

# AI-Generated Disinformation and the Public Trust in Indonesia's Digital Democracy

Ahmad Fazar Maulana<sup>1</sup>, Dendi Superdi<sup>2</sup>, Ariffin Rismansyah<sup>3</sup>, Taqiya Arini Putri<sup>4</sup>  
*UIN Sunan Gunung Djati Bandung*

**Corresponding author: Taqiya Arini Putri**  
Email: Taqiya.arinip@uinsgd.ac.id

## Abstract

This study investigates the phenomenon of artificial intelligence (AI) driven disinformation and its implications for public trust within Indonesia's digital democracy. It centers on a deepfake video portraying a figure resembling Sri Mulyani Indrawati, the Indonesian Minister of Finance, falsely claiming that "honorary teachers are a burden to the state." Employing a qualitative research design, the study utilizes Norman Fairclough's Critical Discourse Analysis (CDA) framework, encompassing three analytical dimensions: textual analysis, discursive practice, and social practice. The findings demonstrate that AI technologies enable the manipulation of meaning through the strategic orchestration of language and visual representation, which is further amplified by social media algorithms prioritizing emotionally resonant content. These dynamics significantly accelerate the dissemination of disinformation and shape public perception within the digital public sphere. As a result, this phenomenon contributes to a growing crisis of public trust (trust deficit) and undermines the legitimacy of digital democracy. By integrating the theoretical lenses of Information Disorder and Digital Democracy, the study contends that AI-driven disinformation should not be viewed merely as a technological concern, but rather as a complex ethical and informational justice issue with far-reaching consequences for the quality and resilience of democracy in the digital age.

**Keywords:** Deepfake Technology; Artificial Intelligence; Digital Democracy; Public Trust.

## Introduction

The rapid advancement of digital technology in the twenty-first century has fundamentally transformed the ways in which individuals communicate, participate, and interact within political systems. One of the most significant yet challenging innovations in this transformation is the emergence of artificial intelligence (AI), which is now capable of realistically imitating human behavior, voice, and facial expressions through deep learning algorithms. This technology, commonly known as deepfake, enables the recreation of an individual's face and voice with remarkable precision (Goodfellow et al., 2014). Initially developed for creative and educational purposes, deepfake technology has increasingly evolved into a tool of digital manipulation that poses serious threats to the reliability of information and the stability of democratic systems (Chesney & Citron, 2019).

Within the framework of digital democracy, technological innovation is expected to function as a mechanism for strengthening civic participation, enhancing governmental transparency, and improving public accountability.

Coleman and Blumler (2009) argue that digital democracy requires open and honest two-way communication between governments and citizens. Ideally, developments in information technology and

artificial intelligence should enhance the quality of democracy by expanding access to information and promoting transparency.

However, the empirical reality often reflects paradoxically: digital technologies, particularly AI, are increasingly used to disseminate disinformation, intensify political polarization, and erode public trust in democratic institutions (Wardle & Derakhshan, 2017).

This phenomenon has become particularly occur in Indonesia through the circulation of a manipulated video depicting the Minister of Finance, Sri Mulyani Indrawati, seemingly delivering a speech in which she allegedly stated that "honorary teachers are a burden to the state." The video, generated through voice-cloning and deepfake synthesis technologies, first appeared on the TikTok platform and rapidly spread across various social media platforms, including WhatsApp, Facebook, and X (formerly Twitter). Due to the highly realistic visual and audio characteristics of the video, the public perceived the content as authentic. As a result, the video triggered widespread public outrage, particularly among educators who viewed the statement as an insult to the teaching profession and as evidence of the government's disregard for teachers' welfare.

Several days later, it was confirmed that the video had been artificially generated using AI

technology. However, the clarification was delivered belatedly. The fabricated video had already sparked collective anger and shaped negative public perceptions of the government. Public trust in state officials and official media sources declined significantly. The situation escalated further when social unrest erupted during mass demonstrations in front of the legislative building. In the course of these protests, the private residence of Sri Mulyani in Bintaro, South Tangerang, was looted by protesters who openly removed belongings from the property. Although the demonstrations were initially driven by broader political grievances, the emotional mobilization fueled by AI-driven disinformation contributed to the escalation of violence and social disruption. Under intense social pressure and amid a crisis of political legitimacy, Sri Mulyani ultimately chose to resign from her position as Minister of Finance. This situation illustrates how AI-driven disinformation can transcend the boundaries of the digital sphere and generate tangible socio-political consequences. Technologies are increasingly used to produce highly realistic falsehoods that mislead the public and weaken democratic legitimacy. In this context, the digital technologies which have to be strengthened democracy, instead, it demonstrates that these technologies can undermine democratic institutions through the erosion of public trust.

Nurdin and Nugraha (2025) argue that the proliferation of deepfake technology in Indonesia has become a multidimensional threat, affecting not only cybersecurity but also social cohesion and political stability, particularly in the context of limited digital literacy and the absence of comprehensive AI regulation. Similarly, Santiko and Bahri (2024) demonstrate that political deepfakes in Indonesia function as instruments for shaping public opinion by manipulating citizens' emotions and perceptions. Zahra et al. (2025) further emphasize the paradoxical nature of artificial intelligence, while it has the potential to strengthen democratic processes, it can also weaken them if deployed without strong ethical principles and adequate digital literacy.

Coleman and Blumler (2009) emphasize that a healthy democracy relies heavily on public trust in government institutions, media organizations, and state authorities. When this trust collapses as a result of AI-driven disinformation, democracy loses its moral foundation of legitimacy. Floridi (2023) further argues that ethical governance of artificial intelligence must become an integral component of democratic governance in order to prevent technological misuse for social and political manipulation.

Based on these conditions, this study is motivated by the growing urgency of AI-driven

disinformation that contributes to a crisis of public trust and threatens the quality of digital democracy in Indonesia. Previous studies (Nurdin & Nugraha, 2025; Santiko & Bahri, 2024) have examined the threats posed by deepfake technology to cybersecurity and political communication, however, relatively few studies have explored its broader social implications for public trust and democratic legitimacy. Therefore, this research aims to analyze the forms, patterns of dissemination, and socio-political impacts of AI-based disinformation on public trust, using the case of the Sri Mulyani deepfake video as a representative example of emerging threats within Indonesia's digital democratic sphere.

The novelty of this study lies in its integrative approach, combining the theoretical perspectives of digital democracy, information disorder, and trust erosion to explain how technologies that are normatively expected to strengthen democracy can instead be used to undermine it. Theoretically, this research contributes to the broader discourse on the relationship between artificial intelligence and democratic legitimacy. Practically, the findings are expected to provide a basis for the development of digital literacy policies and ethical AI regulations aimed at strengthening democratic resilience in Indonesia's digital era. The widespread circulation of the manipulated video generated public anger and triggered protests in several regions, placing Sri Mulyani under intense social pressure and creating a crisis of political legitimacy within the digital public sphere. Accordingly, this study seeks to explain how AI-driven disinformation not only influences public opinion but also challenges the ethical foundations and legitimacy of digital democracy in Indonesia.

### Method

This study is situated within a qualitative research paradigm, employing an interpretive approach to understand how meaning, authority, and legitimacy are constructed and contested within digital discourse. Specifically, it utilizes Critical Discourse Analysis (CDA) in conjunction with digital ethnography to examine the discursive dynamics surrounding AI-driven disinformation in Indonesia's digital public sphere. Although the study applies a sampling strategy, it does not seek statistical generalization; rather, it adopts purposive sampling to select discursive materials that are theoretically relevant, influential, and representative of dominant narrative patterns. This approach reflects the epistemological foundation of qualitative inquiry, where data selection is guided by interpretive depth and contextual significance rather than numerical representativeness. Through this methodological positioning, the research

prioritizes the analysis of discourse as a socially constructed practice, enabling a deeper exploration of how digital texts, public reactions, and institutional responses interact in shaping political perception and public trust. Instead of merely documenting the presence of disinformation, the research focused on how meaning, authority, and legitimacy are constructed, negotiated, and contested within Indonesia's digital public sphere following the circulation of a deepfake video portraying Indonesia's Minister of Finance, Sri Mulyani Indrawati.

Data collection was conducted through a multi-layered digital ethnographic approach, combining online discourse observation, content archiving, and institutional narrative tracing. This approach allows the researcher to examine how digital texts interact with broader sociopolitical structures and communicative practices.

The primary data consisted of three interconnected discursive materials. First, the study collected the deepfake video content itself, treating it as a multimodal text composed of visual, audio, and narrative elements that simulate political authority. Second, the research gathered public responses and user-generated discourse, particularly comments, reposts, and reactions on TikTok, where the video gained significant traction. These comments were systematically documented to observe patterns of interpretation, skepticism, amplification, or contestation among digital audiences. Third, the study compiled institutional counter-discourses, including official statements and clarifications issued by government institutions and fact-checking organizations such as Tempo Cek Fakta and Kominfo.go.id. These responses represent authoritative attempts to reassert epistemic credibility and stabilize public understanding amid the circulation of manipulated content.

Complementing these sources, secondary data were obtained from academic publications, scholarly articles, policy reports, and credible online news coverage discussing AI-generated disinformation, digital political communication, and trust in democratic institutions. These materials provided contextual grounding and theoretical enrichment for interpreting the empirical findings.

Data collection took place during October 2025, a period following the widespread dissemination of the deepfake video and preceding the reported looting incident at Sri Mulyani's residence on 31 August 2025. This temporal scope was chosen to capture the early stages of discursive formation, when interpretations, rumors, and institutional responses were actively unfolding within digital platforms.

The population of this research encompasses

the broader field of digital discourse related to the Sri Mulyani deepfake video across Indonesia's online public sphere. Given the expansive nature of online communication, the study employed purposive sampling, selecting discursive fragments that were particularly influential, widely circulated, or representative of dominant narrative patterns. The purposive sampling criteria prioritized posts with high engagement metrics, comments that articulated distinct interpretive frames, and institutional responses that shaped public discourse.

To preserve analytical rigor, all collected data were archived, categorized, and coded according to their discursive function, such as narrative framing, legitimacy claims, emotional mobilization, or distrust toward institutions. This systematic organization enabled the identification of recurring rhetorical patterns and ideological positioning within the dataset. Through this comprehensive data collection process, the study assembled a discursive corpus that reflects the complex interaction between technological manipulation, public interpretation, and institutional authority. This corpus serves as the empirical foundation for the subsequent Critical Discourse Analysis (CDA), allowing the research to interrogate how AI-generated disinformation operates not merely as false content but as a discursive practice capable of reshaping political perception and public trust in the context of digital democracy.

## Discussion

This section presents the findings of the study concerning the phenomenon of artificial intelligence (AI)-driven disinformation in Indonesia, focusing on the case of a deepfake video depicting a figure resembling the Minister of Finance, Sri Mulyani Indrawati. The analysis is conducted using Norman Fairclough's Critical Discourse Analysis (CDA) framework (1989), which consists of three analytical dimensions: the level of text, the level of discursive practice, and the level of social practice.

### Characteristics of the Deepfake Video Case

This study indicates that AI-driven disinformation in Indonesia is increasingly widespread and has a significant impact on public perception within the context of digital democracy. The object of this research is a deepfake video depicting a figure resembling the Indonesian Minister of Finance, Sri Mulyani Indrawati, uploaded by the TikTok account @humaidiabi27 on 19 August 2025. The 20-second video shows a manipulated representation of Sri Mulyani appearing to state that "honorary teachers are a burden to the state."

The video clip is accompanied by a visual

montage showing a veiled woman crying, which intensifies the emotional tone of the content and creates the impression that the video represents an authentic recording. This emotional framing contributes to the construction of perceived authenticity, making the manipulated content appear credible and convincing to viewers despite its fabricated nature.

The technical investigation conducted in this study indicates that the video was generated using voice-cloning and lip-synchronization technologies based on artificial intelligence. These technologies manipulate both audio and visual elements with a high degree of realism, thereby producing what Chesney and Citron (2019) describe as the illusory truth effect, in which information that appears realistic is more likely to be perceived as true even when it lacks factual basis. This phenomenon illustrates how AI technology can be utilized to create simulated realities capable of influencing public perception of political figures.

The Ministry of Communication and Digital Affairs of the Republic of Indonesia (Komdigi RI), through a press release issued on 20 August 2025, confirmed that the video was a hoax. The Head of the Bureau of Communication and Information Services of the Ministry of Finance, Deni Surjantoro, explained that the video was manipulated from an original speech delivered by Sri Mulyani during the Indonesian Convention on Science, Technology, and Industry (KSTII) held at the Bandung Institute of Technology (ITB) on 7 August 2025. In the original speech, Sri Mulyani emphasized the importance of appreciating and improving the welfare of educators rather than demeaning the teaching profession.

Sri Mulyani also issued an official clarification through her personal Instagram account on 19 August 2025, stating:

"I have never said that teachers are a burden to the state. The circulating video is a misleading result of digital manipulation."

Despite this clarification, the deepfake video had already shaped negative public perceptions. Many social media users were unaware of the context of the original speech and assumed that the statement presented in the video was genuine. This situation demonstrates that the speed of digital disinformation dissemination often surpasses the ability of official institutions to provide timely clarification.

Technical analysis indicates that the video was generated using artificial intelligence-based voice-cloning and lip-synchronization technologies, which enable the replication of speech patterns and facial movements with a high level of realism. Through the integration of synthetic audio and manipulated visual cues, these technologies produce what Chesney and

Citron (2019) describe as the illusory truth effect, a cognitive phenomenon in which information that appears visually and audibly convincing is more likely to be perceived as factual, even when it lacks empirical validity. In this context, the deepfake video demonstrates how AI technology can construct a simulated reality capable of influencing public interpretation and shaping perceptions of political actors.



Figure 1 Deepfake video depicting Sri Mulyani, Source: TikTok (@humaidiabi27, August 19th 2025).

The ability of deepfake technology to generate highly persuasive audiovisual content highlights a growing challenge in contemporary digital communication environments. When manipulated content is presented in formats that resemble authentic political communication. Such as speeches or official statements, audiences may interpret such content as credible without engaging in critical verification. Consequently, AI-driven media manipulation not only distorts factual information but also reshapes the epistemic boundaries between authenticity and fabrication in digital public discourse.

In the original context of the speech, Sri Mulyani emphasized the importance of acknowledging and improving the welfare of educators rather than demeaning the teaching profession.

The clarification illustrates the significant disparity between the rapid dissemination of disinformation and the relatively slower process of institutional fact-checking. By the time official clarification was issued, the manipulated content had already circulated widely across digital platforms, shaping public perception and contributing to the spread of misleading narratives. Such dynamics

demonstrate how AI-generated disinformation can disrupt informational credibility and challenge the authority of government communication within the digital public sphere.



Figure 2 Original speech of Sri Mulyani at the Indonesian Convention on Science, Technology, and Industry (KSTII). Source: YouTube channel @CNNIndonesia.

Sri Mulyani also issued a public clarification through her personal Instagram account on 19 August 2025, stating: *"I have never stated that teachers are a burden to the state. The circulating video is a misleading result of digital manipulation."*



Figure 3 Sri Mulyani Indrawati's official clarification of the deepfake video posted on her personal Instagram account on 19 August 2025. Source: Inilah.com (2025), Instagram post @smindrawati, 19 August 2025.

Despite the issuance of its clarification, the circulation of the deepfake video had already shaped negative perceptions among the public. Many social media users were unaware of the original context of the speech and assumed that the statement presented in the video was authentic. This situation demonstrates that the speed of digital disinformation dissemination often exceeds the capacity of government institutions to provide timely clarification.

## Social Media Reach and Public Response

Based on digital observation conducted during the period of 19 August to 18 October 2025, the deepfake video accumulated 1,400,000 views, 97,500 likes, 8,066 comments, and 7,419 shares. This high level of interaction indicates the extensive reach of the content and its strong emotional resonance among online audiences.

To understand the dynamics of public opinion, the researchers analyzed the ten most prominent public comments on the video post. These comments were classified into three primary categories:

1. Emotional (Believing): comments reflecting direct acceptance of the video's content, often accompanied by moral outrage directed toward public officials.
2. Skeptical (Critical): comments demonstrating digital awareness and encouraging others to verify the authenticity of the information.
3. Reflective/Neutral: Comments expressing evaluative perspectives that connect the video's narrative with broader social issues.

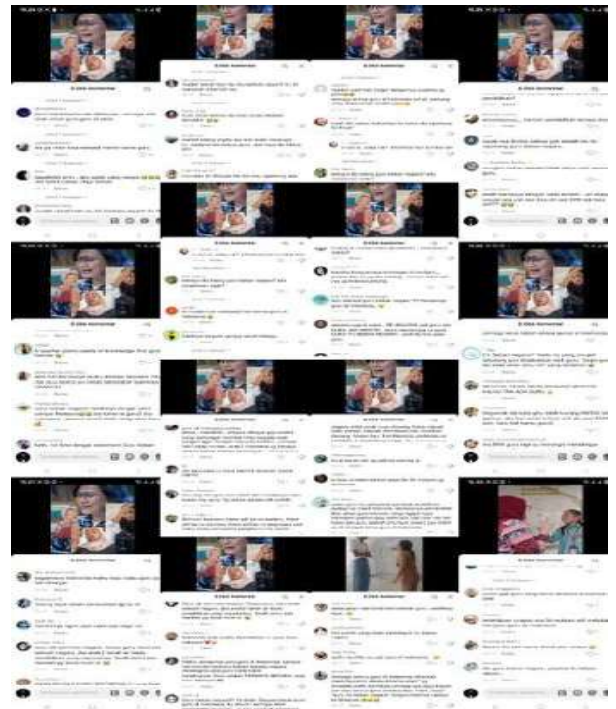


Figure 4 Screenshot of public comments on the TikTok video posted by the account @humaidiabi27 on 19 August 2025. Source: Author's documentation, 2025.

## Critical Discourse Analysis (CDA) Based on Fairclough's Model

The analysis is conducted using Norman Fairclough's Critical Discourse Analysis (CDA) framework, which examines discourse through three

interconnected levels: textual analysis, discursive practice, and social practice. This framework enables a comprehensive understanding of how language, media structures, and social contexts interact in shaping political meaning and public perception in the digital sphere.

#### a. Textual Level

At the textual level, the statement “honorary teachers are a burden to the state” contains a negative framing that positions teachers as a socio-economic meaning. Statements such as “A minister would never become a minister without teachers” illustrate how audiences interpret the video as a moral offense against the teaching profession. This indicates that the language embedded in the deepfake video successfully shaped public perception in accordance with the narrative intended by its creators.

#### b. Discursive Practice Level

At the level of discursive practice, the dissemination of the video occurred primarily through the TikTok platform, whose algorithmic structure tends to prioritize emotionally engaging content. This phenomenon creates what can be described as algorithmic amplification, in which emotionally provocative messages gain significantly greater visibility than factual corrections or official clarifications.

The video was produced within the context of the attention economy, where the value of digital content is determined by its ability to attract engagement, reactions, and emotional responses from users. Public comments subsequently became part of the discursive reproduction process: some users reproduced the hoax by resharing the video, while others acted as counter-discourse agents by questioning or correcting the information.

This analysis demonstrates that discourse production and consumption in digital environments is no longer a one-way communication process. Instead, audiences function as active participants who either reinforce or challenge the spread of disinformation through their interactions, comments, and algorithmic engagement.

#### c. Social Practice Level

At the level of social practice, this case reflects a broader transformation in the structure of power within the ecosystem of digital democracy. Artificial intelligence technologies function as new socio-technical actors capable of producing and distributing political meanings on a massive scale without passing through traditional journalistic verification processes.

As a consequence, a crisis of public trust (trust

deficit) emerges toward government officials, media institutions, and state authorities. Many individuals place greater trust in viral content circulating on social media than in official government clarifications. This phenomenon indicates a shift in the locus of informational authority, from formal institutions to algorithm-driven digital platforms shaped by collective emotions.

Moreover, the circulation of the deepfake video produced tangible social consequences, including polarization of public opinion, anti-government sentiment, and even the mobilization of protest narratives in demonstrations in Jakarta. These developments demonstrate that AI-driven disinformation is not merely a technological issue but a socio-political phenomenon capable of influencing social stability and democratic legitimacy in Indonesia.

### Interpreting AI Disinformation in Digital Democracy

From the perspective of Information Disorder, the Sri Mulyani deepfake video represents a clear example of disinformation, defined as false information deliberately created to mislead audiences (Wardle & Derakhshan, 2017). Unlike misinformation, which may be shared unintentionally, disinformation is intentionally constructed to manipulate public perception. The deepfake video analyzed in this study demonstrates how AI technology can generate a fabricated reality that closely resembles authentic communication. By imitating the voice, facial expressions, and visual presence of a public official, the manipulated content creates a convincing narrative that encourages audiences to interpret the information as credible.

The effectiveness of such disinformation lies primarily in its emotional appeal rather than its factual accuracy. This pattern is evident in the analysis of public responses to the video, where emotional reactions constituted a significant portion of the comments analyzed. Emotional responses, including expressions of anger and moral outrage toward public officials, illustrate how audiences often respond to the moral implications of a message rather than engaging in critical verification. This finding supports Wardle and Derakhshan’s argument that modern disinformation spreads effectively by mobilizing affective responses, thereby shaping perceptions through emotional resonance rather than rational evaluation.

Furthermore, the spread of disinformation in this case cannot be attributed solely to the intentions of content creators. The dynamics of digital dissemination are also influenced by algorithmic structures embedded in social media platforms. These algorithmic systems

prioritize content that generates engagement, such as emotional reactions, controversy, and user interaction, thereby amplifying the visibility of sensational or provocative narratives. As a result, manipulated content can spread more rapidly than factual clarifications. In this study, the official clarification issued by Komdigi and Sri Mulyani circulated more slowly and gained less visibility than the original deepfake video. This phenomenon illustrates a structural shift in digital communication from truth-based communication toward attention-based communication, in which the perceived credibility of information is increasingly determined by its visibility and engagement rather than by its factual validity.

Within the framework of Digital Democracy, this situation reflects a broader crisis of public trust. Coleman and Blumler (2009) argue that democratic communication systems rely on three fundamental principles: public trust, informational transparency, and deliberative participation. These principles require an information environment in which citizens can engage with political communication based on reliable and verifiable information. However, the findings of this study suggest that the proliferation of AI-generated disinformation undermines these foundations by distorting the informational environment within which democratic deliberation occurs.

The viral circulation of the Sri Mulyani deepfake video illustrates how manipulated information can erode public confidence in political institutions. When audiences encounter persuasive but false content, the resulting confusion and mistrust may weaken their confidence in official sources of information. In this sense, the empirical reality diverges from the normative ideal of digital democracy. Ideally, digital technologies should facilitate open communication, transparency, and greater citizen participation. In practice, however, these technologies can accelerate the dissemination of misinformation and undermine the rational foundations of democratic discourse.

Viewed through the lens of Fairclough's Critical Discourse Analysis, the findings of this research demonstrate the interconnected relationship between language, algorithmic systems, and social structures in shaping political realities within digital environments. At the textual level, language within the deepfake video functions as a symbolic instrument used to manipulate moral meaning and construct misleading representations of public figures. The phrase "honorary teachers are a burden to the state" serves as a negative framing device designed to provoke moral outrage and public criticism toward political authorities.

At the level of discursive practice, the production and dissemination of the deepfake video are

mediated by digital platforms whose algorithmic structures prioritize emotionally engaging content. This dynamic accelerates the diffusion of disinformation by allowing provocative narratives to reach wider audiences than factual corrections. The circulation of manipulated content therefore reflects the influence of algorithmic amplification in shaping public discourse within social media ecosystems.

At the social practice level, the spread of AI-generated disinformation contributes to a broader erosion of public trust and democratic legitimacy. The widespread acceptance of manipulated information indicates a weakening of institutional credibility, where viral content circulating in digital networks is often perceived as more trustworthy than official communication from governmental institutions. As a result, artificial intelligence technologies increasingly function as new instruments of informational power capable of shaping political perceptions and influencing public opinion.

Taken together, these findings demonstrate that the Sri Mulyani deepfake case reflects the interaction between language, algorithms, and social structures in the production of political meaning within digital environments. Artificial intelligence technologies, which were originally expected to strengthen democratic participation by expanding access to information, can also function as instruments of manipulation capable of undermining public trust and destabilizing democratic legitimacy.

## Conclusion

This study demonstrates that the emergence of artificial intelligence (AI), particularly deepfake technology, has generated a paradox within Indonesia's digital democracy. While digital innovation was expected to enhance access to information and strengthen civic participation, it simultaneously introduces new vulnerabilities by enabling the production of highly realistic disinformation that distorts public understanding and reshapes political narratives. The case of the deepfake video depicting Sri Mulyani Indrawati illustrates how AI-generated content can influence public perception through the strategic combination of language, visual imagery, and emotionally charged narratives. The findings reveal that public reactions were predominantly driven by emotional responses rather than critical verification, while official clarifications circulated more slowly and reached fewer audiences. This imbalance reflects the structural logic of algorithm-driven platforms, which tend to amplify provocative content and marginalize factual corrections, thereby facilitating the rapid diffusion and perceived legitimacy of misleading

information.

More broadly, the study shows that AI-driven disinformation is not merely the product of individual manipulation but is embedded within the structural dynamics of digital communication systems that prioritize engagement and visibility. As a result, traditional epistemic authorities are increasingly challenged by decentralized and algorithmically mediated information flows, contributing to the erosion of public trust and the weakening of democratic legitimacy. When manipulated content becomes widely accepted as credible, the boundary between truth and fabrication becomes blurred, transforming digital democracy into a space dominated by emotional mobilization rather than rational deliberation. Therefore, this study underscores the urgency of strengthening digital literacy, establishing ethical frameworks for AI governance, and enhancing transparency within digital platforms. Ultimately, the sustainability of digital democracy depends on the capacity of institutions and society to ensure that technological innovation reinforces, rather than undermines, democratic values and public trust.

### References

- Andriani, V. W. (2019). Bahasa sebagai representasi demokrasi digital di Indonesia: Analisis wacana kritis. *Jurnal Bahasa dan Sastra*, 3(2), 231–260.
- Aris Sarjito. (2024). Hoaks, disinformasi, dan ketahanan nasional: Ancaman teknologi informasi dalam masyarakat digital Indonesia. *Journal of Governance and Local Politics*, 5(2), 175–186.
- Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, 107(6), 1753–1819.
- Coleman, S., & Blumler, J. G. (2009). *The internet and democratic citizenship: Theory, practice and policy*. Cambridge University Press.
- Firdaus, Y. A. (2025). Analisis wacana kritis terhadap pernyataan Dedi Mulyadi tentang penghapusan wisuda sekolah: Perspektif Norman Fairclough. *RELASI: Jurnal Penelitian Komunikasi*, 5(4), 378–388.
- Floridi, L. (2023). *Ethics of artificial intelligence and governance*. Oxford University Press.
- Nurdin, S. W., & Nugraha, I. F. (2025). Ancaman deepfake dan disinformasi berbasis AI: Implikasi terhadap keamanan siber dan stabilitas nasional Indonesia. *Jurnal Ilmu Manajemen dan Riset (JIMR)*, 4(1), 73–92.
- Santiko, J. A., & Bahri, S. (2024). Analisis wacana pada fenomena penggunaan AI dalam konten

pemilu: Studi kasus deepfake Soeharto. *INNOVATIVE: Journal of Social Science Research*, 4(3), 13215–13231.

- Wardle, C., & Derakhshan, H. (2017). *Information disorder: Toward an interdisciplinary framework for research and policy making*. Council of Europe.
- Zahra, S., Azizan, A., & Sophia, S. (2025). Enhancing electoral integrity: AI's role and challenges in election monitoring. *Jurnal Huma Betang Demokrasi*, 3(1), 19–31.