Strengthening Trust: Countering Digital Deception in Elections

Dr. Nasim Zaidi, Former Chief Election Commissioner of India

Table of contents

Part I: Introduction

The internet, a transformative force in modern society, has redefined how individuals access information and engage in public discourse, thus reshaping the traditional state-subject relationship.[1] Central to this transformation is the rise of social media, which has not only democratised information but also reinvigorated political engagement, empowering citizens to voice their opinions, organise movements, and hold governments accountable, evident in numerous social movements worldwide, showcasing social media's role in challenging state authority.[2] Conversely, these platforms also serve as tools for states to engage in surveillance, set political agendas, and influence public opinion. Political campaigns have increasingly relied on social media to engage with voters directly, revolutionising dissemination of messages and support mobilisation.[3] For instance, during the 2016 U.S. presidential election, candidates used platforms like Twitter and Facebook to reach millions of followers daily, bypassing traditional media channels to control their messaging.[4] However, while these platforms have enhanced political engagement and transparency by facilitating open discourse, they are also susceptible to misuse,[5] highlighting the need for careful management and regulation of digital spaces to maintain the integrity of democratic engagements.

The effects of misinformation and disinformation are profound, altering public discourse, intensifying polarisation, and eroding public confidence in the electoral process and broader democratic institutions.[6] Additionally, the emergence of deep fakes — highly realistic AI-generated fake videos or audio recordings — presents a new frontier in digital deception.[7] Accordingly, this paper reviews and examines the influence of social media on elections across jurisdictions, referring to international standards and measures, and concludes with actionable recommendations to combat misinformation and disinformation, including deepfakes.

---

[1] Schrape & Jan-Felix, *Technology and the Promise of Decentralization: Origins, Development, Patterns of Arguments*, Department of Organizational Sociology and Innovation Studies, University of Stuttgart (2019), https://www.econstor.eu/bitstream/10419/194289/1/1067704019.pdf; Philip N Howard, *Deep Democracy, Thin Citizenship: The Impact of Digital Media in Political Campaign Strategy*, The Annals of the American Academy of Political and Social Science (2005), https://www.jstor.org/stable/25046067.

[2] Daud Isa and Itai Himelboim, A Social Networks Approach to Online Social Movement: Social Mediators and Mediated. Content in #FreeAJStaff Twitter Network, 1(14) Social Media + Society (2018), https://journals.sagepub.com/doi/pdf/10.1177/2056305118760807.

[3] Dommett, K. & Temple, L., 2018. Digital campaigning: The rise of Facebook and satellite campaigns. *Parliamentary Affairs*, 71(suppl_1), 189-202.

[4] Hendricks, J. A. & Schill, D., 2017. The Social Media Election of 2016. In *The 2016 US Presidential Campaign: Political Communication and Practice*. 121-150.

[5] Murali, C. & Charulatha, A. S., 2017. New Media and the 2016 US Presidential Election: A Case Study. *Research Journal of Humanities and Social Sciences*, 8(2), 244-254.

[6] *Id.*

[7] Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C., 2020. Deepfakes: Trick or Treat? *Business Horizons*, 63(2), 135-146; Maria Pawelec, *Deepfakes and Democracy (Theory): How Synthetic Audio-Visual Media for Disinformation and Hate Speech Threaten Core Democratic Functions*, Springer Nature (2022), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9453721/.

Part II: Global Regulation of Social Media Manipulation

A. Regional Guidelines and Policy Instruments

a. The European Union

The EU, a frontrunner in effective policy making, has established robust regulations for social media platforms and search engines, underscoring its commitment to transparency and accountability.[8] This dedication extends to election integrity, where the EU implements stringent measures to protect democratic processes. The EU Code of Practice on Disinformation, a key policy initiative, calls on major online platforms to voluntarily combat disinformation and promote transparency. First introduced in 2018 and then reinforced in 2022, the Code encompasses 44 commitments and 128 specific measures, spanning areas such as demonetisation, transparency of political advertising, user empowerment, and a rapid response system for cooperation during elections.[9]

Currently, the Code of Practice on Disinformation has garnered significant support, with 44 entities, including Meta, TikTok, Google, and Microsoft, signing on. The EU is now considering the transformation of the Code of Practice on Disinformation and the Code of Conduct on Countering Illegal Hate Speech Online into Codes of Conduct under the co-regulatory framework of the DSA. Building on the Code of Practice of Disinformation and other relevant policy instruments,[10] on 26th March 2024, the EU Commission published Guidelines on recommended measures to 'Very Large Online Platforms' (VLOPs) and Very Large Online Search Engines (VLOSEs) - as defined under the recently enacted Digital Services Act (DSA) - to mitigate systemic risks online that may impact the integrity of elections, after a consultative period of one month.[11] These guidelines are non-binding and became applicable in late April 2024. The guidelines propose the implementation of mitigation measures and the adoption of best practices by Very Large Online Platforms and Search Engines throughout the entire electoral process, consisting of periods before, during, and after electoral events.

As noted, the Guidelines, *inter alia*, aim to:[12] (i) strengthen the internal processes of VLOPs and VLOSEs by setting up internal teams utilising local context-specific risk information and considering user search behaviour throughout the electoral cycle, (ii) implement tailored risk mitigation measures for each

---

[8] *Tackling online disinformation*. (2024, June 6). Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/policies/online-disinformation.

[9] *The 2022 Code of Practice on Disinformation*. (2024, June 6). Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation.

[10] The 2022 Code of Practice on Disinformation. (n.d.). *European Commission*. https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation.

[11] *Guidelines for providers of VLOPs and VLOSEs on the mitigation of systemic risks for electoral processes*. (2024, April 26). Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/library/guidelines-providers-vlops-and-vloses-mitigation-systemic-risks-electoral-processes; *Press corner*. (n.d.-b). European Commission - European Commission. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_1707; *Digital Services Act: Summary report of the public consultation on guidelines for providers of very large online platforms and search engines on the mitigation of systemic risks for electoral processes*. (2024, March 26). Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/library/digital-services-act-summary-report-public-consultation-guidelines-providers-very-large-online.

[12] Calvet-Bademunt, J. (2024b, April 10). *Digital Services Act Roundup: February - March 2024*. Tech Policy Press. https://www.techpolicy.press/digital-services-act-roundup-february-march-2024/.

electoral period, such as promoting official electoral information, labelling political ads, conducting media literacy initiatives, and adjusting recommender systems to reduce harmful content., (iii) adopt specific mitigation measures for generative AI, such as via labelling, (iv) cooperation with EU and national authorities, independent experts, and civil society for efficient information exchange, (v) adopt incident response mechanisms, (vi) assess the effectiveness via post-elections reviews, and publish a non-confidential version of these reviews.

Recently, the EU conducted a simulation exercise - 'stress test' - with the designated platforms, Digital Services Coordinators, and civil society organisations to test their readiness against election manipulation and interference in relation to the European election.[13] As noted, the test aimed to understand platforms' readiness to tackle manipulative behaviour which could impact elections - this included (i) Information manipulation enabled by 'deep fakes' or other uses of AI to distort and manipulate audio-visual content, (ii) Information manipulation through coordinated inauthentic behaviour, (ii) Attempts of suppression of voices, including through harassment and threats online, (iii) Intentional spread of false information on the electoral process to mislead voters, (iv) Online incitement to violence based on manipulated information, (v) Cyber-enabled activity that is used for information manipulation.[14] The results of the stress test are not available in the public domain.

b. The African Union

The African Union has issued the "Principles and Guidelines for the Use of Digital and Social Media in Elections in Africa"[15] to address the growing challenge of misinformation and disinformation during elections on the African continent. These guidelines are a non-binding, human rights-inspired instrument aimed at enhancing the capacities of Election Management Bodies (EMBs) and other electoral stakeholders to harness the advantages of social media while mitigating its adverse effects. The guidelines address, *inter alia*, the follow-

- **Regulating Data Use:** (i) It mandates the measures to prevent recommender systems from amplifying disinformation and misinformation, such as prohibiting the processing of personal data, including special categories, within a group of undertakings or sharing it with other entities, (ii) It also prohibits the use of real-time bidding (RTB) technology to broadcast data about voters, as this infringes on the principle of data confidentiality and integrity.
- **Empowering Election Management Bodies:** (i) It empowers African EMBs to adopt clear and comprehensive plans for the responsible use of social media during electoral periods, (ii) It also encourages government and regulatory bodies to support EMBs in safeguarding elections and combating the threat of digital disinformation and misinformation.
- **Promoting Transparency and Accountability:** The guidelines emphasise on the importance of transparency and accountability in the use of digital and social media during elections, with a focus on ensuring that the processing of personal data respects the principles of lawfulness, fairness, and transparency.

---

[13] *Commission stress tests platforms' election readiness under the Digital Services Act*. (2024, April 24). Shaping Europe's Digital Future. https://digital-strategy.ec.europa.eu/en/news/commission-stress-tests-platforms-election-readiness-under-digital-services-act

[14] *Id.*

[15] Principles and Guidelines for the Use of Digital and Social Media in Elections in Africa. (n.d.). *Electoral Commission of South Africa*. https://www.elections.org.za/pw/Elections-And-Results/Principles-and-Guidelines-for-the-use-of-the-Digital-and-Social-Media-in-Elections-in-Africa.

c. Association of South-East Asian Nations (ASEAN)

The ASEAN guidelines[16] emphasise a multi-stakeholder approach involving governments, media, technology companies, and civil society to combat the spread of misinformation and disinformation during the elections. The guidelines, *inter alia*, recommend that (i) governments should prioritise media literacy education to empower citizens to combat fake news effectively, fostering critical thinking and resilience against misinformation; (ii) governments should implement strategies to detect and counter disinformation campaigns and their origins, whether originating within the country or internationally; (iii) fact-checking organisations should receive sufficient funding and support to effectively combat fake news and disinformation; (iv) media organisations must uphold ethical journalism practices, including rigorous fact-checking, source verification, and avoidance of sensationalism, to combat false news and disinformation effectively; (v) ASEAN Member States should collaborate on sharing articles and international news while establishing verification channels to combat international fake news effectively; (vi) governments and media should leverage technology, including AI and blockchain, to detect and counter fake news and disinformation effectively, enhancing information security and transparency. This said, the guidelines are not focused on elections. ASEAN members in their upcoming meetings should consider deliberating on this.

B. Jurisdictions across the Globe

a. France

In 2018, France adopted a pair of new legislation to fight against false information; it highlighted electoral events demonstrating the impact of massive campaigns spreading false information to disrupt the electoral process. Acknowledging existing civil and criminal liabilities for spreading false information but notes the insufficiency of current laws in enabling the rapid removal of such content.[17] The law, *inter alia*, provides that public prosecutors, candidates, political parties, or interested parties can appeal to a judge to curb the dissemination of "false information" in the three months leading up to an election, whereby the judge has to make a decision within 48 hours of complaint. Additionally, it granted the French broadcasting agency (CSA) the power to suspend television channels "controlled by a foreign state or influenced" by that state if they "deliberately spread false information that could undermine the integrity of the election."[18] In addition to the "notice and takedown" obligation to promptly remove illegal content flagged to them, the service providers now have enhanced obligations: a) Implement an easily accessible and visible system for users to report content containing misinformation, b) Promptly relay these user reports about misinformation to the relevant public authorities, c) Publicly disclose the resources/efforts they are devoting to combating misinformation dissemination on their platforms. These efforts advocate

[16] ASEAN guidelines on management of government of information in combating fake news and disinformation in the media. (n.d.). *Ministry of Communications and Informatics Republic of Indonesia*. https://asean.org/wp-content/uploads/2023/06/ASEAN-Guideline-in-Combating-Fake-News-and-Disinformation-in-the-Media-ISBN.pdf.

[17] Fiorentino R, M. (2018). *France passes controversial 'fake news' law*. Euro News. https://www.euronews.com/2018/11/22/france-passes-controversial-fake-news-law; Boring, N. (2019, April). *Initiatives to Counter Fake News: France.* Library of Congress Law. https://maint.loc.gov/law/help/fake-news/france.php#:~:text=The%20new%20law%20also%20provides,fake%20or%20misleading%20information%20online.

[18] Fake news, French law and democratic legitimacy Lessons for the United Kingdom? (2019). *Journal of Media Law*. https://www.pure.ed.ac.uk/ws/portalfiles/portal/120126408/CraufurdSmithJML2019FakeNewsFrenchLaw.pdf.

for a balance between regulation and self-regulation, encouraging proactive steps by platforms in collaboration with governments and civil society. These new provisions aim to reinforce the duty of technical intermediaries to cooperate proactively in tackling online misinformation.[19] Notably, the law was rejected twice by the Senate due to heavy criticism of the proposed law, especially with regards to a disproportionate curtailment of freedom of speech and expression.

In 2021, a 50-person department called Virginum ("Vigilance and Protection against Foreign Digital Interference Service") was established to counter foreign digital interference and disinformation campaigns during elections and major events like the Olympics.[20]

Despite these efforts, the recent French elections were still wrought with the dissemination of misinformation and disinformation. For instance, it was recently reported that an allegedly Russian fake news machinery endeavoured to impact the French legislative elections by impersonating French media organisations,[21] or even when, after the elections, misinformation surfaced regarding the nature of the celebrations where videos promoting an anti-immigrant narrative spread widely on social media, alleging that the jubilations prominently featured Palestinian and Algerian flags.[22] As noted, the reporting media house attempted to dispel the false information and contain the impact.

b. South Africa

South Africa emphasises the importance of partnering with major social media companies to combat disinformation. The South African Election Commission has recently signed a Framework of Cooperation with platforms like Google, Meta, TikTok, and non-profit organisation Media Monitoring Africa (MMA) to enable more effective and efficient content moderation and curb the spread of false information.[23] The South African Framework, *inter alia*, establishes a working group between the Election Commission and its partners to coordinate efforts, promote access to accurate information, conduct awareness campaigns, and provide training to election stakeholders. It also enables signatories to cooperate with the Election Commission and MMA on initiatives like Real411.org (a complaints platform)[24] and PADRE.org.za (a repository of election-related information). South Africa's approach includes forming partnerships,

---

[19] *Id.*

[20] Louis, L. (2024, April 12). France fights disinformation as Olympics, elections loom. *dw.com*. https://www.dw.com/en/france-fights-disinformation-as-olympics-elections-loom/a-68759644;Vigilance and Protection against Foreign Digital Interference Service (VIGINUM). (n.d.). *General Secretariat for Defence and National Security*. https://www.sgdsn.gouv.fr/files/files/Publications/RA-Viginum-Annee1-32p-V20_EN_LQP-1.pdf.

[21] Bahl, V. (2024, July 9). *Here's how Moscow's fake news machine tried to interfere with French elections.* France 24. https://www.france24.com/en/tv-shows/truth-or-fake/20240709-here-s-how-moscow-s-fake-news-machine-tried-to-interfere-with-the-french-elections.

[22] Bahl, V. (2024, July 8). *"Not a French flag in sight": Fake news after left-wing victory in snap elections*. France 24. https://www.france24.com/en/tv-shows/truth-or-fake/20240708-not-a-french-flag-in-sight-after-the-left-s-win-fake-news-stirs-up-anti-immigrant-rhetoric.

[23] *Electoral Commission : News Article*. (n.d.-b). https://www.elections.org.za/content/About-Us/News/Electoral-Commission-partners-with-social-media-giants-to-combat-disinformation-in-2024-National-and-Provincial-Elections/.

[24] Real411 digital disinformation reporting platform, enabling the public to report cases of election-related disinformation. These complaints are assessed by a panel of experts, and the IEC can take actions such as referring cases to the Electoral Court or social media platforms.

establishing dedicated reporting platforms, aligning with electoral laws, launching awareness campaigns, and utilising specialised investigative bodies.

c. South Korea

The National Election Commission of South Korea has amended the Public Official Election Act to specifically ban the creation and distribution of deepfakes content related to election campaigns within 90 days of an election. The country ramped up efforts by law enforcement and prosecutors to identify and take action against individuals spreading misinformation.[25] Election monitors have set guidelines to mitigate the risks of AI-generated content. These guidelines mandate transparency in the use of AI for political communication and require clear disclosure of any AI-assisted content, ensuring that voters are not deceived by AI-generated falsehoods.

South Korea's leading search engine, Naver, has intensified its monitoring efforts to combat new forms of abuse, such as AI-generated comments and deepfakes. The platform also introduced features that allow users to directly report election misinformation, with a dedicated reporting centre established to facilitate communication with the National Election Commission. Furthermore, KakaoTalk, the country's leading messaging app, introduced the "Karlo AI Profile," which adds watermarks to AI-generated content. Deepbrain AI, a Korean generative AI company, announced a collaboration with South Korea's National Police Agency to develop a detection tool for tracking deepfakes and addressing election-related crimes. Additionally, private sector initiatives have emerged, such as the joint declaration to prevent the malicious use of election-related deepfakes, signed by Naver, Kakao, and SK Communications, and the global AI Elections Accord.[26]

d. United States of America

Social media platforms have increasingly become conduits for manipulating political narratives in the U.S. In 2016, it was alleged that Russian agents utilised social networking sites to disseminate divisive content to millions of Americans before the presidential election. This alleged dissemination, reaching a staggering 126 million users on a social media platform, involved approximately 80,000 posts and 131,000 tweets on another, often centring on polarising topics such as race, religion, and gun rights.[27] Consequently, these occurrences elicited calls for enhanced accountability and transparency from social media intermediaries, prompting a reevaluation of social media's role in electoral processes and advocating for improved advertising disclosure policies and heightened security measures to mitigate future manipulative efforts.

---

[25] South Korea contends with AI and electoral integrity. (2014). *East Asian Forum*. https://eastasiaforum.org/2024/04/01/south-korea-contends-with-ai-and-electoral-integrity/.

[26] Lee, S. (2024b, May 13). AI and Elections: Lessons From South Korea. *The Diplomat*. https://thediplomat.com/2024/05/ai-and-elections-lessons-from-south-korea/.

[27] *Russia-backed Facebook posts 'reached 126m Americans' during US election*, The Guardian https://www.theguardian.com/technology/2017/oct/30/facebook-russia-fake-accounts-126-million; David E Sanger & Nick Corasaniti, *David E. Sanger and Nick Corasaniti, D.N.C. Says Russian Hackers Penetrated its Files, Including Dossier on Donald Trump, N.Y. TIMES*, The New York Times (Apr. 14, 2016), https://www.nytimes.com/2016/06/15/us/politics/russian-hackers-dnc-trump.html.

Lawmakers scrutinised executives from top tech firms in a hearing, addressing concerns about foreign interference in U.S. elections.[28] Revelations of extensive Russian disinformation campaigns on social media platforms prompted calls for increased transparency and regulation. Despite efforts to combat meddling, tech companies fell short of endorsing legislation demanding disclosure standards for political advertisements akin to traditional media. Certain social media platforms highlighted efforts to curb abuse on their platforms, but lawmakers remained sceptical, urging greater cooperation and transparency. Amidst calls for regulation, a proposed bill titled "Honest Ads Act"[29] seeks to amend the Federal Election Campaign Act of 1971 to compel digital platforms to publicly disclose political advertisement information, along with their extant obligation for print, newspaper, and television. While tech companies have launched their transparency initiatives, the debate over regulating online political advertising continues.

e. United Kingdom

After the Brexit referendum in 2016, concerns about misinformation and disinformation in the UK began to escalate, primarily driven by internet propaganda that spread incorrect information.[30] The 2019 U.K. general election, dubbed "The Brexit Election"[31] by Sky News, turned into a battleground of framing and messaging, intensifying global concerns about the spread of misinformation and disinformation on social media platforms.[32] The Labour Party's complaint to Ofcom regarding this labelling was overshadowed by the relentless focus on Boris Johnson's "Get Brexit Done" message, which dominated the discourse. Despite efforts to shift the conversation towards topics like taxes and the economy, the media spotlight often reverted to sloganeering rather than providing in-depth policy analysis.[33]

In response to these challenges, the UK Electoral Commission introduced the Online Harms White Paper in 2019[34], which laid the groundwork for establishing a fundamental "duty of care" principle towards users. After significant amendments and extensive deliberations, this initiative culminated in the passage

---

[28] Craig Timberg et al., *Tech executives try to placate lawmakers on disclosure*, The Washington Post (Oct. 31, 2017), https://www.washingtonpost.com/business/technology/tech-executives-try-to-placate-lawmakers-on-disclosure/2017/10/31/7d0831ea-be7f-11e7-8444-a0d4f04b89eb_story.html.

[29] Congressional Record, House of Representatives, Proceedings and Debates of the 166th Congress, First Session, https://www.congress.gov/116/crec/2019/08/01/CREC-2019-08-01.pdf.

[30] Marco T Bastos & Dan Mercea, *The Brexit Botnet and User-Generated Hyperpartisan News*, 37 Social Science Computer Review 38-54 (2019), https://journals.sagepub.com/doi/10.1177/0894439317734157.

[31] *Labour complains to Ofcom about Sky's 'Brexit Election' slogan*, The Guardian https://www.theguardian.com/politics/2019/nov/15/labour-complains-to-ofcom-about-skys-brexit-election-slogan.

[32] *Will fake news wreck the coming general election?*, The Guardian https://www.theguardian.com/media/2019/oct/06/will-fake-news-wreck-next-british-general-election.

[33] *Labour complains to Ofcom about Sky's 'Brexit Election' slogan*, The Guardian https://www.theguardian.com/politics/2019/nov/15/labour-complains-to-ofcom-about-skys-brexit-election-slogan; Cristian Vaccari et al., *The Campaign Disinformation Divide Believing and Sharing News in the 2019 UK General Election*, 40 School of Social Sciences and Humanities (2023), https://www.tandfonline.com/doi/epdf/10.1080/10584609.2022.2128948?needAccess=true.

[34] The Secretary of State for Digital, *Online Harms White Paper: Full Government Response to the Consultation*, Presented to Parliament (2020),the Secretary of State for Digital, *Online Harms White Paper: Full Government Response to the Consultation*, Presented to Parliament (2020), https://data.parliament.uk/DepositedPapers/Files/DEP2020-0111/Online_Harms_White_Paper-Initial_consultation_response.pdf.

of the "Online Safety Bill" by the UK Parliament in 2023. This legislation governs not only social media platforms but also encompasses file-sharing sites, discussion forums, and e-commerce websites, mandating them to take responsible measures to ensure user safety, transparency, and to address harmful content effectively. Furthermore, the bill emphasises the importance of media literacy, equipping users with the necessary digital skills to combat misinformation and other online threats.[35] The Office of Communications (OFCOM) is designated as the primary regulator under this new law, endowed with the authority to impose significant fines on social media platforms and their key personnel. The law adopts a risk-based approach, aiming to ensure that regulatory actions are proportionate to the level of harm experienced by individuals.[36]

Notably, instances of disinformation were observed during the recent 2024 U.K. elections., such a false audio clip - a deepfake - of a Labour politician using foul and abusive language towards a member of the public who disagreed with him over the war in Gaza. The media gained substantive traction before it was flagged as "manipulated media" on X.[37]

f. Brazil

In 2021, the Superior Electoral Court (TSE) in Brazil launched the Electoral Justice Permanent Program on Countering Disinformation to mitigate the harmful effects of disinformation related to the electoral process. The initiative excludes disinformation targeting pre-candidates, candidates, political parties, coalitions, and federations from its scope, except when the content has the potential to harm the integrity, credibility, and legitimacy of the electoral process.[38] The program functions as a collaborative network that integrates Electoral Justice entities internally through a Management Group, a Strategic Committee dedicated to combating disinformation, and an Analysis and Monitoring Group. Externally, it engages with media outlets, internet service providers, political organisations, public entities, technology firms, and academic institutions.[39] Further, several major social media platforms, including Twitter, TikTok, Facebook, WhatsApp, Google, Instagram, YouTube, and Kwai, have signed individual agreements with the TSE outlining measures to combat the spread of false and misleading information. These agreements, part of the TSE's efforts against disinformation, do not involve financial resources.

Notably, Google and YouTube have committed to providing reliable information about the electoral process, developing educational programs in collaboration with TSE. Facebook and Instagram will introduce features such as labels that guide users to official information and a chatbot to help voters access relevant election-related content. WhatsApp, which has been involved in previous electoral controversies, plans to enhance its chatbot capabilities and conduct training seminars for TSE personnel. Twitter will introduce search prompts for election-related information and prioritise tweets from TSE and fact-checking agencies. TikTok has pledged to create a dedicated page for trustworthy electoral content and

---

[35] Online Safety Act 2023, c. 50 (U.K.).

[36] *UK: Online Safety Bill risks undermining privacy around the world*, Article 19 (Sept. 5, 2023), https://www.article19.org/resources/uk-online-safety-bill-risks-undermining-privacy-around-the-world/.

[37] Spring, M. (2024, July 8). *Marianna Spring: This wasn't the social media election everyone expected*. https://www.bbc.com/news/articles/cj50qjy9g7ro.

[38] *Superior Electoral Court*. (n.d.). Plone Site. https://international.tse.jus.br/en/misinformation-and-fake-news/brazil-electoral-justice-permanent-program-on-countering-disinformation.

[39] *Superior Electoral Court*. (n.d.-b). Plone Site. https://international.tse.jus.br/en/misinformation-and-fake-news/brazil-electoral-justice-permanent-program-on-countering-disinformation.

establish a channel for reporting disinformation, while Kwai will also promote reliable information and support educational initiatives.[40]

C. IFES Guidelines for Social Media Manipulation

This part will discuss the International Foundation for Electoral Systems' (IFES) observations on social media interference in elections and outline the various measures taken by certain countries across the globe to tackle the issue

a. IFES Observations on Elections in the Information Age

IFES has, *inter alia*, made the following observations regarding social media interference:[41]

i. Defining and Addressing Disinformation:

- *Definition*: There is no internationally agreed-upon legal definition of disinformation, making it challenging for democracies to mitigate its harms while upholding freedom of expression. IFES notes that definitions of disinformation and misinformation vary across jurisdictions.[42]
- *Intentionality*: Disinformation always involves intentionality, where actors spread information deliberately to cause harm.[43]

ii. Threats to Electoral Bodies:

- *Entanglement with Disinformation Campaigns*: Electoral Management Bodies (EMBs), judges, and the judiciary can become central targets in disinformation campaigns before, during, and after court proceedings.
- *Frivolous Cases and Direct Attacks*: Disinformation campaigns often involve frivolous cases targeting courts and judges, necessitating sanctions against such actions to deter them.

iii. Judicial Intervention

- *Timely and Effective Responses*: Rapid adjudication of post-election cases, swift addressing of disinformation issues, and collaboration with social media platforms are crucial. Publicising judicial decisions widely can help neutralise disinformation efforts.

---

[40] Mari, A. (2022, February 17). Social networks partner with Brazil's electoral justice to tackle fake news during elections. *ZDNET*. https://www.zdnet.com/article/social-networks-partner-with-brazils-electoral-justice-to-tackle-fake-news-during-elections/.

[41] Rozumiłowicz, Dr. B. M., & Kužel, R. (2019). Social Media, Disinformation and Electoral Integrity. *IFES Working Paper*, 17–20. https://www.eods.eu/library/IFES_2019_SocialMediaDisinfromationElectoralIntegrity.pdf; Richard Nash, Jordan Shipley, and Typhaine Roblot, Lessons on Disinformation and Election Disputes Election Case Law Analysis Series, *IFES*, https://www.ifes.org/publications/election-case-law-analysis-series-lessons-disinformation-and-election-disputes; Alexandra Brown, Lisa Reppell, Patrick Quimby, and Typhaine Roblot, Lessons for Regulating Campaigning on Social Media Election Case Law Analysis Series, *IFES*, https://www.ifes.org/publications/ifes-election-case-law-analysis-series-lessons-use-technology-elections.

[42] Richard Nash, Jordan Shipley, and Typhaine Roblot, Lessons on Disinformation and Election Disputes Election Case Law Analysis Series, *IFES*, https://www.ifes.org/publications/election-case-law-analysis-series-lessons-disinformation-and-election-disputes.

[43] Mark Wilson, Lessons on Disinformation and Election Disputes Election Case Law Analysis Series, *IFES*, https://www.ifes.org/Election-Case-Law-Analysis-Series/Lessons-on-Disinformation-and-Election-Disputes/what-do-we-mean-disinformation#:~:text=Disinformation%20and%20misinformation%20are%20distinct,can%20reach%20a%20significant%20audience.

- *Summary Judgments*: Rapid and summary judgments can be effective in combating disinformation.

iv. Collaboration to Combat Disinformation:

- *Strategic and Innovative Practices*: Drawing from a range of practices related to elections, there are opportunities to enhance global and regional dissemination of these lessons through networks of judges and civil society organisations.

v. Communication Strategies for Judiciary (EMBs):

- *Countering Disinformation*: Courts must implement communication strategies during elections to counter attacks against judges and provide training on crisis communication and digital tools.[44]
- *Information Dissemination by EMBs*: Ensuring maximum information dissemination to all stakeholders before an election is critical. These strategic reforms could range from inviting cameras into the rooms, training and capacity building, enhanced cooperation with media, dialogue with political parties, partnerships with social media platforms to strengthen existing evidentiary procedure, etc.[45]

vi. Campaign Regulations and Social Media:

- *Beyond Traditional Media*: Campaign regulations must extend beyond traditional media, recognizing significant differences between social media and traditional outlets.[46]
- *Equitable Playing Field*: Legislators and regulators must create rules ensuring fairness for all candidates, considering the lack of editorial processes and new actors like citizens and media influencers.
- *Proportional Remedies*: Proportional remedies and nuanced provisions are essential to avoid overly broad restrictions on speech.
- *Understanding Social Media*: Courts must understand how social media platforms operate, and training, education, and resources can help judges identify online campaign violations.

vii. Measures to Combat Disinformation:

- *Regulatory Framework*: Laws regulating social media and disinformation should be consistent with international standards and not infringe on fundamental rights like freedom of speech and expression.[47]
- *Information Integrity*: Governments can mandate or amend curricula to educate voters and students about information integrity threats and media literacy.

---

[44] IFES reiterates the observation by the US's National Center for State Courts (2023) to identify four new specific themes here: "(i) The justice system ignores voting irregularities and fraud, allowing elections to be stolen from certain candidates (ii) The justice system tips the electoral map in favor of a particular party (iii) The justice system is unaccountable. Therefore, judges should be subject to threats of violence to keep them in line (iv) Decisions by the court are political and can be leaked for political purposes."

[45] Richard Nash, Jordan Shipley, and Typhaine Roblot, Lessons on Disinformation and Election Disputes Election Case Law Analysis Series, *IFES*, https://www.ifes.org/publications/election-case-law-analysis-series-lessons-disinformation-and-election-disputes.

[46] Alexandra Brown, Lisa Reppell, Patrick Quimby, and Typhaine Roblot, Lessons for Regulating Campaigning on Social Media Election Case Law Analysis Series, *IFES*, https://www.ifes.org/publications/ifes-election-case-law-analysis-series-lessons-use-technology-elections.

[47] Rozumiłowicz, B. M., & Kužel, R. (n.d.). Social Media, Disinformation and Electoral Integrity. *IFES Working Paper*, 16. https://www.eods.eu/library/IFES_2019_SocialMediaDisinfromationElectoralIntegrity.pdf.

- *Accountability and Proactive Steps*: Regulation provides accountability while self-regulation allows platforms to address disinformation proactively. This includes industry-led initiatives and government oversight.
- *Fact-Checking Initiatives*: (i) Civil society organisations, independent media, and trained journalists play a crucial role in investigating misleading content and providing transparent fact-checking processes. (ii) Fact-checking capacities must be independent, impartial, and transparent, including clear explanations of their processes and ensuring accessible findings.
- *Role of EMBs*: EMBs should develop strategies to counter disinformation, including fact-checking, social media monitoring, and public awareness campaigns. Ensuring necessary resources and capacity for EMBs is essential.

b. IFES Voluntary Election Integrity Guidelines for Technology Companies

The Voluntary Election Integrity Guidelines for Technology Companies, published by IFES, aim to establish foundational practices for technology companies to enhance election integrity and provide reliable information to voters.[48] These guidelines, applicable to various technology sectors, emphasise prioritising resources for global elections based on democratic principles and human rights,[49] engaging with civil society,[50] and creating transparent policies related to election content and activities[51]. They also stress centralising information for election authorities,[52] ensuring access to authoritative election data,[53] combating misinformation,[54] providing communication channels for election authorities,[55] and disclosing information on paid political content[56]. Post-election, companies are encouraged to maintain coordination mechanisms[57] and support analyses by election stakeholders[58]. These guidelines are designed to be adaptable to different contexts and are expected to be refined based on practical experiences.

The Voluntary Guidelines emphasise consulting with global civil society organisations to understand electoral contexts,[59] and aligns with the EU Digital Services Act's designation of civil society as trusted flaggers for identifying illegal content.[60] Both frameworks stress clear policies on election content and authoritative information,[61] paralleling the DSA's requirements for transparency and user rights protection.[62] Addressing misinformation and disinformation is crucial in both, with the guidelines

[48] IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
available at https://electionsandtech.org/election-integrity-guidelines-for-tech-companies/.
[49] Commitment 1, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[50] Commitment 2, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[51] Commitment 3, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[52] Commitment 4 and 5, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[53] Commitment 6, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[54] Commitment 7, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[55] Commitment 8, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[56] Commitment 9, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[57] Commitment 10, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[58] Commitment 11, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[59] Commitment 2, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[60] Article 19, DSA.
[61] Commitments 3 and 6, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[62] Articles 13 and 17, DSA.

advocating strategies to manage such content[63] and the DSA mandating risk mitigation measures.[64] Additionally, the need for effective communication channels with election authorities[65] is reflected in the DSA's provisions for prompt handling of notices and cooperation with authorities.[66]

The Voluntary Guidelines emphasise clear policies on election-related content and access to authoritative information,[67] and align with the Indian IT Rules' requirement for intermediaries to publish their content removal and user data protection policies.[68] Both frameworks also address misinformation, with the guidelines calling for strategies to manage it[69] and the IT Rules mandating the removal of unlawful content and the use of technology to detect misinformation.[70] Additionally, the guidelines and the IT Rules both stress the importance of communication channels with election authorities, with the IT Rules requiring social media intermediaries to appoint a nodal contact for coordination with authorities,[71] reflecting the guidelines' emphasis on effective communication.[72]

[63] Commitment 7, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[64] Article 26, DSA.
[65] Commitment 8, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[66] Article 12, DSA.
[67] Commitments 3 and 6, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[68] Rule 3(1)(a) and (b), 2021
[69] Commitment 7, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.
[70] Sec. 79 IT Act and Rule 4(4) IT Rules, 2021
[71] Rule 4(1)(b) IT Rules, 2021
[72] Commitment 8, IFES. Voluntary Election Integrity Guidelines for Technology Companies, version 1.0.

Part III: Digital Deception and mitigation measures in Indian Elections

This section explores the Election Commission of India's (ECI) enforcement of the Model Code of Conduct (MCC) and its authority to manage social media challenges during elections. It highlights how the ECI navigates its regulatory roles, particularly in the context of modern digital campaigning.

The ECI, a constitutional body, is entrusted with the superintendence, direction, and control of elections in India.[73] Alongside multiple laws such as the Representation of the People Act (RP Act), 1951, and the Bhartiya Nyaya Sanhita 2024 (earlier one Indian Penal Code, 1860), which allow it to penalise electoral offences and corrupt practices, the ECI has also developed various guidelines and advisories over the years. One such important guideline is the Model Code of Conduct (MCC), a set of rules that outlines the minimum standards of behaviour for political parties and candidates.[74] Although the MCC does not carry the force of law, the ECI can prosecute political parties and candidates under relevant penal laws as well.[75] Notably, while the RP Act, 1951, addresses electoral offences and corrupt practices that can be pursued even after the elections, the MCC is applicable immediately after the announcement of elections and remains in effect until the end of the electoral process. Thus, the MCC fills a crucial 'legal vacuum' and acts as a set of moral principles guiding political conduct during elections.[76]

In Indian context, media exposure significantly influences electoral choices, with voters generally favouring parties that effectively harness social media's capabilities.[77] As internet and social media use has increased, major political parties have made social media a primary communication channel with the public, significantly boosting their investment in online campaigns.[78]

Over the last decade, political parties and individuals have actively utilised and sometimes weaponised social media platforms for political gain.[79] This trend has become more pronounced with the rapid surge in Indian internet users which has grown by over 20% year on year in the last ten years.[80] Research identifies five primary methods of social media manipulation: (i) Creating fake accounts to amplify certain messages or distort social media metrics, (ii) Messaging and valence, which includes promoting pro-party messages and launching attacks on the opposition, (iii) Generating content that influences or misleads

---

[73] B. L Fadia, *Reforming the Election Commission*, 53 The Indian Journal of Political Science 78-88. (1992), http://www.jstor.org/stable/41855597; R Ramesh, *Historical Perspectives of the Electoral Reforms in India*, 72 Proceedings of the Indian History Congress 1325-1336 (2011), http://www.jstor.org/stable/4414574; Ramaswamy R Iyer, *The Election Commission and the Judgment*, 31 Economic and Political Weekly 37-42 (1996), http://www.jstor.org/stable/4403648.

[74] Election Commission of India, Model Code of Conduct for the Guidance of Political Parties and Candidates, available at https://www.eci.gov.in/mcc/.

[75] Manjari Katju, *Mass Politics and Institutional Restraint: Political Parties and the Election Commission of India*, 4 Lokniti 37-42 (1996), https://doi.org/10.1177/2321023016634945.

[76] *Id.*

[77] *Shreyas Sardesai, Media Exposure and Vote Choice in India, 1996–2019, 11 Studies in Indian Politics 317-334 (2023),* https://doi.org/10.1177/23210230231203795.

[78] Aakash Shaw, *Role of Social Media in Social Mobilization*, 7 Global Media Journal (2013), https://www.caluniv.ac.in/global-mdia-journal/COMMENT-2016-NOV/C-5-F.pdf.

[79] Qureshi, W. A., 2019. The Militarization of Social Media. *U. Haw. L. Rev.*, 42, 169.

[80] Ministry of External Affairs, India had over 700 mn active internet users by Dec '22: Report (2022), available at https://indbiz.gov.in/india-had-over-700-mn-active-internet-users-by-dec-22-report/.

public opinion on social media, (iv) Targeted advertising, which uses user data to reach specific demographics with tailored political messages, (v) Exploiting social media platforms' features to spread these messages and advertisements effectively.[81]

The 2014 Lok Sabha General Elections marked a pivotal turning point in India, as it was the first election where social media emerged as a significant platform for political communication and engagement.[82] Political parties and candidates utilised platforms like Facebook and Twitter to target younger voters, disseminate their agendas, and shape public opinion.[83] Building on the digital advancements of the 2014 elections, the 2019 Lok Sabha General Elections saw an intensified use of social media for political campaigning. Political parties not only expanded their social media presence but also increasingly employed misinformation and disinformation tactics to enhance their political images and garner votes. The proliferation of websites, media houses, and online news portals have enhanced misinformation and disinformation, which has become immortal on the internet and is difficult to track, flag, and remove.[84] The creation of private groups and the mobilisation of 'cyber troops' — including volunteer networks, private companies, and social media influencers — highlighted a sophisticated and strategic approach to targeting and customising messaging campaigns.[85] This shift underscored the evolving landscape of digital political communication, emphasising the increased role of social media in shaping electoral outcomes amidst concerns about the spread of false information. Despite voluntary protocols like the IAMAI's "Voluntary Code of Ethics for the General Election 2019,"[86] developed with major social media platforms, the voluntary nature falls short of holding signatories accountable.

The challenges of misinformation, disinformation, and fake news, which became particularly prominent after the 2019 general elections, continue to escalate. The online political sphere has turned more complex and pervasive since then. During the COVID-19 pandemic, the spread of such false information not only posed serious threats to India's public healthcare system but also undermined efforts to combat the virus.

---

[81] Abdul Fahad & Ezaleila Mustafa, *Religious-Political Discussion on Instagram and WhatsApp and Perception of Religion Among Youths in Delhi*, Journal of Asian and African Studies (2023), http://dx.doi.org/10.1177/00219096231200592.

[82] Ranganathan, M., 2014. Indian Elections, 2014: Commercial Media Pushes Social Media into Focus. *Asia Pacific Media Educator*, 24(1), 23-38; Vishal Sharma, *Young India, Social Networking Sites & Indian Politics*, 73 The Indian Journal of Political Science 149-154 (2012), https://www.jstor.org/stable/41856570.

[83] Narasimhamurthy, N., 2014. Use and Rise of Social Media as Election Campaign Medium in India. *International Journal of Interdisciplinary and Multidisciplinary Studies*, 1(8), 202-209; Ahmed et al., *Leveling the Playing Field: The Use of Twitter by Politicians During the 2014 Indian General Election Campaign*, 34 Telematics and Informatics (2017), https://escholarship.org/content/qt9nb592x4/qt9nb592x4.pdf?t=pakrzb; Ahmed et al., *The 2014 Indian Elections on Twitter: A Comparison of Campaign Strategies of Political Parties*, 33 Telematics and Informatics (2016), https://doi.org/10.1016/j.tele.2016.03.002; Safiullah, Md & Pathak, Pramod & Singh, Saumya. (2022). The impact of social media and news media on political marketing: an empirical study of 2014 Indian General Election. International Journal of Business Excellence. 26. 536-550. 10.1504/IJBEX.2022.122765.

[84] Sangeeta Mahapatra & Johannes Plagemann, *Polarisation and Politicisation: The Social Media Strategies of Indian Political Parties*, 3 GIGA Focus (2019), http://www.jstor.org/stable/resrep24806.

[85] Billy Perrigo, How Volunteers for India's Ruling Party Are Using WhatsApp to Fuel Fake News Ahead of Elections, TIME (Jan. 25, 2019), https://time.com/5512032/whatsapp-india-election-2019/ ; Ualan Campbell-Smith & Samantha Bradshaw, GLOBAL CYBER TROOPS COUNTRY PROFILE: INDIA(2024), https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2019/05/India-Profile.pdf; Vindu Goel, *In India, Facebook's WhatsApp Plays Central Role in Elections*, The New York Times (May 14, 2018), https://www.nytimes.com/2018/05/14/technology/whatsapp-india-elections.html.

[86] Press Information Bureau, Government of India, IAMAI-ECI Voluntary Code of Ethics,, available at https://static.pib.gov.in/WriteReadData/userfiles/IAMAI-ECI%20VCE.pdf.

Such incidents have clarified the harmful effects of misinformation and disinformation on social media platforms.

Regardless of whether political parties sanction misinformation and disinformation campaigns, these tactics have undermined meaningful public discourse on politics in India. Notably, in response to these challenges, political parties have significantly increased their budgets for social media campaigning to enhance their outreach and influence effectively and engage more directly with voters.[87]

Moreover, the 2024 Freedom House assessment, which evaluates the level of political rights and civil liberties in various regions, has specifically identified online disinformation as a critical issue that impedes freedom of expression and belief.[88] The World Economic Forum's Global Risk Report 2024 highlighted that the unprecedented proliferation of misinformation and disinformation in India poses a significant threat.[89] Such trends emphasise the need for robust strategies to uphold the integrity of elections, and preserve democratic processes.

Although the ECI has a constitutional mandate, there is a lack of a legislative framework that empowers the ECI to directly regulate online disinformation and misinformation before, during, or after elections. Presently, the ECI does not have a statutory authority to intervene for the purposes of elections and demand a takedown of any impugned content. The ECI is not directly a part of the takedown notice mechanism in India, which is primarily operationalised via the Information Technology Act, 2000 and the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021. These two legal instruments empower the government or the courts to either block public access to any information if it is deemed necessary,[90] or to issue takedown notices to intermediaries, including social media intermediaries.[91]

Further, the Model Code of Conduct (MCC) does not specifically address the regulation of social media during elections and primarily targets traditional forms of political campaigning, however, the general principles of the MCC are adaptable and can extend to digital platforms.[92] Under the MCC's framework, the ECI can direct stakeholders, including social media intermediaries, to promote responsible campaigning and regulate online misinformation and disinformation.[93] These directives, while impactful,

---

[87] Kumari, S. (2024, June 20). *Behind the Scenes: Digital Ad Spend Soared 10x in Lok Sabha Elections 2024.* The Quint. https://www.thequint.com/elections/digital-advertisement-spending-lok-sabha-elections-2024-political-parties.

[88] Freedom House, *India: Freedom in the World 2024 Country Report*, Freedom House, https://freedomhouse.org/country/india/freedom-world/2024.

[89] *Global Risks 2024: At a Turning Point*, World Economic Forum (2024), https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf.

[90] *See* section 69A, IT Act 2000.

[91] Rule 3(1)(d), IT Rules 2021.

[92] Economic Times, Model Code, Political Ad Rules Will Apply to Social Media Too (2019), available at https://economictimes.indiatimes.com/news/elections/lok-sabha/india/model-code-political-ad-rules-will-apply-to-social-media-too/articleshow/68350634.cms?from=mdr.

[93] For example, the ECI recently instructed the Ministry of Electronics and Information Technology to halt the dissemination of "Viksit Bharat Sampark" messages through a popular messaging app. These messages, which endorsed the policies of the ruling government, were deemed inappropriate during the period when the MCC was in effect, following multiple public complaints. HT News Desk, *Election Commission tells government to stop sending 'Viksit Bharat' messages on WhatsApp*, Hindustan Times (Mar. 21,

lack binding legal authority and are enforceable only if the offences fall within the provisions of the RP Act, 1951 or penal laws. Social media intermediaries, therefore, are not legally obligated to follow the ECI's directives. Furthermore, since MCC includes only political parties and candidates, it does not directly regulate the social media platforms.

The ECI introduced several measures for the 2024 general elections to educate the general public about social misinformation and disinformation. One notable measure was the release of the "Myth v. Reality" register in April 2024, during the election process, to educate voters and ensure that voters have access to accurate and verified information throughout the electoral process.[94] The register is designed in a user-friendly format, addressing myths and misinformation concerning (i) Electronic Voting Machine (EVM)/VVPAT, (ii) Electoral Roll/Voter Services, (iii) the conduct of elections, and related topics. It provides stakeholders with previously debunked election-related false information, probable myths circulating on social media platforms, FAQs on significant issues, and reference materials under various sections.[95]

Despite its efforts, the volume and nature of cases addressed via the register remain inadequate and most cases may have escaped attention and mitigation. The ECI has managed to tackle only a limited number of social media manipulation incidents. Considering India's population of approximately 1.4 billion, this limited response is insufficient to effectively combat the extensive spread of misinformation and disinformation in the country. The ECI has only (i) 'busted' 17 disinformation news regarding EVMs, 'busted' 23 myths about EVMs; answered 120 FAQs regarding EVMs; (ii) 'busted' 4 disinformation news regarding electoral roll/voter services, 'busted' 4 myths about electoral roll/voter services, answered 56 FAQs regarding electoral roll/voter services; (iii) 'busted' 17 disinformation news regarding conduct of elections, 'busted' 13 myths about conduct of elections, answered 244 FAQs regarding conduct of elections.[96]

The information regarding number of posts taken down by social media platforms voluntarily during election process is neither available in public domain nor on ECI's website. The ECI in collaboration with the concerned statutory authority under the IT Act should direct the social media platforms to publish such information in public domain.

Pending legislative empowerment of ECI and going forward for future elections, ECI should develop a tripartite collaborative framework amongst ECI, Tech companies, and the Ministry of Electronics and Information Technology (MeitY), for devising efficient mechanisms for the expeditious removal of election-related misinformation and deepfakes. This collaboration must be informed by structured discussions aimed at identifying effective solutions to mitigate the proliferation and impact of such content during electoral periods.

2024), https://www.hindustantimes.com/india-news/election-commission-to-it-ministry-stop-sending-viksit-bharat-messages-on-whatsapp-101711006490551.html.

[94] *ECI introduces 'Myth vs Reality Register' to proactively combat mis-information in General Elections 2024*. PRESS INFORMATION BUREAU (April, 2024). https://pib.gov.in/PressReleaseIframePage.aspx?PRID=2016941#:~:text=The%20'Myth%20vs%20Reality%20Register'%20serves%20as%20a%20comprehensive%20repository,them%20to%20make%20informed%20decisions..

[95] Id.

[96] *Myth v. Reality Register*. ELECTION COMMISSION OF INDIA. https://mythvsreality.eci.gov.in/index.

Part IV. The AI Paradigm Shift: Deepfakes and Beyond

This part delves into the rapidly evolving domain of AI, focusing on deepfakes, which represent a new challenge in digital deception. The first part introduces the problem, exploring how deepfakes are created to deceive, and their implications for misinformation campaigns. The second part examines the potential impacts of these technologies on the elections across the world in 2024 highlighting the challenges they pose to democratic processes and electoral integrity.

A. Deepfakes: The New Frontier of Digital Deception

Unlike other synthetic media designed for entertainment or education,[97] deepfakes are crafted specifically to deceive, making them powerful tools for spreading misinformation.[98] By training AI on extensive datasets of real clips, deepfakes replicate an individual's appearance and voice with alarming accuracy, producing content where people appear to say or do things they never did. This blurs the line between reality and fiction and has been primarily used to misrepresent political figures or celebrities, potentially swaying public opinion or sparking outrage.[99] Their capacity to bypass traditional detection cues, such as those visible in amateur photoshopping, poses a significant challenge in the era of social media, where the rapid dissemination of information outstrips the slower pace of authenticity verification.[100]

B. Impact of Deepfakes on Elections

In the 2024 elections globally, the influence of deepfakes became a critical concern, whereby such sophisticated digital fabrications posed significant impact electoral outcomes by crafting narratives that can quickly spread across social media platforms, reaching vast audiences before the authenticity of the content can be verified.[101] The deployment of deepfakes in election campaigns can be particularly disruptive. For example, in Gabon in 2019, a suspected deepfake video of President Ali Bongo giving a New Year's speech fueled doubts about his fitness and whereabouts, contributing to a brief attempted coup.[102] In another instance, a Belgian political party created a deepfake video in 2018 of the Belgian Prime Minister announcing drastic climate actions, which was actually a campaign to raise climate change

[97] Javahir Askari, *Deepfakes and Synthetic Media: What are they and how are techUK members taking steps to tackle misinformation and fraud*, techUK (June 18, 2023), https://www.techuk.org/resource/synthetic-media-what-are-they-and-how-are-techuk-members-taking-steps-to-tackle-misinformation-and-fraud.html.

[98] Rosa Gil et al., *Deepfakes: Evolution and Trends*, 27 Soft Computing 11295–11318 (2023), https://doi.org/10.1007/s00500-023-08605-y.

[99] Faragó, T., 2019. Deep Fakes–An Emerging Risk to Individuals and Societies Alike.

[100] Stuart A Thompson, *Making Deepfakes Gets Cheaper and Easier Thanks to A.I.*, The New York Times (Mar. 12, 2023), https://www.nytimes.com/2023/03/12/technology/deepfakes-cheapfakes-videos-ai.html.

[101] Mira Patel, *How big is the threat of artificial intelligence over elections*, The Indian Express (Mar. 2, 2023), https://indianexpress.com/article/research/how-big-is-the-threat-of-artificial-intelligence-over-elections-9188317/.

[102] Cahlan, S., 2020. How a Sick President & Suspect Video Helped Spark an Attempted Coup in Gabon, *The Washington Post* (Feb. 13, 2020), available at https://www.washingtonpost.com/politics/2020/02/13/how-sick-president-suspect-video-helped-sparked-an-attempted-coup-gabon/.

awareness.[103] Such uses demonstrate how political operatives or malicious actors could release deepfake videos that show candidates engaging in behaviour that could damage their reputations or alienate their support base. Even if these videos are later debunked, the initial impact and lingering doubt could irreparably harm a candidate's image and reduce their chances of securing votes.[104] Furthermore, deepfakes could be used to exacerbate societal divisions, inflaming tensions on topics like race, immigration, or public policy by simulating offensive or polarising statements by leaders.[105] Further, the psychological impact of seeing a "video" of a candidate saying or doing something controversial can be much more powerful than reading about their alleged actions. Visual misinformation has been shown to be more memorable than textual misinformation, which could affect how people vote, especially when the authenticity of the content is not immediately questioned or when the fake content aligns with a voter's pre-existing biases.[106]

C. Global Policy Response to Deepfakes

Countries have begun to recognise the threat posed by deepfakes and are taking steps to address it.[107] In the United States, there are both federal and state mandates aimed at regulating the creation and distribution of deepfakes, especially with regard to elections. Legislation like the DeepfakesAccountability Act has been proposed to criminalise the malicious creation and distribution of deepfake content.[108] Similarly, several states have enacted laws specifically addressing election-related deepfakes, setting legal precedents for punishing those seeking to influence elections through digital deception.[109]

In India, advisories have been issued to raise awareness about the issue, and discussions are ongoing about integrating guidelines on digital content into existing laws to safeguard elections from misinformation. Despite these efforts, the rapidly evolving nature of deepfake technology continues to pose significant challenges. The arms race between deepfake creation and detection technologies means that as soon as new detection methods are developed, newer methods to evade these detections are also being devised.

---

[103] The Brussels Times, XR Belgium Posts Deepfake of Belgian Premier Linking COVID-19 with Climate Crisis (Apr. 14, 2020), available at https://www.brusselstimes.com/106320/xr-belgium-posts-deepfake-of-belgian-premier-linking-covid-19-with-climate-crisis.

[104] Janna Anderson & Lee Rainie, *As AI Spreads, Experts Predict the Best and Worst Changes in Digital Life by 2035*, Pew Research Center, https://www.pewresearch.org/internet/wp-content/uploads/sites/9/2023/06/PI_2023.06.21_Best-Worst-Digital-Life_2035_FINAL.pdf.

[105] Janna Anderson & Lee Rainie, *Deepfakes and international conflict*, Brookings https://www.brookings.edu/articles/deepfakes-and-international-conflict/.

[106] Myrto Pantazi et al., *Social and Cognitive Aspects of the Vulnerability to Political Misinformation*, 42 Political Psychology (2021), https://doi.org/10.1111/pops.12797; Sameer Patil & Shourya Gori, *Deep fake, disinformation, and deception*, The Observer Research Foundation (June 25, 2023), https://www.orfonline.org/expert-speak/deep-fake-disinformation-and-deception.

[107] Charlotte Stanton, *How Should Countries Tackle Deepfakes*, Carnegie Endowment (Jan. 28, 2019), https://carnegieendowment.org/2019/01/28/how-should-countries-tackle-deepfakes-pub-78221.

[108] *The US is drafting new laws to protect against AI-generated deepfakes*, World Economic Forum (Feb. 27, 2024), https://www.weforum.org/agenda/2024/02/ai-deepfakes-legislation-trust/.

[109] *The US is drafting new laws to protect against AI-generated deepfakes*, World Economic Forum (Feb. 27, 2024), https://www.weforum.org/agenda/2024/02/ai-deepfakes-legislation-trust/.

This dynamic makes it incredibly challenging to create long-term, effective countermeasures that are impactful and sustainable.

Part V: Recommendations

As social media continues to shape electoral landscapes and technologies like deepfakes evolve, it is imperative to adopt a strategic and collaborative approach to mitigate risks and ensure the integrity of democratic processes. This part provides targeted recommendations for various stakeholders, including the ECI, technology companies, and government bodies, and suggests a multistakeholder approach for broader collaborative efforts.

A. EMBs

1. **Enhance Monitoring and Response:** Establish specialised units dedicated to monitoring social media activities, especially during elections. These units should consist of experts in digital media, cybersecurity, and electoral law, tasked with identifying and addressing instances of misinformation, disinformation, and deepfakes swiftly. Comprehensive enhancements in financial resources and structural reforms are essential. The Commission's monitoring capabilities should be strengthened. This support should encompass advanced technological tools and comprehensive training for staff, enabling them to effectively manage digital threats during electoral processes

2. **Partnerships with Civil Society and Fact-Checkers:** Engage actively with civil society organisations and professional independent fact-checking groups to create a broad network of stakeholders committed to maintaining the integrity of information during elections. These collaborations can enhance the EMB's capabilities in monitoring content, verifying facts, and spreading accurate information.

3. **Enhance Literacy and Awareness: As an example,** the ECI has taken an important step with initiatives like the 'Myth v. Reality' to counter disinformation and implementation of voter education program. Building on these efforts is crucial, as enhancing and expanding literacy initiatives will equip voters with the critical skills necessary to navigate the complex information landscape effectively and safeguard the integrity of the democratic process.

4. **Develop a Reputation Management Strategy:** Develop a comprehensive strategy to inspire public trust in the EMB, especially considering the spread of false information about the incumbents, the EMB, electoral processes and their management. While initiatives such as c-Vigil[110] and Myth v. Reality Register launched by ECI helps to some extent, it would be crucial to develop a comprehensive real time communication strategy for the same; the Australian Electoral Commission has established an exemplary, holistic approach towards restoring public trust.[111]

B. Technology Companies

[110] It is an online application by the ECI for election information dissemination and grievance redressal for MCC violations by citizens.

[111] Australian Electoral Commission. The AEC Reputation Management System. https://www.aec.gov.au/About_AEC/reputation-management.htm.

1. **Advance Detection Technologies:** Continue to invest in the development of advanced detection technologies that can identify deepfakes. This commitment should include ongoing research to stay ahead of rapidly evolving misinformation tactics.

2. **Enhance Transparency and Reporting:** Increase transparency regarding the efforts to combat misinformation and routinely publish reports detailing the effectiveness of these measures, especially during election periods. This should include statistics on the volume of content reviewed, the accuracy of detection algorithms, and the speed of response actions.

3. **Clearly Label Synthetic Media:** Automate the process of labelling content that has been digitally altered or generated. Users should be clearly informed when the media they are viewing may have been modified, helping them to make better-informed decisions about the content's trustworthiness.

C. Government

1. **Empower the EMB:** To bolster the EMB's ability to safeguard the integrity of elections against digital threats, appropriate legislative framework should be framed empowering the EMB to counter the adverse impact of misinformation etc. on the integrity of elections particularly by vesting authority in EMB to issue take down notices to social media platforms during election process.

2. **Invest in Public Education:** Allocate funds for public education programs focused on digital literacy, which would equip citizens with the skills to identify misinformation and understand the implications of synthetic media. These initiatives are essential for empowering voters to make informed decisions.

3. **Strengthen Inter-agencies Collaboration:** Promote enhanced cooperation between different governmental bodies, EMB, and law enforcement agencies. This collaboration should aim to create a unified strategy for addressing and mitigating digital threats to electoral integrity.