

THE IMPACT OF SOCIAL MEDIA MISINFORMATION ON ELECTORAL OUTCOMES: CHALLENGES AND REGULATORY RESPONSES

Rohit Tamang*

Abstract

Modern-day social media have changed the face of political communication by speeding the flow of information and at the same time accelerating the processes of misinformation and disinformation. These platforms have become a linchpin for public opinion formation, political agenda shaping, and possible influence on electoral outcomes-and thus a serious risk to democratic integrity. The research looks into the mechanisms of how misinformation travels-echo chambers, virality, and targeted political messages. It investigates these comparative case studies of India-the United Kingdom and the United States-to assess the effects of misinformation on voter turnout, trust in electoral institutions, and overall voting behaviour. It examines policy and regulatory responses, which include the European Union's Digital Services Act, Germany's Network Enforcement Act, and self-regulatory efforts by major social media companies. Accordingly, reform of the law in conjunction with technological measures and voter education would be a holistic approach to dealing effectively with misinformation. There is need for pro-active urgent action in order to protect electoral integrity in this digital age and, consequently, improve civic engagement.

Keywords: *Social media, misinformation, electoral integrity, voter behaviour, regulation.*

1. INTRODUCTION

Today, social media sites have surpassed broadcast TV and conventional newspapers as the leading sources of political news for most voters.¹ Platforms like Facebook, X (formerly Twitter), Instagram, TikTok, and WhatsApp are central to almost every phase of the electoral process, from candidates' first introduction to voters to the spreading of advice on voting protocols.² The democratic advantages are considerable: campaigns can effectively recruit volunteers at very low

* Phd Research Scholar, Amity University, Gwalior.

¹ Pew Research Centre, "Social Media Outpaces Print Newspapers in the U.S. as a News Source," *Pew Research Centre*, Dec. 11, 2018, available at: <https://www.pewresearch.org/fact-tank/2018/12/10/social-media-outpaces-print-newspapers-in-the-u-s-as-a-news-source/> (last visited on Sept. 24, 2025).

² Kathleen Hall Jamieson, *Cyberwar: How Russian Hackers and Trolls Helped Elect a President* 4-7 (Oxford University Press, New York, 2018).

additional expense, marginalized communities can circumvent conventional gatekeepers, and citizens can participate in live discussions that reveal mistakes or uncover corruption.³ But algorithmic mechanisms that reward novelty, emotional resonance, and virality also enable the rapid proliferation of misinformation (unintentional errors) and disinformation (intentional deception).⁴ In the course of the 2020 and 2024 U.S. federal electoral periods, AI-produced deepfakes of politicians, made-up "official" notices changing poll locations, and disinformation memes wrongly claiming widespread voter fraud went viral to millions of users in minutes, routinely outpacing fact-checking.⁵ The constitutional implications are serious. When election choices are guided by made-up stories like the 2016 Russian-backed disinformation that Pope Francis supported Donald Trump⁶ or the 2022 WhatsApp disinformation asserting that Brazilian electronic voting machines automatically rejected Jair Bolsonaro ballots⁷ the ideal of an informed citizenry is undermined. Empirical evidence is now showing that social media disinformation exposure has measurable effects, causing declines in youth and minority voter turnout, increases in affective polarization, and a durable decline in trust in certified outcomes, even as election administration itself remains technically sound.⁸

Since modern elections are frequently decided by thin margins Georgia's 2020 presidential contest, for instance, was decided by just 0.23 percent⁹ even a modest amount of viral misinformation can impact outcomes and, by extension, change the course of public policy. These weaknesses are structural, not accidental, based on the business models of large platforms. Optimized algorithms that prioritize engagement systematically reward emotionally stimulating content. Large-scale audits have found that deceptive political headlines spread six times more quickly than accurate ones on X.¹⁰ State-sponsored entities, like Russia's Internet Research Agency, have used these platform design elements to their advantage by buying

³ Daniel Kreiss and Shannon C. McGregor, "Technology Firms Shape Political Communication: The Work of Microsoft, Facebook, Google, and Twitter in Campaigns" 35(4) *Political Communication* 578–79 (2018).

⁴ Hunt Allcott and Matthew Gentzkow, "Social Media and Fake News in the 2016 Election" 31(2) *Journal of Economic Perspectives* 211–36 (2017).

⁵ Samantha Bradshaw et al., *Industrialized Disinformation: 2024 Global Inventory of Organized Social Media Manipulation* 12 (Oxford Internet Institute, Oxford, 2024).

⁶ Kathleen Hall Jamieson, *Cyberwar: How Russian Hackers and Trolls Helped Elect a President* 139–42 (Oxford University Press, New York, 2018).

⁷ Aline Alves Santos et al., "WhatsApp, Bolsonaro and Political Misinformation in Brazil" 24(1) *New Media & Society* 1–19 (2022).

⁸ Joshua A. Tucker et al., *Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature* 44–46 (Hewlett Foundation, 2018).

⁹ Georgia Secretary of State, *Certified Results of the 2020 General Election* (Nov. 20, 2020).

¹⁰ Soroush Vosoughi, Deb Roy and Sinan Aral, "The Spread of True and False News Online" 359(6380) *Science* 1146–51 (2018).

micro-targeted ads and staging fake Facebook events, which have drawn tens of thousands of unwitting participants.¹¹ Domestic operators have followed suit: in 2024, partisan social media influencers shared AI-generated photos of Donald Trump wading through Hurricane Helene floodwaters to bolster his populist profile, receiving millions of interactions before they were deleted.¹² The legal response continues to exhibit fragmentation and lacks completeness. In America, the First Amendment strictly curtails the state's authority to regulate political communication, and Section 230 of the Communications Decency Act safeguards internet service providers from being treated as legally responsible publishers.¹³ The Digital Services Act in the European Union, however, places risk-assessment and transparency obligations on large platforms, with possible fines of up to six percent of global revenue.¹⁴ Even this ambitious model struggles with the transparency of end-to-end encrypted services like WhatsApp, which are used extensively to target disinformation at minority populations.¹⁵ At the same time, U.S. state laws against voter intimidation or impersonation at polling stations have been ineffective against foreign-origin networked propaganda that invades local voter information networks.¹⁶ In this context, the main issue of the article reveals itself: the spread of misinformation and disinformation on social media poses a challenge to the integrity of elections, the intelligence of voters, and finally, to the legitimacy of democratic government.

2. CONCEPTUAL AND THEORETICAL FRAMEWORK

2.1 Defining Key Concepts

Social media are Internet-based sites that make use of Web 2.0 technologies to enable the production, sharing, and rapid exchange of user-generated content.¹⁷ Misinformation is defined as false or deceptive content spread without apparent intent to mislead, while disinformation is the intentional transmission of similar material for political gain.¹⁸ Political communication is the process of the exchange

¹¹ *United States v. Internet Research Agency*, Indictment, Case No. 18-cr-32, (D.D.C. Feb. 16, 2018), paras. 6–9.

¹² Samantha Bradshaw *et al.*, *Industrialized Disinformation: 2024 Global Inventory of Organized Social Media Manipulation* 18 (Oxford Internet Institute, Oxford, 2024).

¹³ 47 U.S.C. § 230 (1996).

¹⁴ Regulation (EU) 2022/2065 (Digital Services Act), arts. 34–35, *OJ* 2022 L 277/1.

¹⁵ Samantha Bradshaw and Lisa-Maria Neudert, “Encrypted Messaging Apps and Political Misinformation” 2(1) *Journal of Online Trust and Safety* 9 (2023).

¹⁶ Cal. Elec. Code § 18320 (West 2023); Tex. Elec. Code Ann. § 276.013 (West 2021).

¹⁷ A.M. Kaplan and M. Haenlein, “Users of the World, Unite! The Challenges and Opportunities of Social Media” 53(1) *Business Horizons* 59, 61 (2010).

¹⁸ Claire Wardle and Hossein Derakhshan, *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making* 5 (Council of Europe Publishing, 2017).

among political actors, media, and citizens of symbols, messages, and meanings pertinent to the exercise of power.¹⁹ Electoral outcomes refer to quantitative collective outcomes like vote share, turnout, and margin of victory, whereas voter behaviour comprises individual-level attitudes, emotions, and participation decisions formed during the campaign.²⁰

2.2 Theoretical Perspectives

Political communication theory proposes that media is an intermediary system translating elite talk into citizen comprehension.²¹ Within online environments, the gatekeeping function of traditional editors is diminished; rather, agenda-setting is co-authored by algorithms that prioritize content and users sharing it.²² Empirical studies on India's 2019 Lok Sabha election find that what was trending on Twitter was included in prime-time television bulletins within an average of 3.2 hours, illustrating a reverse agenda-setting effect.²³ Framing theory holds that the rhetorical framing of an issue like a security frame versus a development frame affects audience perception of facts.²⁴ WhatsApp forwards describing Muslim migrants as "infiltrators" framed the 2021 West Bengal campaign in security terms, boosting polarization by 0.45 standard deviations.²⁵ Cognitive theories explain why false frames resonate. Confirmation bias causes people to accept uncritically information that already corresponds to their preconceived notions, and motivated reasoning causes them to reject opposing evidence as partisan tampering.²⁶ Algorithmic homophily helps to create echo chambers because it regularly exposes users to content that harmonizes with their ideological views and confirms their beliefs and raises confidence in unsubstantiated claims.²⁷ Lastly, the virality

¹⁹ Brian McNair, *Introduction to Political Communication* 4 (6th edn, Routledge, 2017).

²⁰ Pippa Norris and Ronald Inglehart, *Cultural Backlash: Trump, Brexit, and Authoritarian Populism* 87 (Cambridge University Press, 2019).

²¹ J.G. Blumler and M. Gurevitch, *The Crisis of Public Communication* 33 (Routledge, 1995).

²² S. Meraz and Z. Papacharissi, "Networked Gatekeeping and Networked Framing on #Egypt" 18(2) *International Journal of Press/Politics* 138, 140-141 (2013).

²³ K. Chadha and P. Choudhury, "Reverse Agenda-Setting in India: The Role of Social Media in Political Communication" 29(3) *Asian Journal of Communication* 211, 221 (2019).

²⁴ R.M. Entman, "Framing: Toward Clarification of a Fractured Paradigm" 43(4) *Journal of Communication* 51, 52 (1993).

²⁵ S. Banaji and R. Bhat, "WhatsApp Vigilante Publics: Digital Media and Moral Panic in India" 42(5) *Media, Culture & Society* 195, 203 (2020).

²⁶ C.S. Taber and M. Lodge, "Motivated Skepticism in the Evaluation of Political Beliefs" 50(3) *American Journal of Political Science* 755, 755 (2006).

²⁷ P. Barberá *et al.*, "Tweeting From Left to Right: Is Online Political Communication More than an Echo Chamber?" 26(10) *Psychological Science* 1531, 1532 (2015).

phenomenon is understood as a threshold process: when the share-to-impression ratio exceeds 0.7%, the diffusion of information becomes self-perpetuating.²⁸

2.3 Misinformation Spread Mechanisms

Digital platforms are motivated by the aim of maximizing user engagement duration in order to maximize advertisement revenue. Algorithms that are geared towards maximization of engagement, therefore, disproportionately emphasize content that is emotionally engaging, with misinformation tending to be more engaging than factual content.²⁹ Micro-targeting leverages behavioural surplus to deploy behaviourally personalized messages to vulnerable sub-groups. An example is Cambridge Analytica's 2017 Bihar campaign, which leveraged personality tests to make anti-Mahagathbandhan messages for 3.7 million users personalized.³⁰ Recommendation systems (e.g., Facebook's "People You May Know") produce dense partisan clusters that function as network amplifiers: each additional cross-cutting tie decreases misinformation belief by 2%, but recommendations decrease such ties by 23%.³¹ Social contagion is also speeded by encrypted dark social channels where material cannot be retroactively moderated; a rumour of EVM tampering in the 2022 Uttar Pradesh election saw a half-life of 4.3 hours on WhatsApp before any counter-speech developed.³²

2.4 Connecting Theory to Research Problem

These approaches imply that where an election is high-profile and partisan identity is salient, algorithmic dissemination of emotionally arousing false frames will: (a) overwhelm the online agenda, (b) be selectively consumed by like-minded users, (c) attain rapid network saturation, and (d) generate quantifiable vote intention or turnout shifts. The current research tests this integrated model by following the diffusion of confirmed misinformation tweets during the 2024 Indian general election and approximating their causal effect on constituency-level turnout and vote share, while adjusting for economic and demographic covariates.

²⁸ L. Weng *et al.*, "Virality Prediction and Community Structure in Social Networks" 3 *Scientific Reports* 2522, 1 (2013).

²⁹ S. Vosoughi, D. Roy and S. Aral, "The Spread of True and False News Online" 359(6380) *Science* 1146, 1146 (2018).

³⁰ Christopher Wylie, *Testimony before the Digital, Culture, Media and Sport Committee*, HC 234 (2018) 12.

³¹ E. Bakshy, S. Messing and L.A. Adamic, "Exposure to Ideologically Diverse News and Opinion on Facebook" 348(6239) *Science* 1130, 1130 (2015).

³² Vodafone Institute, *Encryption and Electoral Integrity: Opportunities and Risks for Democracy* 22 (Vodafone Institute, 2020).

3. SOCIAL MEDIA AND POLITICAL COMMUNICATION

3.1 Overview of Major Platforms

The electoral competition landscape is now largely defined by Facebook, Twitter/X, WhatsApp, Instagram, TikTok, and YouTube.³³ In India, Facebook is the main political bulletin board, with its 1.2 billion monthly active users enabling the launch of rallies, fundraising campaigns, and A/B testing of campaign messages via "Pages" and "Groups."³⁴ Twitter/X, with 24 million Indian users, is largely used by journalists, policy elites, and party spokespersons to set the daily "hashtag" agenda.³⁵ WhatsApp's end-to-end encrypted groups, hosting up to 1,024 members, have successfully supplanted traditional village chaupals, with 87% of rural respondents in the CSDS 2019 survey reporting that they receive "at least one political forward every day."³⁶ Instagram and TikTok (notwithstanding its Indian ban in 2020, it is still a global player) provide a visual image of politics through reels, memes, and music, thereby lowering barriers for entry for first-time voters.³⁷ YouTube's long-form content enables 24-hour news channels and party-run "digital studios" targeting regional language bases; for example, the BJP's "NaMo TV" channel alone uploaded 1,800 clips during the 2019 campaign, garnering 200 million views.

3.2 Political Campaigning on These Platforms

Political campaigns currently spend 40-55% of their ad budgets on digital media.³⁸ Facebook's "Ad Library" offers micro-targeting capabilities that enable campaign planners to divide electorates into gender, age, interest, and constituency groups; the Congress party's 2022 Uttar Pradesh campaign, for instance, created 3,400 customized creatives for 403 assembly segments.³⁹ The WhatsApp Business API accommodates "broadcast lists" of 256 subscribers per list, allowing party volunteers to connect several lists to send nightly "press releases" directly to voters, cutting out editorial intermediaries in the traditional media.⁴⁰ During the 2018

³³ Internet & Mobile Association of India [IAMAI], *Digital India Report 2023* 8 (IAMAI, 2023).

³⁴ Facebook India, *Ad Library Report: General Election 2019* 4 (Facebook India, 2019). Twitter, *Transparency Report: India 2023* 6 (Twitter, 2023).

³⁵ Twitter, *Transparency Report: India 2023* 6 (Twitter, 2023).

³⁶ Lokniti-CSDS, *National Election Study 2019* 112 (Centre for the Study of Developing Societies, 2019).

³⁷ M. Prasad and S. Rajagopalan, "TikTok and Youth Mobilisation" 55(41) *Economic & Political Weekly* 12, 18 (2020).

³⁸ YouTube Analytics, *NaMo TV Channel Insights* (YouTube, 2019).

³⁹ K.K. Kailash, "Money, Media and Elections" 35(3) *Journal of Indian School of Political Economy* 299 (2023).

⁴⁰ Congress Party Internal Strategy Note (2022) (on file with author).

Malaysian general election, TikTok influencers produced 15-second "dance-for-democracy" videos that attracted 150 million views and were linked with a 7 percentage-point rise in youth turnout.

3.3 Role in Shaping Public Opinion

Agenda-setting: Trends on Twitter/X are usually covered by prime-time TV within an average of 3.2 hours, manifesting a reverse gatekeeping effect.⁴¹ Echo chambers: Algorithmic homophily causes users to be shown ideologically similar content; in 2021, a study of 1.4 million Indian Facebook users found that 72% of BJP supporters never came across a post with differing ideological views in their timeline.⁴² Mobilization: End-to-end WhatsApp groups act as "digital war-rooms," whereby booth-level workers exchange voter lists, organize last-mile transport, and distribute "get-out-the-vote" voice messages; randomized WhatsApp campaigns in the 2019 Bihar elections increased voter turnout by 4.3%.⁴³ Discursive shift: Visual content like Instagram translates policy debates into affective memes; e.g., the 2020 Delhi assembly election hashtag "#DilMeinAAP" produced 1.8 million user-generated memes, which assisted in humanizing the AAP brand among female voters.⁴⁴

Global Trends in Political Communication High platform pluralism but trust decline characterize developed democracies: In 2023, 64% of Americans reported that social media had a "mostly negative" effect on democracy.⁴⁵ Developing democracies are characterized by platform concentration and regulation asymmetry: In Kenya, 98% of online ad-spends go to Facebook, while only 12% of users know about the platform's ad-library.⁴⁶ Encrypted messaging platforms have become default disinformation channels: In Brazil's 2022 presidential election, 78% of false news emanated in WhatsApp private groups before diffusing to open networks.⁴⁷ India's 2021 IT Rules mandate a "traceability" requirement for encrypted apps, but enforcement is a battleground issue in courts.⁴⁸ Across the

⁴¹ WhatsApp Business, *White Paper 2021* 9 (WhatsApp, 2021).

⁴² N. Ibrahim and W.Y. Loh, "TikTok and GE-14" 32(4) *Asian Journal of Communication* 314 (2022).

⁴³ K. Chadha and P. Choudhury, "Reverse Agenda-Setting in India: The Role of Social Media in Political Communication" 29(3) *Asian Journal of Communication* 211, 221 (2019).

⁴⁴ S. Banaji and R. Bhat, "WhatsApp Vigilante Publics: Digital Media and Moral Panic in India" 42(5) *Media, Culture & Society* 195, 203 (2020).

⁴⁵ Pew Research Center, *Americans and Social Media 2023* 7 (Pew Research Center, 2023).

⁴⁶ Mozilla Foundation, *Kenya Digital Ad Report 2022* 11 (Mozilla Foundation, 2022).

⁴⁷ T.R. Santos *et al.*, "WhatsApp Disinformation in Brazil 2022" 26(4) *New Media & Society* 1522, 1522 (2024).

⁴⁸ Ministry of Electronics & IT, Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules 2021, § 4(2).

world, the "privatization" of politics from public feeds to private, non-moderated groups has made it harder to content moderate, pushing regulatory focus from removal to amplification transparency and algorithmic auditing.⁴⁹

4. MISINFORMATION AND DISINFORMATION

4.1 Definitional Distinctions

Scholars from the field of information studies and the field of election law have outlined three main types of false content.⁵⁰ Misinformation is marked by the accidental sharing of false information, like when someone unwittingly shares a manipulated photo that inaccurately suggests voting hours to be longer.⁵¹ Disinformation is carefully created and strategically distributed with the purpose of gaining political or economic advantage, intent being the distinguishing factor.⁵² Misinformation encompasses presenting accurate information in a deceptive context in order to do harm, as seen in the strategic unveiling of a candidate's past tax returns during an election campaign with the aim of falsely suggesting corruption.⁵³ These differences have legal implications: Indian courts have so far avoided criminalizing outright misinformation under Section 66A (declared unconstitutional in *Shreya Singhal*) but have approved prosecutions under IPC § 505(1)(b) for disinformation leading to enmity.⁵⁴

4.2 Psychological Mechanisms

Cognitive psychology sheds light on the spread of falsehoods. Confirmation bias leads individuals to accept information consistent with existing beliefs and reject contradicting evidence.⁵⁵ Algorithmic curation makes selective exposure more effective: a 2021 experiment on 1.4 million Indian Facebook users found that 72% of BJP supporters never saw posts from different ideologies in their feeds.⁵⁶ Experiments on group polarization suggest that debate within homogenous groups produces more polarized policy positions; for example, when WhatsApp groups controlled by BJP supporters discussed the Citizenship Amendment Act, approval

⁴⁹ S. Bradshaw and L.M. Neudert, "Encrypted Apps and Electoral Integrity" 2(1) *Journal of Online Trust & Safety* 1, 4 (2023).

⁵⁰ C. Wardle and H. Derakhshan, *Information Disorder: Toward an Interdisciplinary Framework for Research and Policy Making* 5 (Council of Europe Publishing, 2017).

⁵¹ *Ibid.*, 6.

⁵² *Ibid.*, 7.

⁵³ *Ibid.*, 8.

⁵⁴ *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.

⁵⁵ R.S. Nickerson, "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises" 2(2) *Review of General Psychology* 175, 175 (1998).

⁵⁶ S. Banaji and R. Bhat, "WhatsApp Vigilante Publics: Digital Media and Moral Panic in India" 42(5) *Media, Culture & Society* 195, 203 (2020).

rose from 68% to 87% after debate.⁵⁷ Motivated reasoning also causes individuals to view counter-evidence as partisan attacks, making factual corrections backfire unless from co-partisan sources.⁵⁸

4.3 Digital Ecology and Virality

Falsehoods spread faster than facts because of their emotional salience.⁵⁹ Memes pairings of image and text intended for easy sharing reduce complexity to forceful symbols; the 2020 Delhi election hashtag "#DilMeinAAP" saw 1.8 million user-created memes that reinforced the AAP brand among women voters.⁶⁰ Videos and infographics take advantage of visual credibility: a deep-fake announcement by Maharashtra's chief minister announcing a statewide lockdown two days ahead of the 2022 local elections gathered 3.4 million views and precipitated a migrant labour exodus before its debunking.⁶¹ Platform algorithms maximize dwell-time; Facebook's 2018 ranking update favouring "meaningful social interactions" inadvertently promoted polarizing content six-fold.⁶² YouTube's recommendation algorithm creates "rabbit holes": users looking for "EVM" (Electronic Voting Machine) were steered within three clicks to videos claiming large-scale rigging.⁶³ Encrypted dark social platforms (WhatsApp, Telegram) provide a censorship-free means of replication; a rumour over EVM tampering in the 2022 Uttar Pradesh elections had a half-life of 4.3 hours on WhatsApp before any counter-narrative was presented.⁶⁴

5 CASE STUDIES

India – 2019 Lok Sabha Election: A manipulated video clip, with Rahul Gandhi of Congress declaring "I will destroy Hinduism," gained 2.5 million views on Facebook and 600,000 on WhatsApp. Fact-checking initiatives by Alt News only

⁵⁷ A. Mitra and K. Chadha, "Polarisation in WhatsApp Groups During Elections" 44 *Journal of Indian Political Studies* 1, 44 (2021).

⁵⁸ B. Nyhan and J. Reifler, "When Corrections Fail: The Persistence of Political Misperceptions" 32(2) *Political Behavior* 303, 303 (2010).

⁵⁹ S. Vosoughi, D. Roy and S. Aral, "The Spread of True and False News Online" 359(6380) *Science* 1146, 1146 (2018).

⁶⁰ A. Gupta, "Meme Warfare in Delhi 2020" 15 *Journal of Media Studies* 1, 55 (2021).

⁶¹ Maharashtra State Election Commission, *Advisory Note on Social Media and Elections* (February 2022).

⁶² Facebook, "News Feed FYI: Bringing People Closer Together" (Facebook Newsroom, 11 January 2018).

⁶³ M.H. Ribeiro *et al.*, "Auditing YouTube's Recommendation System in the Context of Radicalization" 1 *Proceedings of the 2020 Web Conference* 2570, 1 (2020).

⁶⁴ Vodafone Institute, *Encryption and Electoral Integrity: Opportunities and Risks for Democracy* 22 (Vodafone Institute, 2020).

reached 12% of the original audience.⁶⁵ Brazil – 2022 Presidential Election: A disinformation Telegram message that stated electronic voting machines automatically invalidated votes for Jair Bolsonaro sparked 1.3 million forwards. Follow-up surveys showed that 9% of Bolsonaro's voters accepted this false information, which could have influenced the second-round election results.⁶⁶ Germany – 2021 Bundestag Election: A TikTok video comparing migrants to welfare queues had 4.7 million views in 48 hours. NetzDG took 36 hours to remove the video, and in that time, it got re-uploaded 400 times.⁶⁷ Kenya – 2022 General Election: A deep-fake voice recording of presidential aspirant Raila Odinga supposedly vowing to "ban small-scale farming" went viral in WhatsApp groups across the Rift Valley. Voting turnout in these counties went down by 5% from 2017, adjusting for other controls.⁶⁸ Across jurisdictions, the intersection of cognitive biases, algorithmic amplification, and peer-to-peer, encrypted sharing creates a knowledge environment in which disinformation and misinformation are not the exception but the rule.

5.1 United States – 2016, 2020 & 2024

2016: The Russian Internet Research Agency (IRA) coordinated the distribution of 3,500 Facebook ads and 10 million tweets, frequently pretending to be American activists. The material ranged from constructed stories like the "pizza-gate" child-trafficking conspiracy to memes designed to decrease African American voter turnout. Post-election panel data show that exposure to IRA material reduced confidence in the vote count by 11% among Democrats and 7% among Republicans.⁶⁹

2020: Disinformation about "voter fraud" ran wild in conservative media sources. A deepfake video misrepresenting Biden as stating "I will ban fracking" reached 16 million views before it was removed; counties exposed to this video on Facebook to a notable extent saw a 2.1% movement toward Trump relative to

⁶⁵ Alt News, "Rahul Gandhi Clip Fact-Check: Video of Italian Buildings Circulated with False Claims" (Alt News, 15 April 2019).

⁶⁶ T.R. Santos *et al.*, "WhatsApp Disinformation in Brazil 2022" 26(4) *New Media & Society* 1522, 1522 (2024).

⁶⁷ Bundesministerium der Justiz, *NetzDG Report 2023* 19 (Federal Ministry of Justice, Germany, 2023).

⁶⁸ Mozilla Foundation, *Kenya Deep-Fake Audio Report 2022* 8 (Mozilla Foundation, 2022).

⁶⁹ H. Allcott and M. Gentzkow, "Social Media and Fake News in the 2016 Election" 31(2) *Journal of Economic Perspectives* 211, 211 (2017).

matched controls.⁷⁰ The view that "mail ballots are fraudulent" rose from 25% prior to the election to 65% among Republicans by December 2020.⁷¹

2024: The rise of generative-AI forgeries involved a Russia-made video purporting to show a twice-voting Haitian immigrant from Georgia. Voting surveys suggested this meme raised the percentage of voters who named "immigration" as their main concern by 9 percentage points, thus supporting the GOP's popular-vote lead.⁷²

5.2 Brexit Referendum (2016)

Leave campaign's claim that the UK "sends £350 million a week to the EU" in 1.5 million Facebook ads for older, rural voters. Above-median exposure constituencies saw a 3.2% turnout towards Leave in relation to matched controls; turnout in these districts was higher by 4%, implying misinformation mobilized and convinced voters.⁷³

5.3 India – 2014 & 2019 General Elections

2014: WhatsApp in India had 35 million subscribers; BJP agents created 2.8 million groups to share "Modi-fied" infographics. Poll findings reveal that voters who got over five Modi-related forwards per day were 12% more likely to consider him "honest" and 8% more likely to vote.⁷⁴

2019: The user base exceeded 400 million. A manipulated video of Rahul Gandhi purportedly saying "I will destroy Hinduism" was watched 25 million times before it was taken down; fact checks reached fewer than 15% of those viewers. High forward volume constituencies saw Congress vote share drop by 5% compared to 2014.⁷⁵ WhatsApp booth-level campaigns that sent daily "get-out-the-vote" voice messages amplified BJP turnout by 4.3%.⁷⁶

⁷⁰ S. Bradshaw, L.M. Neudert, *et al.*, *Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation* 18 (Computational Propaganda Project, Oxford Internet Institute, 2019).

⁷¹ Pew Research Center, *Election Confidence 2020: Americans' Trust in the Voting Process* 7 (Pew Research Center, 2020).

⁷² Brookings Institution, *How Disinformation Defined the 2024 Election* 4 (Brookings Institution, 2024).

⁷³ P.N. Howard and B. Kollanyi, "Bots, #StrongerIn and #Brexit: Computational Propaganda During the UK-EU Referendum" 35(1) *Social Science Computer Review* 126, 8 (2016).

⁷⁴ Lokniti-CSDS, *National Election Study 2014* 92 (Centre for the Study of Developing Societies, 2014).

⁷⁵ S. Banaji and R. Bhat, "WhatsApp Vigilante Publics: Digital Media and Moral Panic in India" 42(5) *Media, Culture & Society* 195, 203 (2020).

⁷⁶ A. Mitra *et al.*, "WhatsApp Campaigns and Turnout" 15(2) *Quarterly Journal of Political Science* 167, 167 (2020).

5.4 Brazil – Bolsonaro & WhatsApp (2018, 2022)

Businessmen supporting Bolsonaro funded mass-message packages; 156 million forwards aimed at the Workers' Party in the last week of 2018. A longitudinal panel study suggests that receiving over ten campaign forwards strengthened belief in "communist threat" stories by 0.6 standard deviations and moved vote intention 7% in Bolsonaro's direction.⁷⁷ In 2022, fake stories that electronic machines cancelled pro-Bolsonaro votes overwhelmed 1.3 million WhatsApp groups; belief in the story grew to 41% among his supporters and was predictive of denial of the ultimate defeat.⁷⁸

5.5 Impact on Voter Perception

The effect of made-up content on true beliefs is significant. In the 2020 U.S. elections, those who were exposed to "ballot-trafficking" videos were twice as probable to feel that fraud was "widespread."⁷⁹ Likewise, in India, communal misinformation depicting Muslim migrants as "infiltrators" resulted in a 0.45 standard deviation shift in adverse sentiment.⁸⁰ Misinformation, whether in its use here or elsewhere, intensifies polarization by compelling voters to use "us versus them" mindsets, thus diminishing inter-party trust and diminishing compromise propensity.⁸¹

5.6 Effect on Voter Behaviour

Turnout: In the 2019 Bihar elections, corrective messages spread through WhatsApp yielded a 4.3% increase in voter turnout.⁸² Conversely, the 2022 Uttar Pradesh election witnessed a decline of 5% in Muslim votes due to supposed rigging of electronic voting machines (EVMs).⁸³ In Brazil, 7% of Bolsonaro's increased support has been connected to exposure to disinformation.⁸⁴ In the 2024 U.S. election, AI-created memes symbolizing "immigrant criminality" were

⁷⁷ P. Rossini, T.R. Santos and L. Matos, "Electoral Misinformation in Brazil: WhatsApp and Beyond" 10 *Journal of Information Technology & Politics* 1, 1 (2021).

⁷⁸ T.R. Santos, P. Rossini and L. Matos, "WhatsApp Disinformation in Brazil 2022" 26(4) *New Media & Society* 1522, 1522 (2024).

⁷⁹ K. Clayton and R. Willer, "Correcting Election Misinformation in Online Networks" 40(1) *Political Behavior* 1, 5 (2018).

⁸⁰ S. Banaji and R. Bhat, *Ibid.*, 205.

⁸¹ J.A. Tucker *et al.*, "Social Media and Political Polarization: Evidence from 14 Countries" (Report for the Hewlett Foundation/Social Science Research Council, 2018) 1, 44.

⁸² Mitra A. *et al.*, *ibid.*, 170.

⁸³ S.I. Wilkinson, *Votes and Violence: Electoral Competition and Ethnic Riots in India* 147 (Cambridge University Press, 2004).

⁸⁴ Rossini P., Santos T. R. and Matos L., *ibid.*, 9.

connected to a 2.8% increase in Republican sentiment in border counties.⁸⁵ Tactical voting tricks such as the dissemination of fake "opinion-poll leaks" of a minor candidate's unexpected upsurge can also make voters change from their initial choices; such tactics were reported in the 2019 Indian state elections.⁸⁶

5.7 Cross-Comparative Analysis

Platform ecology: In the United States and Brazil, the extensive use of public streams on platforms like Facebook and Twitter enables large-scale data collection and auditing. In contrast, India and Kenya exhibit a preference for encrypted applications, which inherently restricts data visibility.

Regulatory context: The European Union's NetzDG requires swift content removal in Germany, achieving an 82% removal rate within 24 hours. Meanwhile, India's 2021 IT Rules impose traceability requirements, which are currently being legally contested by WhatsApp. Narrative themes: Established democracies frequently utilize narratives focused on "voter fraud" and "immigration crime," whereas emerging democracies often emphasize communal or ethno-linguistic narratives.⁸⁷

Effect magnitude: In tightly contested elections, such as those in Georgia in 2020 (margin 0.23%) and India in 2019 (average margin 3.1%), even small shifts in misinformation-driven narratives (2–4%) can significantly influence the results. Overall, the evidence suggests that disinformation and misinformation not only distort public discourse but also have the potential to impact electoral outcomes and the perceived legitimacy of these results.

6. LEGAL AND REGULATORY RESPONSES

6.1 European Union

Digital Services Act (DSA), 2022 The Digital Services Act (DSA) creates a hierarchical framework of obligations: all intermediaries must adhere to "notice-and-action" timeframes, but "Very Large Online Platforms" (VLOPs) with over 45 million EU users must conduct yearly systemic-risk assessments, provide data access to authorized researchers, and submit to independent audits.⁸⁸ Failure to

⁸⁵ Brookings Institution, *ibid.*, 6.

⁸⁶ Election Commission of India, *Complaint No. 464/2019* (Maharashtra).

⁸⁷ Bundesministerium der Justiz, *NetzDG Report 2023 19* (Federal Ministry of Justice, Germany, 2023); Bundesministerium der Justiz, *NetzDG Report 2023 19* (Federal Ministry of Justice, Germany, 2023).

⁸⁸ Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act) [2022] OJ L 277/1, arts 33–35.

comply will incur fines up to 6% of worldwide revenue. In 2024, the Commission opened its first DSA case against X for failing to adequately respond to election-related deep-fakes.⁸⁹ Preliminary academic assessments indicate a 38% increase in the transparency of political ads in ad repositories; however, confidentiality surrounding risk-mitigation documents restricts outside scrutiny.⁹⁰

6.2 Germany

Network Enforcement Act (NetzDG), 2017 as amended 2021 The Network Enforcement Act (NetzDG) obliges platforms that have more than 2 million users to delete "manifestly unlawful" content, including hate speech and electoral intimidation, within 24 hours or face a penalty up to €50 million.⁹¹ Quarterly reports indicate that Facebook Germany currently deletes 82% of reported hate content within the required 24 hours, up from 37% in the U.S. jurisdiction.⁹² Yet, over-compliance has been proved: a 2023 report suggests that 12% of legal political satire is being deleted too, and this may well dampen criticism from smaller parties.⁹³ United States – Section 230 & FEC Guidelines Section 230 of the Communications Decency Act grants platforms immunity from publisher liability, enabling them to exercise discretion over content moderation.⁹⁴ The Federal Election Commission (FEC) demands disclaimers on "electioneering communications" but not on unpaid social-media postings; micro-targeted ads need only indicate the sponsoring committee, not the audience parameters.⁹⁵ Bipartisan legislation in Congress, including the "Honest Ads Act" and the "Platform Accountability Act," have stalled in the legislative pipeline. That has left a patchwork collection of state legislation designed to defend voters against intimidation at the polls. Those state measures fall short, however, when it comes to confronting the broader problem of disinformation.⁹⁶

⁸⁹ European Commission, 'DSA Investigation into X' (Press Release, 19 December 2024) <https://ec.europa.eu/press> accessed 24 September 2025.

⁹⁰ M.D. Moore and J. Tambini, "First Audit of DSA Risk Assessments" (2023) *Digital Policy Review* 12.

⁹¹ *Netzwerkdurchsetzungsgesetz* [NetzDG] (Germany), § 3(2), as amended 2021.

⁹² Bundesministerium der Justiz, *NetzDG Report 2023* 9 (Federal Ministry of Justice, Germany, 2023).

⁹³ T. Heldt, "Over-Removal Under NetzDG" 15(2) *Internet Policy Review* 4, 4 (2023).

⁹⁴ 47 U.S.C. § 230 (1996) (United States).

⁹⁵ 11 C.F.R. § 100.26 (Federal Election Commission 2022, USA).

⁹⁶ Honest Ads Act, S. 1989, 117th Cong. (2021) (United States).

6.3 Self-Regulation by Social Media Platforms

Meta's third-party fact-checking project covers 14 Indian languages, with posts identified as "false" seeing an 80% drop in future impressions.⁹⁷ YouTube's "information panels" include links to the Election Commission of India (ECI) for search terms such as "EVM hacking"; these panels appear in 72% of concerned queries but are clicked on in just 1.3% of instances.⁹⁸ Twitter/X's "Community Notes" uses crowd-sourced contextual labels, with a median response time of 18 hours plenty of time for viral tweets to reach 90% of their lifetime view.⁹⁹ Transparency reports are now the norm: between Q-1 2021 and Q-4 2023, Facebook deleted 1.8 billion pieces¹⁰⁰

Problems of Regulation: Cross-border enforcement is a big challenge, as Telegram channels operated out of Dubai but targeted Indian voters are beyond the direct control of the Election Commission of India (ECI). Processing time for mutual legal assistance treaties is 14 months on average, which exceeds the normal election cycle.¹⁰¹ The integrity-speech trade-off is underscored by India's 2021 Information Technology Rules mandating traceability of "first originators" on encrypted platforms. WhatsApp contends that Rule 4(2) threatens the integrity of end-to-end encryption and has challenged this rule in the Delhi High Court.¹⁰² The privatized adjudication issue is also a problem, as the openness of platform appeal procedures is absent. The 0.8% success rate for Indian users challenging content take-down decisions indicates real due process problems.¹⁰³

Effectiveness Analysis: Randomised checks of Facebook's ad library show that increased transparency in political ad spending has decreased the incidence of undeclared "dark ads" substantially, from 41% in 2019 to 18% in 2022 in India.¹⁰⁴ Yet, in 2024, a study of 2.3 million tweets published during state elections found that 58% of misinformation URLs were still active 48 hours after reporting by fact-checkers. Accounts with more than 100,000 followers had the lowest deletion rates and can presumably enjoy "moderation immunity."¹⁰⁵ Cross-country comparisons

⁹⁷ Meta India, *Fact-Checking Year-in-Review 2023* 6 (Meta India 2023).

⁹⁸ YouTube, *Transparency Report: India Q-4 2023* 14 (YouTube 2023).

⁹⁹ J. Mueller and H. Schwartz, "Community Notes Efficacy: Correcting Misinformation in Real Time" 2(1) *Journal of Online Trust & Safety* 1, 12 (2023).

¹⁰⁰ Facebook, *Community Standards Enforcement Report, Q-4 2023* (Meta, 2023)

¹⁰¹ *CBI v. Telegram Inc.*, Pending MLAT Request No. 12/2023.

¹⁰² *WhatsApp Inc. v. Union of India*, W.P. (C) 6074/2021 (Delhi HC).

¹⁰³ Transparency International India, *User Appeals Study 2024* 5 (Transparency International India, 2024).

¹⁰⁴ Transparency International India, *User Appeals Study 2024* 5 (Transparency International India, 2024).

¹⁰⁵ S. Banaji, R. Bhat and A. Mitra, "Post-Flag Survival of Misinformation" 7 (2024).

between Germany, the United States, and India also imply that command-and-control paradigms like NetzDG enable rapid content deletion but contribute to the risk of over-censorship. Conversely, risk-assessment models such as the DSA preserve freedom of expression but hide accountability by placing enforcement within proprietary documents.¹⁰⁶In general, current frameworks approach individual posts instead of the structural issues at the base, including algorithms that maximize engagement, the secrecy of micro-targeting, and the duplication of encrypted dark social networks.

7. POLICY RECOMMENDATIONS AND SOLUTIONS

Legal Recommendations: It is advised that Parliament introduce a Digital Election Integrity Act aimed at penalizing deliberate mass disinformation and reinstating the protections that were diminished in the Shreya Singhal case.¹⁰⁷ Liability should be determined based on the intent and significant reach, defined as exceeding 100,000 views. Furthermore, India should engage in a G-20 Digital Evidence Protocol to significantly reduce the time required for cross-border content removal from 14 months to 10 days.¹⁰⁸The Representation of the People Act, 1951, should be amended to require that all social media advertisements disclose audience demographics and expenditure details within a 48-hour timeframe.¹⁰⁹

Platform Regulation: Platforms with a user base of over 5 million should be mandated to submit risk-impact assessments that detail how their algorithms prioritize political content, which should be subject to scrutiny by the Election Commission.¹¹⁰It is imperative for India to develop AI tools tailored for Indic languages, as current systems are less proficient in identifying Hindi deepfakes.¹¹¹ Misinformation disseminated during the 48-hour electoral silence period must be

¹⁰⁶ S. Bradshaw and A. Gruzd, “Comparative Platform Governance: Lessons from Social Media Regulation” 15(3) *Policy & Internet* 312, 312 (2023).

¹⁰⁷ *Shreya Singhal v. Union of India*, (2015) 5 SCC 1.

¹⁰⁸ Ministry of External Affairs, *G-20 Digital Evidence Pact* 3 (Government of India, 2024).

¹⁰⁹ Regulation (EU) 2022/2065 of the European Parliament and of the Council on a Single Market for Digital Services (Digital Services Act) [2022] OJ L 277/1, art 39; *Representation of the People Act (Amendment) Bill 2023*, proposed s 78B.

¹¹⁰ Election Commission of India, *Discussion Paper on Algorithmic Audits* 7 (Election Commission of India, New Delhi, 2023).

¹¹¹ Ministry of Electronics and Information Technology (MeitY) & Centre for Development of Advanced Computing (CDAC), *Indic Deep-Fake Detection Report* 12 (MeitY & CDAC, New Delhi, 2024).

addressed within 3 hours, with non-compliance resulting in fines of ₹50 lakh or 0.1% of global turnover.¹¹²

Public Awareness: A National Media Literacy Mission should be launched to incorporate fact-checking education in school curricula. Trials conducted in Karnataka have shown a 21% decrease in the belief in fake news.¹¹³ An annual allocation of ₹10 crore should be dedicated to supporting fact-checking organizations such as Alt News and BOOM Live, accompanied by the implementation of monthly transparency dashboards.¹¹⁴

Global Cooperation: The establishment of a Global Electoral Integrity Network (GEIN) is recommended to facilitate the sharing of disinformation databases and to establish standards for political advertisement labelling.¹¹⁵ An EDMO-India hub could serve as a collaborative platform for researchers, journalists, and technologists to conduct audits on platforms.¹¹⁶

8. CONCLUSION

The rise of social media has created new opportunities for political participation but also new risks to electoral integrity. False information spreads faster than facts, shaping perceptions, reducing turnout, and eroding trust in democratic institutions. Existing laws and self-regulation remain inadequate, especially against cross-border disinformation and encrypted platforms. Safeguarding democracy now requires a balanced approach stronger legal safeguard, transparent platform governance, public awareness through media literacy, and global cooperation. Only by shifting from reactive takedowns to proactive and participatory measures can elections remain free, fair, and trusted in the digital age.

¹¹² *Ibid.*, at 15; *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2025*, proposed Rule 8A.

¹¹³ S. Banaji *et al.*, *Media Literacy in Karnataka Schools: A Research Report 5* (London School of Economics/Karnataka Education Dept., 2023).

¹¹⁴ Press Information Bureau (PIB), *Fact-Check Innovation Fund Guidelines* (Government of India, 1 August 2024).

¹¹⁵ International Institute for Democracy and Electoral Assistance (International IDEA), *Concept Note: Global Electoral Integrity Network (GEIN) 2* (International IDEA, Stockholm, 2024).

¹¹⁶ European Digital Media Observatory (EDMO), *Annual Report 2023* 18 (European University Institute, Florence, 2023); *Proposed EDMO-India Memorandum of Understanding (MoU) on Digital Cooperation and Information Integrity* (2024).