

Snapshot

GenAI and Electoral Disinformation

The Case of the Philippines

Maria Elize H. Mendoza, Aries A. Arugay

SAIL Snapshots is a peer-reviewed article series featuring key research findings from Safer Internet Lab researchers and its associates. The views presented are exclusively those of the authors and do not reflect the official stance of SAIL, CSIS, Google, or any other affiliated organisation.

GenAI and Electoral Disinformation

The Case of the Philippines

Maria Elize H. Mendoza¹, Aries A. Arugay²

Editor: Noory Okthariza³

INTRODUCTION

The Philippines will hold its next general elections on 12 May 2025. In September 2024, the Commission on Elections (COMELEC) issued Resolution No. 11064⁴ which seeks to regulate the use of social media, specifically Generative Artificial Intelligence (GenAI), in the upcoming elections. The long absence of a legal regulatory framework to counter disinformation in Philippine elections has exacerbated a problem that has plagued the country since 2016, capturing three electoral cycles, including the 2016 and 2022 presidential elections. This report presents an overview of the disinformation problem in the Philippines and its implications in the upcoming 2025 elections, regulatory steps being undertaken to address this issue, and the challenges in countering disinformation and the use of GenAI in elections.

¹ Assistant Professor, Department of Political Science, University of the Philippines-Diliman

² Chairperson, Department of Political Science, University of the Philippines-Diliman

³ Researcher, Department of Politics and Social Change, CSIS Indonesia

⁴ COMELEC Resolution No. 11064: Guidelines on the Use of Social Media, Artificial Intelligence, and Internet Technology, for the Digital Election Campaign, and the Prohibition and Punishment of its Misuse for Disinformation, and Misinformation, in connection with the 2025 National and Local Elections and the BARMM Parliamentary Elections

The 2016 Philippine presidential elections were known as the country's first "social media elections" (Sinpeng et al., 2020). The electoral victory of strongman and populist Rodrigo Duterte (2016-2022) was fueled by unprecedented levels of online disinformation orchestrated by "architects of networked disinformation" (Ong and Cabañes, 2018) and spread on platforms, especially Facebook and YouTube. Social media influencers and online "trolls" amplified online disinformation to justify Duterte's draconian policies, such as the deadly war on drugs. On Facebook, pro-Duterte pages circulated a photo of a rape incident supposedly committed by a drug addict to garner sentiments that drug addicts deserve to be killed. This was eventually proven false (Feldstein, 2021). In the 2019 midterm elections, massive online attacks against the political opposition led to the absolute defeat of the opposition in the senatorial race and a sweeping victory for pro-administration candidates (Chua and Soriano, 2020). In the 2022 presidential elections, the victory of Ferdinand Marcos Jr., son of the late dictator Marcos Sr., was largely attributed to the proliferation of online disinformation across different social media platforms intended to twist historical narratives, cleanse their family name, and further demonize the opposition (Mendoza, 2022). In 2021, *VERA Files* released an investigative report showing that Marcos Jr. was the biggest beneficiary of election-related disinformation on social media. On TikTok, Marcos supporters created and spread videos of the deceased dictator using fake manufactured audio. Moreover, a research report found that different candidates spent an overall amount of approximately 600 million to 1.5 billion Philippine pesos (approximately USD 10.9 million to 25 million) on political influencers during the 2022 campaign (Gaw et al., 2023).

Through the years, the Philippine government has failed to address the problem of electoral disinformation. Proposals to counter "fake news" and regulate social media campaigns lingered in the legislature. The burden of countering disinformation fell on the shoulders of civil society actors, particularly media groups and the academe, who launched fact-checking initiatives, digital literacy campaigns, and voter education programs. The impact of these commendable initiatives is

constrained by the lack of similar efforts, especially a strict legal framework, on the part of the government. The enactment of COMELEC Resolution No. 11064, which seeks to establish a formal collaboration network between the COMELEC and civil society actors, is a step towards the right direction. However, its belated nature presents huge challenges for its effective implementation in the upcoming 2025 elections.

The threat of misusing GenAI in the upcoming elections aggravates the existing yet unaddressed disinformation problem. In February 2022, a Social Weather Stations (SWS) survey found that the majority of Filipinos find it difficult to detect fake news. Results of a September 2022 survey by Pulse Asia show that Filipinos consider the Internet or social media as the top source of disinformation, mostly spread by social media influencers. These findings, together with the fact that Filipinos are among the top social media users in the world, are a dangerous combination. The incumbent administration of President Ferdinand Marcos Jr. was not spared: in April 2024, the government flagged a deepfake audio of the president allegedly ordering a military attack in light of the rising tensions between the Philippines and China. In July 2024, the government flagged another deepfake video of the president supposedly taking drugs in his younger years (De Leon, 2024). More of these kinds of videos are expected in the months leading to the May 2025 elections.

REGULATORY PREPAREDNESS: GAPS, INTERVENTIONS, AND LIMITATIONS

The sole key regulatory instrument specifically addressing the issue of disinformation and GenAI in the upcoming 2025 Philippine elections is COMELEC Resolution No. 11064. Enforced on 24 September 2024, the resolution aims to regulate the use of digital platforms and protect against the misuse of digital technology and AI. It does not completely prohibit the use of AI but provides parameters for its allowable use, which must be fully disclosed by relevant actors. Its main implementing

arm is the Task Force *Katotohanan, Katapatan, at Katarungan sa Halalan* [Truth, Honesty, and Justice in Elections] led by the Education and Information Department (EID) and the Law Department of COMELEC. This Task Force was formed to foster collaboration between deputized law enforcement agencies, relevant government instrumentalities, accredited citizens' arms, the legal community, and other non-partisan organizations in detecting the misuse of social media, AI, and other digital technology in the upcoming elections.

Major interventions of the resolution include the following: registration of social media accounts, websites, and digital campaign platforms of candidates, political parties, and their supporters; disclosure of the use of AI within its allowable parameters in election and campaign materials; and sanctions against prohibited acts, including the use of fake accounts and hiring of paid users or "trolls" to amplify campaign messages. As of writing, the provisions of the resolution, including its sanctions, have not been applied to any specific electoral case. The registration of social media accounts is ongoing until 13 December 2024, while the campaign period commences in February 2025.

On the other hand, civil society actors are consistent in their efforts to combat disinformation. The fact-checking initiatives of *VERA Files*, an independent media organization, are a continuation of its efforts since 2022. A similar initiative *#FactsFirstPH* is carried out by *MovePH*, the civic engagement arm of media outlet *Rappler*. Both *VERA Files* and *MovePH* engage local and international partners in these projects. Moreover, the accredited citizen arms of COMELEC such as the National Citizens' Movement for Free Elections (NAMFREL) and Parish Pastoral Council for Responsible Voting (PPCRV) have ongoing voter education programs. What is clearly lacking is the sustained engagement between these civil society groups and the government, particularly COMELEC. The formation of COMELEC's Task Force is long overdue. For the longest time, civil society groups have been operating without legal backing in the absence of formal government policies against disinformation.

The COMELEC resolution requires the registration of social media accounts in support of a particular candidate. However, the Philippine information ecosystem is saturated with “independent media practitioners” known as vloggers or influencers who are not necessarily formally affiliated with a candidate or political party. They exert significant influence when it comes to shaping public opinion compared to official campaign teams and political parties. Their online contents are extensively consumed by the general public, and there is evidence that they have been hired by politicians in previous elections in the Philippines (Gaw and Arugay, 2024). This presents a challenge for COMELEC and its partners in accounting for such users given their number. The domain of these vloggers and influencers remains largely unregulated given the lack of substantial action from tech platforms. Meta and TikTok have expressed their commitment to cooperate with the COMELEC in fighting disinformation and preventing the misuse of AI (Gutierrez, 2024), but concrete results remain to be seen.

CONCLUSIONS AND RECOMMENDATIONS

Electoral disinformation spread on social media and amplified by influencers and vloggers is a major issue in the upcoming 2025 Philippine elections. This is further complicated by the integration of AI technology in the creation and proliferation of misleading, fake, or malicious contents. COMELEC and civil society actors recognize the inadequacies of current policies and are pushing for the Philippine legislature to pass a social media regulation bill (De Leon, 2024). To this date, blatant fake contents from previous electoral cycles are still present on platforms such as Facebook and TikTok. Thus, proactive content moderation measures and accountability must be demanded from and exercised by social media platforms. The COMELEC resolution is expected to temporarily fill the gap left by the absence of formal legislation. Absent a formal law and institutional capability-building, current interventions will be highly limited. In the meantime, it is necessary for COMELEC to sustain its engagements with civil

society, demand full cooperation from tech platforms, and exhaust all means to ensure the credibility and integrity of the upcoming elections.

REFERENCES

- Chua, Y. & Soriano, J. (2020). Electoral Disinformation: Looking Through the Lens of Tsek.ph Fact Checks. *Plaridel*, 17(1), 285-295. <https://doi.org/10.52518/2020.17.1-10chusrn>
- De Leon, D. (2024, October 23). Lack of legal teeth vs online disinformation casts long shadow over 2025 polls. *Rappler*. <https://www.rappler.com/philippines/elections/lack-legal-teeth-against-online-disinformation-casts-long-shadow-2025/>
- Feldstein, S. (2021). Social Manipulation and Disinformation in the Philippines. In *The Rise of Digital Repression* (pp. 135-176). Oxford University Press.
- Gaw, F. & Arugay, A. (2024). Like, Subscribe and Vote: The Role of Political Influencers in the 2022 Philippine Elections and Beyond. *ISEAS Perspective* 2024/19. https://www.iseas.edu.sg/wp-content/uploads/2024/01/ISEAS_Perspective_2024_19.pdf
- Gaw, F., Bunquin, J. B., Cabbuag, S., Lanuza, J.M., Sapalo, N. & Yusoph, A. (2023). Political Economy of Covert Influence Operations in the 2022 Philippine Elections. *Internews*. https://internews.org/wp-content/uploads/2023/07/InternewsPH_July2023_Political-Economy-of-Covert-Influence-Operations-in-the-2022-Philippine-Elections-2.pdf
- Gutierrez, P. (2024, October 17). Comelec coordinating with social media firms vs AI, deepfakes. *ABS-CBN News*. <https://news.abs-cbn.com/news/2024/10/17/comelec-coordinating-with-social-media-firms-vs-ai-deepfakes-1611>

- Ong, J. & Cabañes, J.V. (2018). Architects of Networked Disinformation: Behind the Scenes of Troll Accounts and Fake News Production in the Philippines. *Newton Tech4Dev Network*.
<https://newtontechfordev.com/wp-content/uploads/2018/02/ARCHITECTS-OF-NETWORKED-DISINFORMATION-FULL-REPORT.pdf>
- Pulse Asia Research Inc. (2022, October 11). Pulse Asia Research's September 2022 Nationwide Survey on Fake News.
<https://pulseasia.ph/wp-content/uploads/2022/10/MR2-UB2022-3-MR-on-Fake-News-Final.pdf>
- Sinpeng, A., Gueorguiev, D., & Arugay, A. (2020). Strong Fans, Weak Campaigns: Social Media and Duterte in the 2016 Philippine Election. *Journal of East Asian Studies*, 20(3), 353-374.
<https://doi.org/10.1017/jea.2020.11>
- Social Weather Stations. (2022, February 25). Fourth Quarter 2021 Social Weather Survey: 69% of adult Filipinos say the problem of fake news in media is serious.
<https://www.sws.org.ph/swsmain/artcldisppage/?artcsyscode=ART-20220225130129>



Safer Internet Lab

 saferinternetlab.org

 Jl. Tanah Abang III no 23-27
Gambir, Jakarta Pusat, 10160

Find Us On



CSIS Indonesia | Safer Internet Lab