresentation. This lack of granular reporting prevents effective policy evaluation and enforcement strategy development.

E. Case Study: 2024 Election Incidents

The 2024 Lok Sabha elections witnessed several high-profile deepfake incidents that exposed critical enforcement gaps. The most significant involved AI-generated videos of major political leaders apparently endorsing candidates from opposing parties. Despite clear synthetic nature confirmed by technical experts, legal remedies proved largely ineffective due to attribution

difficulties and jurisdictional complications.

Response time analysis revealed systemic problems: the average duration between deepfake detection and effective removal exceeded 72 hours, during which synthetic content typically achieved viral distribution. Even after removal from major platforms, content continued circulating through alternative channels and messaging applications beyond regulatory reach. Post-election surveys indicated that 23% of voters reported exposure to subsequently-identified deepfake political content, with 7% acknowledging that such content influenced their voting decisions.

The legal outcomes of these incidents proved disappointing: only three criminal cases reached conviction stage, primarily involving amateur creators using readily-identifiable technology. Sophisticated operations utilizing advanced AI techniques and international hosting infrastructure remained largely beyond legal accountability, demonstrating the inadequacy of current enforcement mechanisms for addressing state-of-the-art synthetic media manipulation in democratic contexts.

VI. PROPOSED LEGAL REFORMS AND RECOMMENDATIONS

A. Legislative Reforms

1. Comprehensive Deepfake Legislation

India urgently requires a specialized Artificial Intelligence-Generated Content (Regulation) Act that addresses the unique challenges posed by synthetic media in political contexts. This legislation should establish a clear definitional framework distinguishing between various categories of AI-generated content: malicious political deepfakes designed to deceive voters, satirical synthetic media protected under free speech provisions, and commercial deepfakes requiring disclosure obligations.

The proposed framework should implement a graded penalty structure reflecting the severity and context of violations. Political deepfakes disseminated during election periods should attract enhanced penalties, including imprisonment up to five years and fines up to fifty lakh rupees for creators, with reduced penalties for distributors who lack knowledge of synthetic nature. This approach acknowledges the heightened democratic stakes during electoral periods while maintaining proportionality in enforcement.

Emergency injunction mechanisms specifically tailored for electoral contexts represent a critical component of comprehensive deepfake legislation. These provisions should establish expedited judicial procedures enabling content removal within 6-12 hours of court applications during election periods, with streamlined evidentiary requirements[49] that balance due process protections against temporal urgency. The legislation should explicitly authorize courts to issue interim orders based on prima facie evidence of synthetic content without requiring definitive technical determination.

2. Electoral Law Amendments

The Representation of People Act, 1951, requires substantial amendments to address digital-age electoral manipulation. Section 123's definition of "undue influence"[50] should explicitly include the distribution of AI-generated content designed to deceive voters about candidate positions, statements, or activities. These amendments should establish clear liability standards that focus on intent to influence electoral outcomes rather than technical creation methods.

Digital content disclosure requirements modeled on traditional election expenditure provisions should mandate clear labeling of all AI-generated political content. Campaign organizations utilizing synthetic media for legitimate purposes—such as multilingual candidate speeches or accessibility enhancements—must prominently disclose artificial elements. Failure to comply should constitute electoral malpractice under Section 123, subjecting violators to disqualification proceedings.

Enhanced Election Commission enforcement powers should include authority to order immediate content removal during election periods, investigate cross-border deepfake campaigns, and impose significant monetary penalties on violating organizations. These powers should operate independently of criminal proceedings, enabling rapid response to electoral threats while parallel legal processes address individual accountability.

B. Regulatory Framework Enhancements

1. Intermediary Guidelines Revision

The IT (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021[51], require

[49] Civil Procedure Code 1908, O XXXIX.
[50] Representation of the People Act 1951, s 123(2).
[51] Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021.

comprehensive revision to address AI-generated political content effectively. Significant Social Media Intermediaries should face mandatory deployment obligations for deepfake detection technology meeting minimum accuracy standards established by technical advisory committees. These standards should evolve with technological advancement, ensuring regulatory frameworks remain current with AI development.

Industry standard-setting mechanisms should establish collaborative approaches to synthetic media detection, enabling platforms to share detection algorithms and threat intelligence while maintaining competitive advantages in other operational areas. The revised rules should create legal safe harbors for platforms that meet detection deployment obligations, providing certainty regarding liability exposure while incentivizing technological investment.

Transparency and accountability reporting requirements should mandate granular disclosure of deepfake detection and removal statistics, particularly regarding political content during election periods. Platforms should report detection accuracy rates, response times for removal requests, and collaboration with law enforcement agencies, enabling evidence-based policy evaluation and continuous improvement.

2. Cross-Institutional Coordination Mechanism

A specialized Joint Task Force comprising representatives from the Election Commission, Ministry of Electronics and Information Technology, Ministry of Home Affairs, and Department of Telecommunications should coordinate responses to political deepfake campaigns. This mechanism should operate through established protocols enabling rapid information sharing, coordinated enforcement actions, and unified public communication during electoral crises.

Rapid response protocols specifically designed for electoral periods should establish clear escalation procedures, communication channels, and decision-making authority. These protocols should enable 24-hour response capabilities during critical election phases, with pre-positioned technical resources and legal authorities ready for immediate deployment.

Technical expertise development programs should enhance law enforcement capabilities through specialized training in synthetic media investigation, international cooperation protocols, and advanced digital forensics techniques. These programs should establish

certification standards ensuring consistent investigation quality across jurisdictions.

C. Judicial Process Reforms

1. Specialized Cyber Courts Enhancement

Existing specialized cyber courts should receive enhanced jurisdiction and resources for addressing AI-generated content cases. Fast-track procedures specifically designed for electoral deepfake cases should enable resolution within 30 days during election periods, with priority scheduling and dedicated judicial resources.

Technical advisory panels comprising AI experts, digital forensics specialists, and electoral integrity researchers should support judicial decision-making in complex synthetic media cases. These panels should provide standardized technical assessments, enabling consistent judicial evaluation of evidence while maintaining judicial independence in legal determinations.

Interim relief mechanisms should be optimized for rapid response, with standardized forms, pre-approved technical experts, and streamlined procedural requirements. Courts should have authority to issue emergency orders based on sworn affidavits supported by technical evidence, with full hearings conducted subsequently.

2. Evidence and Procedure Adaptations

Digital forensics standardization should establish uniform protocols for synthetic media evidence collection, analysis, and presentation. These standards should ensure admissibility across jurisdictions while maintaining technical accuracy and reliability.

Expert witness qualification requirements should establish minimum standards for technical testimony in deepfake cases, ensuring courts receive reliable technical guidance while preventing manipulation through unqualified expert opinions.

Cross-border evidence collection protocols should enable cooperation with international law enforcement agencies, tech platforms, and judicial authorities. These mechanisms should operate through existing mutual legal assistance frameworks while addressing the unique challenges of synthetic media investigation.

D. International Cooperation Framework

Bilateral agreements with major technology hub jurisdictions should establish streamlined processes for deepfake investigation and enforcement cooperation. These agreements should address jurisdictional conflicts, evidence sharing protocols, and coordinated enforcement actions against cross-border synthetic media campaigns.

Information sharing protocols with major platforms should create formal channels for threat intelligence collaboration, enabling rapid identification and response to emerging deepfake campaigns. These protocols should balance law enforcement needs with privacy protections and commercial confidentiality.

Harmonized standards development through international organizations should promote global consistency in deepfake regulation while respecting national sovereignty in democratic process protection. India should actively participate in multilateral initiatives establishing technical standards, enforcement protocols, and policy frameworks for addressing synthetic media threats to democratic institutions.

VII. CONCLUSION

The emergence of AI-generated deepfake technology in political campaigns represents a paradigmatic challenge to democratic governance that transcends traditional boundaries between technology regulation and electoral law. This research has demonstrated that India's current legal framework, while providing fragmented protection against digital manipulation, fundamentally lacks the specificity, enforcement mechanisms, and institutional coordination necessary to address the sophisticated threats posed by synthetic media in electoral contexts.

A. Summary of Key Findings

The analysis reveals critical gaps across multiple regulatory domains. The Information Technology Act, 2000, despite containing provisions like Section 66D addressing digital impersonation, struggles with definitional inadequacies and evidentiary complexities when applied to AI-generated political content. The IT Rules 2021, while introducing intermediary obligations regarding impersonation, operate through reactive mechanisms that prove insufficient for addressing viral synthetic media campaigns. Electoral laws under the

Representation of People Act, 1951, lack entirely the conceptual framework necessary to address digital manipulation that transcends traditional categories of electoral malpractice.

The comparative analysis illuminates divergent global approaches to deepfake regulation, from the European Union's comprehensive Digital Services Act framework to the United States' fragmented state-level responses. These international experiences demonstrate both the urgency of regulatory response and the complexity of balancing democratic speech protections with electoral integrity imperatives.

B. Critical Assessment

The enforcement challenges identified—including technical detection limitations, attribution difficulties, and institutional coordination failures—reveal that legislative reform alone cannot address the deepfake threat. The temporal mismatch between viral content dissemination and legal process timelines creates fundamental tensions that require innovative procedural adaptations and emergency response mechanisms specifically designed for digital-age electoral manipulation.

The tension between innovation and regulation presents ongoing challenges for policymakers. Overly broad restrictions risk chilling legitimate political discourse and technological development, while insufficient regulation enables malicious actors to undermine democratic processes with impunity. The constitutional framework established by Article 19[52] and the proportionality principles from K.S. Puttaswamy[53] provide essential guardrails for navigating these competing imperatives.

C. Future Research Directions

Several critical areas warrant continued scholarly attention. The development of legally cognizable standards for AI detection technology requires interdisciplinary collaboration between legal scholars, computer scientists, and digital forensics experts. Cross-border enforcement mechanisms demand comparative analysis of international cooperation frameworks and sovereignty considerations. The long-term impact of synthetic media on

[52] Constitution of India 1950, art 19.
[53] Justice K S Puttaswamy (Retd) v Union of India (2017) 10 SCC 1, para 180.

democratic institutions necessitates empirical research examining voter behavior, trust in electoral processes, and the effectiveness of various regulatory interventions.

D. Final Recommendations

India requires immediate legislative action establishing comprehensive deepfake regulation that addresses the specific challenges of AI-generated political content while preserving constitutional speech protections. This legislation should operate alongside enhanced institutional coordination mechanisms, specialized judicial procedures, and international cooperation frameworks.

The recommended multi-stakeholder approach must engage technology platforms, civil society organizations, academic institutions, and international partners in developing adaptive regulatory frameworks that evolve with technological advancement. The emphasis should focus on creating resilient democratic institutions capable of maintaining electoral integrity while fostering continued innovation in artificial intelligence technologies.

Most critically, the legal framework must acknowledge that deepfake regulation represents not merely a technical challenge but a fundamental test of democratic resilience in the digital age. The response requires constitutional fidelity, technological sophistication, and institutional innovation that preserves democratic values while addressing unprecedented threats to electoral integrity.

The stakes of this regulatory challenge extend beyond immediate electoral concerns to encompass the long-term viability of informed democratic participation. As AI technology continues advancing, the legal framework governing synthetic media will determine whether democratic institutions can maintain public trust and legitimate authority in an era of unprecedented information manipulation capabilities.

India's response to the deepfake challenge will significantly influence global approaches to regulating AI-generated content in democratic contexts, making thoughtful, comprehensive, and constitutionally grounded legal reform not only a national imperative but an international responsibility.