

The Ethics by Design Paradigm in Communication Regulation: Countering the Hegemony of Falsehood in Synthetic Media Architecture

Ahmad Sofian^{1*}, Hasan Sazali², Muhammad Randicha Hamandia³, Zhila Jannati⁴

^{1,2} Universitas Islam Negeri Sumatera Utara Medan, Indonesia

^{3,4} Universitas Islam Negeri Raden Fatah Palembang

DOI:

<https://doi.org/10.47134/villages.v7i2.519>

*Correspondence: Zhila Jannati

Email: ahmad4004253002@uinsu.ac.id

Received: 16-04-2026

Accepted: 16-05-2026

Published: 16-06-2026



Copyright: © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Abstract: This study examines the hegemony of falsehood in synthetic media and the limitations of conventional communication regulation in the age of generative artificial intelligence. Using a critical library-based method, the research synthesizes recent national and international literature, regulatory documents, and platform policy texts to analyze deepfakes, AI-generated disinformation, and post-facto legal responses. The findings show that synthetic media transforms falsehood from a representational distortion into an algorithmically produced reality that is scalable, personalized, and difficult to verify. Indonesia's current cyber-regulatory framework, including the Electronic Information and Transactions Law, remains largely reactive because it intervenes after harmful content has circulated. This creates epistemological and legal gaps involving velocity, accountability, intent, and platform responsibility. The study proposes Ethics by Design as a future communication policy paradigm that embeds transparency, traceability, algorithmic accountability, privacy protection, and human oversight into system architecture. It concludes that effective synthetic media governance requires shifting from *Lex Scripta* to *Lex Informatica* through co-regulation among the state, industry, academics, and civil society.

Keywords: Ethics by Design, Communication Regulation, Synthetic Media, Hegemony of Falsehood, Simulacra

Introduction

The global communication landscape is undergoing a fundamental shift from representational media to synthetic media, where text, images, audio, and video can be produced by generative AI with a level of realism that challenges ordinary public perception. This transformation intensifies the post-truth crisis because falsehood is no longer created only by human propaganda actors, but can also be automated, personalized, and distributed through algorithmic systems at high speed and scale. Recent studies show that AI-generated disinformation, deepfakes, and synthetic media can weaken digital authenticity, undermine journalistic trust, and create new vulnerabilities in democratic communication. Therefore, communication policy cannot rely only on reactive sanctions after harmful content has already circulated; it must address the technological architecture that enables synthetic deception to become socially influential ([Bontridder & Pouillet, 2021](#); [Fatimah et al., 2024](#); [Feuerriegel et al., 2023](#); [Verma, 2026](#); [von Sikorski & Hameleers, 2025](#)).

In this context, Ethics by Design becomes an urgent regulatory paradigm because it requires transparency, accountability, fairness, and respect for truth to be embedded into AI systems from the beginning of technological development. For synthetic media governance, this means that platforms and developers should integrate content authentication, digital watermarking, traceability mechanisms, and clear labeling of AI-generated materials directly into system design. Such an approach shifts communication law from merely punishing users toward regulating the code, architecture, and institutional responsibility behind digital circulation. In Indonesia, the idea of Lex AI is also relevant because it encourages a *sui generis* governance model for artificial intelligence that moves beyond conventional legal instruments. Thus, future communication regulation should protect the public's right to authentic information without falling into either digital authoritarianism or unregulated post-truth anarchy ([Brey & Dainow, 2024](#); [Lisaldy et al., 2024](#); [Luo et al., 2025](#)).

Methodology

This research employs a critical library-based methodology to examine the hegemony of falsehood, synthetic media, post-facto regulatory weakness, and Ethics by Design as conceptual and discursive problems in communication policy. The study does not collect empirical field data, but relies on systematic exploration, evaluation, and synthesis of national and international academic literature, indexed journal articles, monographs, regulatory documents, and digital platform policy texts published mainly between 2020 and 2026. This approach is appropriate because literature-based research can function as an independent scientific inquiry when it is conducted transparently, systematically, and critically to generate theoretical contribution rather than merely summarize previous studies ([Jamaluddin et al., 2025](#); [Lim et al., 2022](#); [Simsek et al., 2023](#); [Snyder, 2024](#)).

The analysis follows an interpretive qualitative design through three stages: organizing sources according to major themes, displaying the theoretical relationship between classical regulation and AI-era communication challenges, and drawing a conceptual synthesis for future regulatory models. Critical discourse analysis is used to examine how power, ideology, and legitimacy are embedded in policy documents and platform governance, while philosophy of communication technology helps reveal the ontological, epistemological, and axiological assumptions behind synthetic media systems. Through this method, the study positions technology not as a neutral instrument, but as a value-laden architecture that requires ethical evaluation from the earliest stage of design and regulation ([Bednar & Spiekermann, 2024](#); [Braun & Clarke, 2021](#); [Khan & Zaki, 2022](#); [Razali, 2024](#)).

Result and Discussion

Ontological Deconstruction of the Hegemony of Falsehood in Synthetic Media Architecture

Conventional communication regulation becomes increasingly inadequate in the era of synthetic media because deepfakes and AI-generated disinformation do not merely

distort existing reality but produce highly realistic informational objects that are difficult for ordinary users to verify. Generative AI can create persuasive text, images, audio, and video at scale, while deepfake circulation may also weaken public trust by making audiences suspicious not only of false content but also of authentic evidence ([Feuerriegel et al., 2023](#); [Twomey et al., 2023](#)).

This regulatory failure is intensified by the personalization logic of digital platforms, where AI-generated political messages can be microtargeted according to users' psychological tendencies and communicative vulnerabilities. Therefore, synthetic-media governance cannot rely only on content removal or post-hoc prohibition; it requires adaptive legal frameworks, clearer platform accountability, disclosure obligations, traceability mechanisms, and public digital-verification literacy to protect democratic communication from AI-driven manipulation ([Labuz, 2023](#); [Paliwang & Swandiani, 2025](#); [Simchon et al., 2024](#)).

The Epistemological Impasse of Conventional Communication Regulation (A Post-Facto Paradigm Critique)

A critical review of Indonesia's Electronic Information and Transactions Law, including its second amendment through Law No. 1 of 2024, shows that conventional cyber-regulation still tends to operate through a post-facto legal logic: the state intervenes after harmful electronic content has circulated, been reported, or generated social disturbance. This model becomes increasingly inadequate in the synthetic-media era because deepfakes and AI-generated disinformation move faster than investigation, prosecution, and adjudication, while current Indonesian frameworks still contain legal gaps in defining AI responsibility, deepfake misuse, and platform-based content moderation ([Anggara, 2026](#); [Febryani, 2025](#); [Hailtik & Afifah, 2024](#); [Republic of Indonesia, 2024](#); [Respati, 2024](#)).

The failure of post-facto regulation is also visible in three structural problems: the pacing problem between rapid AI innovation and slow legal reform, the ambiguity of mens rea when autonomous or semi-autonomous AI systems generate harmful outputs, and the tension between content removal, freedom of expression, and public legitimacy. Therefore, regulating synthetic media requires more than blocking or deleting content; it demands adaptive legal design, risk-based accountability, transparent moderation standards, stronger attribution mechanisms, and human-rights-oriented governance that can respond to deepfake harms before they become irreversible public damage ([Currie et al., 2025](#); [Kozyreva et al., 2023](#); [Romero Moreno, 2024](#); [van der Sloot & Wagenveld, 2022](#)).

The Construction of the Ethics-by-Design Paradigm in Communication Regulation

Facing the limitations of reactive digital regulation, this study argues that communication governance must shift toward an Ethics-by-Design paradigm, in which ethical values are embedded from the earliest stages of technological design rather than imposed only after social harm occurs. This paradigm is consistent with value-sensitive and socio-technical approaches to AI, which emphasize that digital systems are never neutral infrastructures but designed environments that can encode transparency, accountability,

privacy protection, fairness, and human oversight into their technical architecture ([Brey & Dainow, 2024](#); [Buijsman, 2024](#); [Sadek et al., 2024](#)).

In the context of synthetic media and AI-driven communication platforms, Ethics-by-Design requires four integrated safeguards: proactive transparency through clear disclosure and traceable digital provenance, algorithmic accountability through independent audits and impact assessment, privacy and verifiability by default to reduce manipulative data extraction, and human-in-the-loop oversight for high-risk decisions involving public information. Such a model is particularly relevant for Indonesia because comparative policy analysis suggests that national AI governance should move toward risk-based regulation, algorithmic auditing, transparency obligations, and stronger institutional accountability in order to respond to the rapid development of AI communication systems ([Radanliev, 2025](#); [Revolusi & Febriandy, 2025](#)).

Transforming Lex Scripta to Lex Informatica: Implementation and Challenges

The implementation of Ethics-by-Design in communication policy requires policymakers to move beyond purely textual regulation and recognize platform architecture, source code, and technical standards as normative instruments that can shape user behavior before harm occurs. However, recent scholarship also warns that the classical “code is law” thesis cannot be applied mechanically to generative AI because these systems are often opaque, adaptive, and emergent; therefore, AI-driven communication infrastructures require hybrid governance through technical standards, platform guardrails, transparency duties, and risk-based oversight ([Cantero Gamito & Marsden, 2024](#); [Dong et al., 2025](#); [Judge et al., 2025](#)).

At the same time, this paradigm faces serious political and ethical challenges because Big Tech platforms possess infrastructural power over data, algorithms, and digital markets, while developing countries such as Indonesia must avoid both weak digital sovereignty and excessive regulation that suppresses local innovation. Therefore, future communication governance should adopt a co-regulatory model in which the state defines public-interest principles, industry embeds safeguards and transparency mechanisms by design, and academics, civil society, and ethical-religious communication experts act as independent auditors to ensure that AI systems remain accountable, culturally contextual, and resistant to algorithmic colonialism ([Ahmed, 2026](#); [Fazriati et al., 2025](#); [Gu, 2023](#); [Putri, 2026](#); [Sitabuana et al., 2024](#)).

Conclusion

This study concludes that the rise of synthetic media has transformed falsehood from a simple problem of distorted representation into an algorithmically generated reality that is scalable, personalized, and difficult to verify. Deepfakes and AI-generated disinformation weaken public trust not only by spreading false content, but also by making authentic information appear doubtful. In this context, conventional communication regulation becomes insufficient because it still operates through a post-facto paradigm, where legal

intervention occurs only after harmful content has circulated and produced social disruption.

The study also finds that Indonesia's existing cyber-regulatory framework, including the Electronic Information and Transactions Law, still faces serious limitations in responding to AI-driven communication harms. These limitations include the speed gap between legal processes and technological innovation, uncertainty over algorithmic responsibility, and the difficulty of balancing content moderation with freedom of expression. Therefore, synthetic media governance cannot rely solely on blocking, deletion, or criminal sanctions, but must address the technological architecture that enables manipulation.

As a conceptual contribution, this study proposes Ethics-by-Design as a future paradigm for communication regulation. This paradigm requires transparency, traceability, algorithmic accountability, privacy protection, and human oversight to be embedded directly into media system design. Future regulation should move from *Lex Scripta* toward *Lex Informatica* through a co-regulatory model involving the state, technology industry, academics, civil society, and ethical-religious communication experts. In this way, communication policy can protect truth, public trust, and human dignity without falling into either digital authoritarianism or unregulated post-truth anarchy.

References

- Ahmed, S. A. (2026). Algorithmic dependence and digital colonialism: A conceptual framework for artificial intelligence in education and knowledge systems of the Global South. *Frontiers in Education*, 11, Article 1720563. <https://doi.org/10.3389/feduc.2026.1720563>
- Anggara, I. G. K. B. (2026). Legal liability of artificial intelligence in criminal law: A comparative study between Indonesia and European Union. *Indonesian Journal of Law and Justice*, 3(3), Article 12. <https://doi.org/10.47134/ijlj.v3i3.5565>
- Bednar, K., & Spiekermann, S. (2024). Eliciting values for technology design with moral philosophy: An empirical exploration of effects and shortcomings. *Science, Technology, & Human Values*, 49(3), 611–645. <https://doi.org/10.1177/01622439221122595>
- Bontridder, N., & Pouillet, Y. (2021). The role of artificial intelligence in disinformation. *Data & Policy*, 3, Article e32, 1–21. <https://doi.org/10.1017/dap.2021.20>
- Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in reflexive thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352. <https://doi.org/10.1080/14780887.2020.1769238>
- Brey, P., & Dainow, B. (2024). Ethics by design for artificial intelligence. *AI and Ethics*, 4(4), 1265–1277. <https://doi.org/10.1007/s43681-023-00330-4>

-
- Brey, P., & Dainow, B. (2024). Ethics by design for artificial intelligence. *AI and Ethics*, 4(4), 1265–1277. <https://doi.org/10.1007/s43681-023-00330-4>
- Buijsman, S. (2024). Transparency for AI systems: A value-based approach. *Ethics and Information Technology*, 26, Article 34. <https://doi.org/10.1007/s10676-024-09770-w>
- Cantero Gamito, M., & Marsden, C. T. (2024). Artificial intelligence co-regulation? The role of standards in the EU AI Act. *International Journal of Law and Information Technology*, 32, Article eaae011. <https://doi.org/10.1093/ijlit/eaae011>
- Currie, W. L., Schlagwein, D., Leimeister, J. M., & Willcocks, L. (2025). Rethinking technology regulation in the age of AI risks. *Journal of Information Technology*, 40(1), 1–10. <https://doi.org/10.1177/02683962251378815>
- Dong, Y., Mu, R., Zhang, Y., Jin, G., Qi, Y., Hu, J., Zhao, X., Meng, J., Ruan, W., & Huang, X. (2025). Safeguarding large language models: A survey. *Artificial Intelligence Review*, 58, Article 382. <https://doi.org/10.1007/s10462-025-11389-2>
- Fatimah, R., Mumtaz, A., Fahrezi, F. M., & Zakaria, D. (2024). AI-generated misinformation: A literature review. *Indonesian Journal of Artificial Intelligence and Data Mining*, 7(2), 241–254. <https://doi.org/10.24014/ijaidm.v7i2.26455>
- Fazriati, R. M., Rosadi, S. D., & Amalia, P. (2025). The urgency of regulating the transparency principle of the “AI system” in Indonesia: The phenomenon of self-preferencing and regulation in the European Union. *Journal of Law, Politic and Humanities*, 5(3), 2061–2074. <https://doi.org/10.38035/jlph.v5i3.1485>
- Febryani, E. (2025). The impact of content moderation policy on the spread of fake news on social media in Indonesia. *The Easta Journal Law and Human Rights*, 3(3), 176–183. <https://doi.org/10.58812/eslhr.v3i03.639>
- Feuerriegel, S., DiResta, R., Goldstein, J. A., Kumar, S., Lorenz-Spreen, P., Tomz, M., & Pröllochs, N. (2023). Research can help to tackle AI-generated disinformation. *Nature Human Behaviour*, 7(11), 1818–1821. <https://doi.org/10.1038/s41562-023-01726-2>
- Feuerriegel, S., DiResta, R., Goldstein, J. A., Kumar, S., Lorenz-Spreen, P., Tomz, M., & Pröllochs, N. (2023). Research can help to tackle AI-generated disinformation. *Nature Human Behaviour*, 7(11), 1818–1821. <https://doi.org/10.1038/s41562-023-01726-2>
- Gu, H. (2023). Data, Big Tech, and the new concept of sovereignty. *Journal of Chinese Political Science*, 29(4), 591–612. <https://doi.org/10.1007/s11366-023-09855-1>
- Hailtik, A. G. E., & Afifah, W. (2024). Criminal responsibility of artificial intelligence committing deepfake crimes in Indonesia. *Asian Journal of Social and Humanities*, 2(4), 776–795. <https://doi.org/10.59888/ajosh.v2i4.222>
-

-
- Jamaluddin, M., Rahmatullah, A., & Farid, M. (2025). Library research methodology in education: Fundamental concepts and implementation. *Cognitive: Jurnal Pendidikan dan Pembelajaran*, 3(2), 128–187. <https://doi.org/10.61743/cg.v3i2.160>
- Judge, B., Nitzberg, M., & Russell, S. (2025). When code isn't law: Rethinking regulation for artificial intelligence. *Policy and Society*, 44(1), 85–97. <https://doi.org/10.1093/polsoc/puae020>
- Khan, M. A., & Zaki, S. (2022). Corpus assisted critical discourse analysis of Pakistan's language education policy documents: What are the existing language ideologies? *SAGE Open*, 12(3). <https://doi.org/10.1177/21582440221121805>
- Kozyreva, A., Herzog, S. M., Lewandowsky, S., Hertwig, R., Lorenz-Spreen, P., Leiser, M., & Reifler, J. (2023). Resolving content moderation dilemmas between free speech and harmful misinformation. *Proceedings of the National Academy of Sciences of the United States of America*, 120(7), Article e2210666120. <https://doi.org/10.1073/pnas.2210666120>
- Łabuz, M. (2023). Regulating deep fakes in the Artificial Intelligence Act. *Applied Cybersecurity & Internet Governance*, 2(1), 1–42. <https://doi.org/10.60097/ACIG/162856>
- Lim, W. M., Kumar, S., & Ali, F. (2022). Advancing knowledge through literature reviews: "What," "why," and "how to contribute." *The Service Industries Journal*, 42(7–8), 481–513. <https://doi.org/10.1080/02642069.2022.2047941>
- Lisaldy, F., Ismail, & Iryani, D. (2024). Lex AI: Solution for governance of artificial intelligence in Indonesia. *DiH: Jurnal Ilmu Hukum*, 20(1), 50–67. <https://doi.org/10.30996/dih.v20i1.9632>
- Luo, H., Li, L., & Li, J. (2025). Digital watermarking technology for AI-generated images: A survey. *Mathematics*, 13(4), Article 651. <https://doi.org/10.3390/math13040651>
- Paliwang, A. N. A. A., & Swandiani, N. L. P. E. (2025). Artificial intelligence regulation in the protection of democracy: A legal analysis of political deepfakes and disinformation in the 2024 election. *Hakim: Jurnal Ilmu Hukum dan Sosial*, 3(4), 1433–1455. <https://doi.org/10.51903/rx25ap29>
- Putri, D. A. K. (2026). How Big Tech determines the state's digital sovereignty: A comparative study of India and Indonesia. *Jurnal Hubungan Internasional*, 14(2). <https://journal.umy.ac.id/index.php/jhi/article/view/28792>
- Radanliev, P. (2025). Privacy, ethics, transparency, and accountability in AI systems for wearable devices. *Frontiers in Digital Health*, 7, Article 1431246. <https://doi.org/10.3389/fdgth.2025.1431246>
-

-
- Razali, G. (2024). The philosophy of communication in technology-based education: Study of research trends in the last three years assisted by NVivo 12 Pro. *Jurnal ASPIKOM*, 9(1), 1–12. <https://doi.org/10.24329/aspikom.v9i1.1405>
- Republic of Indonesia. (2024). *Law Number 1 of 2024 concerning the second amendment to Law Number 11 of 2008 concerning Electronic Information and Transactions*. State Gazette of the Republic of Indonesia Year 2024 Number 1. <https://peraturan.bpk.go.id/details/274494/uu-no-1-tahun-2024>
- Respati, A. A. (2024). Reformulasi UU ITE terhadap artificial intelligence dibandingkan dengan Uni Eropa dan China AI Act Regulation. *Jurnal USM Law Review*, 7(3), 1737–1758. <https://doi.org/10.26623/julr.v7i3.10578>
- Revolusi, P., & Febriandy, R. K. (2025). Developing AI regulations in Indonesia: Policy recommendations based on comparative policy analysis from the European Union, the United States, and Singapore. *Jurnal Indonesia: Manajemen Informatika dan Komunikasi*, 6(2), 1035–1049. <https://doi.org/10.63447/jimik.v6i2.1380>
- Romero Moreno, F. (2024). Generative AI and deepfakes: A human rights approach to tackling harmful content. *International Review of Law, Computers & Technology*, 38(3), 297–326. <https://doi.org/10.1080/13600869.2024.2324540>
- Sadek, M., Calvo, R. A., & Mougnot, C. (2024). Designing value-sensitive AI: A critical review and recommendations for socio-technical design processes. *AI and Ethics*, 4, 949–967. <https://doi.org/10.1007/s43681-023-00373-7>
- Simchon, A., Edwards, M., & Lewandowsky, S. (2024). The persuasive effects of political microtargeting in the age of generative artificial intelligence. *PNAS Nexus*, 3(2), Article pgae035. <https://doi.org/10.1093/pnasnexus/pgae035>
- Simsek, Z., Fox, B., & Heavey, C. (2023). Systematicity in organizational research literature reviews: A framework and assessment. *Organizational Research Methods*, 26(2), 292–321. <https://doi.org/10.1177/10944281211008652>
- Sitabuana, T. H., Sanjaya, D., Amri, I. F., & Nethan, N. (2024). The urgency of artificial intelligence code of ethics. *Indonesia Law Review*, 14(3), Article 5. <https://scholarhub.ui.ac.id/ilrev/vol14/iss3/5>
- Snyder, H. (2024). Designing the literature review for a strong contribution. *Journal of Decision Systems*, 33(4), 551–558. <https://doi.org/10.1080/12460125.2023.2197704>
- Twomey, J., Ching, D., Aylett, M. P., Quayle, M., Linehan, C., & Murphy, G. (2023). Do deepfake videos undermine our epistemic trust? A thematic analysis of tweets that
-

discuss deepfakes in the Russian invasion of Ukraine. *PLOS ONE*, 18(10), Article e0291668. <https://doi.org/10.1371/journal.pone.0291668>

van der Sloot, B., & Wagenveld, Y. (2022). Deepfakes: Regulatory challenges for the synthetic society. *Computer Law & Security Review*, 46, Article 105716. <https://doi.org/10.1016/j.clsr.2022.105716>

Verma, A. (2026). Deepfakes and the crisis of digital authenticity: Ethical challenges in the age of synthetic media. *Journal of Information, Communication and Ethics in Society*, 24(1), 59–76. <https://doi.org/10.1108/IICES-04-2025-0083>

Von Sikorski, C., & Hameleers, M. (2025). Disinformation in the age of artificial intelligence (AI): Implications for journalism and mass communication. *Journalism & Mass Communication Quarterly*, 102(4). <https://doi.org/10.1177/10776990251375097>
