



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IX **Month of publication:** September 2026

DOI: <https://doi.org/10.22214/ijraset.2026.84936>

www.ijraset.com

Call: ☎ 08813907089

E-mail ID: ijraset@gmail.com

Ethical Issues in AI-Generated Political Content: Deepfakes, Softfakes, and Responsibility in Political Communication

Xinran Wang¹, Dr. Mohd Haizra Hashim²

Faculty of Business and Communication, INTI International University

Abstract: *Generative artificial intelligence has created new possibilities for political communication while raising ethical concerns about authenticity, disclosure, manipulation, public trust, and accountability. This article examines these issues through a qualitative comparative analysis of two 2024 cases: the AI-generated robocall imitating Joe Biden during the New Hampshire Democratic primary and the AI-assisted “Gemoy” persona used in Indonesia’s presidential campaign. Drawing on Hans Jonas’s responsibility ethics and rule-utilitarianism, the study distinguishes between deepfakes that fabricate political statements and softfakes that enhance political representation through AI-generated imagery. The analysis argues that synthetic political content is not inherently unethical; its ethical status depends on how it is produced, disclosed, distributed, and used. Deepfakes may undermine the evidentiary value of authentic political communication, while softfakes may influence political judgement through affective framing. Transparency and labelling are therefore necessary but insufficient. The article proposes a shared responsibility approach in which political actors, platforms, media institutions, developers, and audiences bear differentiated responsibilities according to their capacity to prevent or amplify harm.*

Keywords: *AI-generated political content, deepfakes, softfakes, political communication, media ethics.*

I. INTRODUCTION

Generative artificial intelligence (AI) has become an increasingly important tool in contemporary political communication, enabling political actors and campaigns to produce synthetic text, images, audio, video, and digital personas at unprecedented speed and scale. While these technologies may support accessibility, multilingual communication, and creative campaigning, they also raise concerns about authenticity, manipulation, transparency, and public trust (Battista, 2024; Floridi et al., 2018).

Deepfakes are particularly significant because AI-generated media can make political figures appear to say or do things that never occurred. Beyond direct deception, deepfakes may increase uncertainty about the authenticity of genuine political evidence, contributing to what Chesney and Citron (2019) call the “liar’s dividend.” Experimental research also suggests that synthetic political videos can increase uncertainty and reduce trust in news on social media (Vaccari & Chadwick, 2020). However, not all AI-generated political content is explicitly deceptive. Softfakes use AI to enhance, stylise, or reshape political representations, potentially influencing how audiences perceive candidates through emotional framing and idealisation (Chowdhury, 2024; Pawelec & Łabuz, 2026).

These differences raise a broader ethical question: responsibility should not be assigned only to audiences, who may have limited ability to identify synthetic content, but also to political actors, platforms, media organisations, and technology providers. Transparency and labelling may improve awareness, but their effectiveness depends on how disclosures are designed and interpreted (Dobber et al., 2024; Chomanski & Lauwaert, 2025).

This article examines these issues through two contrasting cases from the 2024 electoral cycle: the AI-generated robocall impersonating Joe Biden during the New Hampshire primary and the AI-assisted “Gemoy” persona associated with Prabowo Subianto’s Indonesian presidential campaign (FCC, 2024; Chowdhury, 2024). The analysis draws on Jonas’s (1984) responsibility ethics and Mill’s (1863/2001) rule-utilitarian perspective to examine technological responsibility and the broader consequences of normalising synthetic political communication.

The study asks three questions: (1) how do deepfakes and softfakes affect authenticity and public trust; (2) how should ethical responsibility be distributed among the actors involved; and (3) are transparency and labelling sufficient responses? The article argues that AI-generated political communication is not inherently unethical; its ethical implications depend on purpose, disclosure, technological power, audience capacity, and potential social consequences.

II. METHODOLOGY

This study employs a qualitative comparative case analysis based on documentary evidence. Two cases from the 2024 electoral cycle were purposively selected because they represent contrasting combinations of fabrication and disclosure: the concealed AI-generated voice used in the New Hampshire robocall and the openly stylised AI-assisted “Gemoy” political persona in Indonesia. The purpose of the comparison is analytical rather than statistical. The cases allow examination of whether different forms of synthetic political communication generate different ethical problems and corresponding responsibilities.

The study relies exclusively on secondary sources, including regulatory records, peer-reviewed academic literature, institutional documents, and credible journalistic accounts. The New Hampshire case is examined primarily through FCC records and relevant scholarship on political deepfakes, synthetic media, and democratic communication. The Gemoy case is examined through academic and journalistic discussions of AI-assisted political branding and digital campaigning (Chowdhury, 2024; Pawelec & Labuz, 2026). No interviews, surveys, experiments, or other primary human-participant research were conducted.

The analysis proceeded in four stages. First, relevant documentary materials were identified according to their direct relationship to the selected cases and to AI-generated political communication. Second, each case was examined in terms of its communicative context, production, distribution, and political function. Third, the material was coded according to four analytical dimensions: authenticity, disclosure, audience capacity, and responsibility allocation. Fourth, the cases were interpreted through Jonas’s (1984) responsibility ethics and a rule-utilitarian approach based on Mill (1863/2001).

The same analytical questions were applied to both cases: whether the content was presented as authentic or synthetic; whether its AI origin was disclosed; whether audiences possessed a reasonable capacity to assess it; what communicative purpose it served; and which actors possessed the practical capacity to prevent, mitigate, or amplify potential harm. The comparison therefore focuses on communicative function rather than audience frequency or statistical effects. Verified documentary evidence, scholarly interpretation, and analytical inference were kept conceptually distinct.

III. RESULTS AND DISCUSSION

The comparison of the New Hampshire robocall and the Gemoy campaign demonstrates that AI-generated political communication can create different forms of ethical risk depending on how synthetic media is produced, presented, and used.

The New Hampshire robocall represents a clear case of fabricated political speech. In January 2024, voters received an automated call using an AI-generated voice that imitated President Biden and discouraged participation in the Democratic primary. The FCC subsequently imposed enforcement measures against the scheme and also addressed the role of the telecommunications provider, demonstrating that the case involved both deceptive content and failures in communication infrastructure (FCC, 2024).

The ethical problem lies primarily in the creation of an apparently authentic political utterance that never occurred. Because the message was delivered through audio only, recipients had limited contextual information for assessing its authenticity, while research indicates that people cannot consistently distinguish synthetic from authentic political media (Groh et al., 2024). The timing of the call, shortly before the primary, further increased the importance of rapid interpretation and limited opportunities for verification. From Jonas’s (1984) responsibility perspective, responsibility should therefore fall primarily on the actors with greater technological and communicative capacity. The case also demonstrates the importance of infrastructural responsibility: the FCC’s action against the transmitting carrier indicates that responsibility extends beyond the original producer to intermediaries involved in distribution (FCC, 2024).

The broader significance of the case concerns the reliability of political evidence. Even when a fabricated message is later exposed, repeated exposure to synthetic political speech may make citizens less certain that genuine recordings are authentic. This corresponds to the “liar’s dividend” described by Chesney and Citron (2019). From a rule-utilitarian perspective, the ethical harm therefore extends beyond individual deception to the potential erosion of trust in recorded political communication.

The Gemoy case presents a different configuration. During Indonesia’s 2024 presidential campaign, AI-generated visual representations of Prabowo Subianto were used across campaign materials to construct a more youthful, friendly, and approachable persona. The campaign used the “Gemoy” identity in visual and social-media communication, including AI-generated imagery.

Unlike the New Hampshire robocall, the ethical issue here is not primarily fabricated speech. Instead, AI-assisted visualisation can shape the affective conditions through which a political actor is perceived. The use of a cute and approachable visual identity may direct attention toward particular characteristics while reducing the visibility of others. The ethical question is therefore less about whether audiences believe a false factual statement and more about how synthetic representation influences political evaluation.

This difference also limits the explanatory power of transparency alone. In a concealed deepfake, disclosure can directly undermine the mechanism of impersonation.

In an openly artificial political image, however, knowing that AI was used does not necessarily reveal its persuasive function. A label can establish provenance without explaining how the representation frames the political actor. The present study therefore does not claim that Gemoy imagery caused particular voting outcomes; rather, it identifies AI-assisted visual branding as a mechanism of affective political representation.

TABLE I
COMPARATIVE ETHICAL STRUCTURE OF TWO FORMS OF AI-GENERATED POLITICAL CONTENT

Dimension	New Hampshire robocall	Gemoy persona
Main form	AI-generated voice impersonation	AI-assisted visual representation
Ethical mechanism	Fabrication and deception	Affective framing and image construction
Transparency problem	Concealed synthetic identity	Artificiality more visible
Primary risk	Uncertainty about political evidence	Selective construction of political image
Main responsibility	Producer and communication intermediaries	Campaign and communication actors
Role of labelling	Potentially disrupts deception	Does not by itself resolve representational concerns

The comparison therefore suggests that deepfakes and softfakes should not be treated as ethically identical. Deepfakes can directly alter what citizens believe a political actor said or did, whereas softfakes may influence how citizens perceive a political actor without necessarily asserting a false proposition. Consequently, transparency should be understood as a necessary condition rather than a complete ethical solution. Research indicates that disclosure can increase awareness of persuasive intent, but its effects depend on label design and audience interpretation (Dobber et al., 2024; Chomanski & Lauwaert, 2025; Conrad et al., 2025).

The cases also demonstrate that responsibility is distributed across multiple actors. Political actors and campaign organisations determine the purpose and strategic use of synthetic content and therefore bear primary responsibility for its communicative function. Platforms and telecommunications providers influence whether content is authenticated, amplified, or distributed at scale. Media organisations affect whether uncertain synthetic material is verified or inadvertently reproduced, while developers influence generation controls, watermarking, and access to relevant technologies. Audiences retain a role in evaluating political content, but their responsibility should remain proportionate to their actual ability to detect synthetic media, which is limited (Groh et al., 2024). This distribution of responsibility corresponds to Jonas’s (1984) principle that responsibility should reflect the capacity to foresee and influence technological consequences. It also addresses the “problem of many hands,” in which complex technological systems make it difficult to identify a single responsible actor (Thompson, 1980; Nissenbaum, 1996). Responsibility should therefore be shared but not equal.

On this basis, a responsible framework should combine functional disclosure, consent for synthetic representations of identifiable individuals, stronger provenance mechanisms, verification-oriented media practices, and civic education focused on reasonable scepticism rather than perfect detection. These measures should not necessarily take the same regulatory form. Some, such as consent and provenance requirements, may be suitable for formal regulation, while functional disclosure and verification practices may be better addressed through professional or electoral codes. The framework should therefore be understood as a capability-based approach to ethical responsibility rather than a universal regulatory model.

Overall, the findings indicate that the ethical status of AI-generated political communication depends less on the mere presence of AI than on its communicative purpose, degree of disclosure, distribution mechanism, and potential consequences for political judgement. Transparency is an important starting point, but responsible political communication requires accountability for what synthetic content is designed to accomplish and how technological power is distributed among the actors involved.

IV. CONCLUSION

This article examined the ethical implications of AI-generated political communication through a comparison of the New Hampshire AI robocall and the Gemoy persona in Indonesia's 2024 presidential campaign. The analysis shows that the ethical implications of synthetic political content depend not simply on the use of AI or the presence of factual falsehood, but on the relationship between technological power, communicative purpose, disclosure, audience capacity, and broader consequences.

The two cases reveal different mechanisms of ethical risk. The New Hampshire robocall fabricated political speech and directly challenged the evidentiary value of recorded communication, while the Gemoy persona illustrates how AI-assisted representation may shape political perception through affective framing without necessarily fabricating a specific proposition. Responsibility should consequently be shared across the communication system but remain proportional to actors' technological and communicative capabilities. Political actors bear primary responsibility for the purpose and use of synthetic content, while platforms, infrastructures, media organisations, developers, and audiences hold different supporting responsibilities.

The analysis also shows that transparency is necessary but not sufficient. Disclosure can undermine concealed impersonation, as in the New Hampshire case, but may not resolve the persuasive or representational function of openly artificial content. The study therefore proposes functional disclosure, which considers not only whether AI was used but also what the synthetic content is intended to accomplish.

The study has several limitations. Its qualitative comparative design limits generalisability, the Gemoy case relies largely on secondary evidence, and the study does not establish audience-level causal effects. Future research should therefore combine documentary analysis with audience experiments, platform data, and multilingual empirical research.

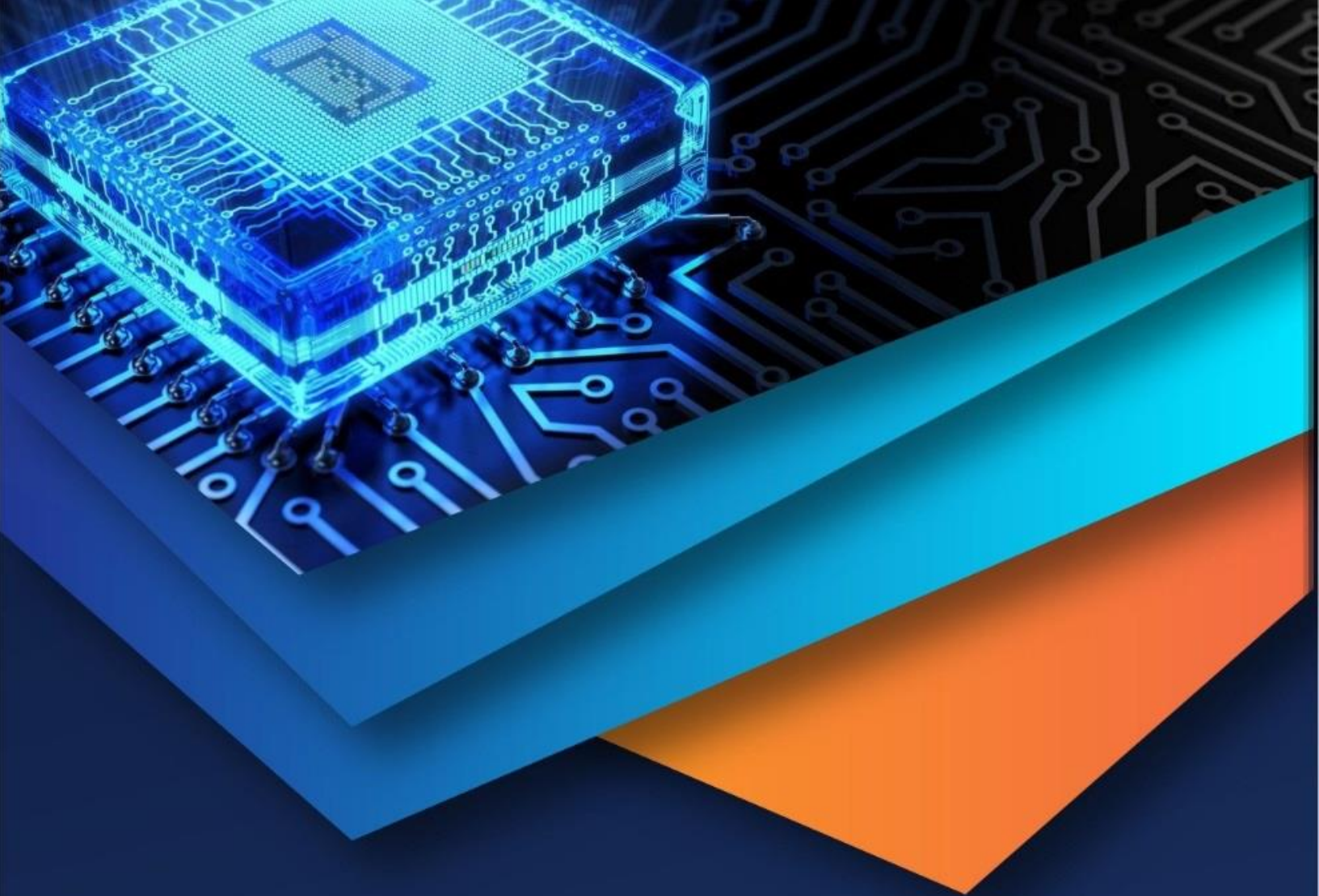
Overall, responsible governance of political AI should address not only whether content is authentic, but also who creates, distributes, amplifies, and contextualises synthetic representations. Transparency should serve as a starting point for accountability, complemented by capability-proportional responsibility across the political communication system.

REFERENCES

- [1] S Battista, D. (2024). Political communication in the age of artificial intelligence: An overview of deepfakes and their implications. *Society Register*, 8(2), 7–24.
- [2] Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107(6), 1753–1820.
- [3] Chomanski, B., & Lauwaert, L. (2025). Automated propaganda: Labeling AI-generated political content should not be required by law. *Journal of Applied Philosophy*, 42(3), 994–1015.
- [4] Chowdhury, R. (2024). AI-fuelled election campaigns are here—where are the rules? *Nature*, 628(8007), 237.
- [5] Coeckelbergh, M. (2020). Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Science and Engineering Ethics*, 26(4), 2051–2068.
- [6] Conrad, C., Nissen, A., Masoumi, K., Ramchandani, M., Fecury Braga, R., & Newman, A. J. (2025). Do truthfulness notifications influence perceptions of AI-generated political images? A cognitive investigation with EEG. *Computers in Human Behavior: Artificial Humans*, 5, Article 100185.
- [7] Diakopoulos, N., & Johnson, D. (2021). Anticipating and addressing the ethical implications of deepfakes in the context of elections. *New Media & Society*, 23(7), 2072–2098.
- [8] Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer.
- [9] Dobber, T., Kruikemeier, S., Helberger, N., & Goodman, E. (2024). Shielding citizens? Understanding the impact of political advertisement transparency information. *New Media & Society*, 26(11), 6715–6735.
- [10] Dobber, T., Metoui, N., Trilling, D., Helberger, N., & de Vreese, C. (2021). Do (microtargeted) deepfakes have real effects on political attitudes? *The International Journal of Press/Politics*, 26(1), 69–91.
- [11] Fallis, D. (2021). The epistemic threat of deepfakes. *Philosophy & Technology*, 34(4), 623–643.
- [12] Federal Communications Commission. (2024). In the matter of Steve Kramer: Notice of apparent liability for forfeiture (File No. EB-TCD-24-00036094, FCC 24-59).
- [13] Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707.
- [14] Groh, M., Sankaranarayanan, A., Singh, N., Kim, D. Y., Lippman, A., & Picard, R. (2024). Human detection of political speech deepfakes across transcripts, audio, and video. *Nature Communications*, 15, Article 7629.
- [15] Hameleers, M., van der Meer, T. G. L. A., & Dobber, T. (2024). Distorting the truth versus blatant lies: The effects of different degrees of deception in domestic and foreign political deepfakes. *Computers in Human Behavior*, 152, Article 108096.
- [16] Helberger, N., Pierson, J., & Poell, T. (2018). Governing online platforms: From contested to cooperative responsibility. *The Information Society*, 34(1), 1–14.
- [17] Jonas, H. (1984). *The imperative of responsibility: In search of an ethics for the technological age*. University of Chicago Press.
- [18] Kreiss, D., & McGregor, S. C. (2018). Technology firms shape political communication: The work of Microsoft, Facebook, Twitter, and Google with campaigns during the 2016 U.S. presidential cycle. *Political Communication*, 35(2), 155–177.



- [19] Kuipers, N., et al. (2024). Third time's the charm: The youth vote and Prabowo's victory in the 2024 Indonesian presidential election. *Pacific Affairs*. (Preprint/Online ahead of print).
- [20] Mill, J. S. (2001). *Utilitarianism* (G. Sher, Ed.). Hackett Publishing. (Original work published 1863)
- [21] Nightingale, S. J., & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. *Proceedings of the National Academy of Sciences*, 119(8), e2120481119.
- [22] Nissenbaum, H. (1996). Accountability in a computerized society. *Science and Engineering Ethics*, 2(1), 25–42.
- [23] Pawelec, M. (2022). Deepfakes and democracy (theory): How synthetic audio-visual media for disinformation and hate speech threaten core democratic functions. *Digital Society*, 1, Article 19.
- [24] Pawelec, M., & Łabuz, M. (2026). Are softfakes really “soft”? Semantic, ethical, and legal perspectives on deepfakes and AI-generated visualizations in political campaigning. *International Journal of Politics, Culture, and Society*.
- [25] Rini, R. (2020). Deepfakes and the epistemic backstop. *Philosophers' Imprint*, 20(24), 1–16.
- [26] Sunstein, C. R. (2017). *#Republic: Divided democracy in the age of social media*. Princeton University Press.
- [27] Thompson, D. F. (1980). Moral responsibility of public officials: The problem of many hands. *American Political Science Review*, 74(4), 905–916.
- [28] Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1).



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)