



## "Hate Speech Communication Patterns in Social Media: A Case Study on Instagram During the 2024 Political Year"

Jefriyanto Saud<sup>1\*</sup>, Ratna Dewanti<sup>2</sup>

<sup>1</sup>Universitas Negeri Gorontalo

<sup>2</sup>Universitas Negeri Jakarta

**Corresponding Author:** Jefriyanto Saud [jefri.saud@ung.ac.id](mailto:jefri.saud@ung.ac.id)

---

### ARTICLE INFO

*Keywords:* Hate Speech, Election Year, Critical Discourse Analysis, Forensic Linguistics, Social Media

*Received :* 21 May

*Revised :* 23 June

*Accepted:* 23 July

©2025 Saud, Dewanti: This is an open-access article distributed under the terms of the [Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



### ABSTRACT

This study looks at trends in hate speech communication on Instagram during the 2024 election year. Four main patterns were identified in this study: linear, spiral, circular, and covert. Information was gathered from recordings, comments, stories, and captions that included hate speech directed at political figures and ideological groups. The findings of the investigation demonstrate that hate speech is not only expressed explicitly but also through memes, disguised narratives, and symbols. The findings demonstrate that hate speech is a discourse about power and symbolic violence that is organized digitally. This research contributes to the advancement of social and linguistic techniques for identifying and curbing hate speech

---

## **INTRODUCTION**

In the digital age, social media has evolved into a dominant discursive space not merely serving as an information sharing platform but also becoming an arena for ideological battles and political expression. Among the most widely used platforms in Indonesia, Instagram holds a strategic position due to its visual nature, interactivity, and rapid content dissemination. Its popularity has transformed it from a channel for self expression into fertile ground for the spread of hate speech, particularly during politically charged periods like the 2024 General Elections. With 89.15 million active users in Indonesia by early 2024, Instagram plays a pivotal role in shaping public opinion and even fueling digitally rooted social conflicts (Sari & Prasetya, 2022; Mustamin et al., 2024).

The phenomenon of hate speech on social media, especially Instagram, has shifted from spontaneous emotional outbursts to structured communication patterns reflecting social fragmentation and political polarization. During elections, digital communication which should ideally serve as a democratic space is instead exploited to propagate negative sentiments based on religious, racial, and ideological identities. This trend is exacerbated by the rise of political buzzers, micro-influencers, and algorithm-driven smear campaigns, which indirectly normalize hate speech in the digital public sphere (Sholihah, 2022; Rizky et al., 2023). Thus, it is crucial to examine hate speech not merely as isolated content but as a recognizable and systematically analyzable communication pattern.

The proliferation of hate speech across social media platforms poses serious threats to democratic quality, social stability, and national cohesion. Within the context of Indonesia's 2024 political year, the nation faces renewed challenges in governing digital spaces as integral components of its democratic ecosystem. Emerging patterns of hate speech communication not only reflect political tensions but also reinforce intolerance, social exclusion, and symbolic violence against targeted groups (Utami & Darmaiza, 2020; Lia, 2021). Without a comprehensive understanding of how hate speech is constructed, disseminated, and received by the public, policy interventions and digital literacy efforts remain largely reactive and ineffective.

The urgency of this study lies in expanding academic inquiry beyond mere content analysis to examine communicative structures. By deconstructing hate speech communication patterns, key societal actors – including government institutions, media organizations, educators, and civil society can develop more effective preventive strategies. This research will contribute significantly to applied linguistics and political communication literature in Indonesia, particularly in the realm of digital discourse analysis. The focus on Instagram proves particularly salient, as the platform has evolved into a primary battleground for symbolic political warfare through comments, reels, and narrative captions (Putri et al., 2024; Hanum, 2024).

Research on hate speech in social media has expanded across multiple disciplines, spanning linguistics, communication studies, social psychology, and legal studies. However, most studies remain focused on identifying explicit slurs or offensive phrases, failing to examine how hate speech operates as part of broader communication patterns. Widyatnyana et al.'s (2023) research, for instance, reveals that social media hate speech often employs sarcasm, political rhetoric, and cultural references that lexical analysis alone cannot capture. This context underscores the importance of critical discourse analysis in unpacking hate speech as a discursive strategy intertwined with power dynamics and ideology.

Instagram presents unique characteristics as a research subject. Unlike Twitter's text-dominant format, Instagram features multimodal content—images, text, and audio creating novel dynamics in hate speech representation. Studies by Rahmawati (2020) and Mustamin et al. (2024) demonstrate how features like Stories and Reels effectively disseminate ephemeral yet high impact hate narratives due to their viral potential. This makes Instagram a rich digital ecosystem for linguistic and sociopragmatic analysis. Yet, few studies have specifically investigated hate speech communication patterns across Instagram's features during election periods, despite their layered complexity and implicit meanings.

Previous studies in national publications have highlighted that election-related hate speech predominantly manifests through identity politics narratives, negative labeling, and covert disinformation targeting public figures and minority groups. For instance, research by Kusumasari & Arifianto (2020) demonstrates how memes and captions frequently serve as effective vehicles for disguising hatred through humor and pop culture references, making hate messages appear innocuous while retaining their ideological impact.

Furthermore, studies emphasize how social media algorithms and viral dynamics accelerate hate speech dissemination, shaping public perceptions of truth and morality. Rizky et al.'s (2023) research reveals that hate speech in Instagram comments often originates as reactive expressions but rapidly evolves into collective discourse through likes, shares, and affirmative replies. This pattern indicates that hate speech does not spread randomly but rather follows strategic communication patterns aligned with specific social functions and objectives.

The research gap in integrated studies examining hate speech communication patterns across Instagram's features represents a critical academic void that this study addresses. Existing literature predominantly focuses on content analysis and psychological impacts, with limited systematic investigation into discourse structures, communicative intentions, and production contexts across Instagram's various features Reels, Stories, and

comments. This study offers conceptual and methodological contributions by integrating two complementary approaches: critical discourse analysis and forensic linguistics.

Focusing on the 2024 political period, this research explores emerging hate speech communication patterns on Instagram. It specifically examines how user interactions form recognizable communicative structures – such as linear, spiral, circular, or echo chamber patterns – while analyzing the distinct roles of Instagram's features (comments, captions, Reels, and Stories) in facilitating both explicit and implicit hate speech. The study investigates three key dimensions:

1. **Message Strategies:** How hate speech is linguistically constructed and framed
2. **User Relational Dynamics:** The interplay between content creators, amplifiers, and targets
3. **Feature Exploitation:** The systematic use of platform tools to disguise or amplify hate messages

This study investigates the complex dynamics of hate speech on Instagram during Indonesia's 2024 political year through two primary research questions. First, it examines the distinctive communication patterns emerging in hate speech exchanges, seeking to identify and categorize both overt and covert forms of digital interaction. Second, the research analyzes how hate speech manifests differently across Instagram's various features - including comments, captions, Reels, and Stories - with particular attention to the interplay between visual and verbal elements in constructing harmful discourse. These questions are designed to uncover not only linguistic patterns but also the underlying social practices and power dynamics that characterize political hate speech on digital platforms.

The investigation employs an innovative methodological synthesis of Critical Discourse Analysis and Forensic Linguistics to address these questions comprehensively. This dual approach enables the study to move beyond surface-level content analysis to reveal how hate speech operates as both a linguistic phenomenon and a social practice embedded in Indonesia's unique digital-political landscape during election periods. The research ultimately aims to contribute new theoretical insights into digital discourse analysis while offering practical implications for platform governance and hate speech mitigation strategies in politically charged contexts.

## LITERATURE REVIEW

### 1. Critical Discourse Analysis (CDA)

Critical Discourse Analysis (CDA) is an interdisciplinary approach to discourse studies that examines the relationship between language, power, and ideology in society. CDA conceptualizes language not merely as a communicative tool but as a social practice that significantly influences the construction and reproduction of power structures and social inequalities (Fairclough, 1995; van Dijk, 1993). In the context of hate speech on social media, CDA serves as an effective framework for deconstructing how such speech is shaped not only by lexical choices or sentence structures but also by underlying socio-political and ideological contexts.

Van Dijk (1998) argues that hate speech in public discourse is often constructed through macro- and micro-linguistic strategies, including lexical selection, narrative framing, "us vs. them" polarization, and the reinforcement of stereotypes against specific groups. A hallmark of CDA is its three-level analytical framework: (1) text (linguistic structures), (2) discursive practice (how discourse is produced and consumed), and (3) social practice (ideological and socio-political contexts). On platforms like Instagram, CDA proves valuable for analyzing how hate speech manifests across features—captions, comments, Reels, and Stories—which are not isolated but interconnected with power agendas, political polarization, and the mobilization of negative sentiments (Fairclough, 2001; van Dijk, 2006).

In the context of Indonesia's 2024 political year, CDA is employed to investigate how hate-driven content functions as part of digital political communication strategies. Hate speech targeting candidates, parties, or political supporters is often framed as "defending national values," "anti-radicalism," or "nationalism," while covertly perpetuating social exclusion and power ideologies (Widyatnyana et al., 2023). CDA enables researchers to trace embedded power dynamics and resistance in such discourse, particularly when hate communication shifts from explicit and vulgar expressions to symbolic, implicit forms disseminated through reposts, memes, and serial Stories.

### 2. Forensic Linguistics and the Language of Violence

Forensic linguistics, an applied branch of linguistics, examines language use in legal and criminal contexts—including hate speech, verbal threats, and digital defamation. Olsson (2008) asserts that hate speech can be analyzed as verbal evidence containing violent or discriminatory intent against specific groups based on race, religion, gender, ideology, or political affiliation. This approach scrutinizes sentence structures, lexical choices, pragmatic intent, and the social consequences of such speech.

A defining feature of forensic linguistics is its focus on utterance function—evaluating both the speaker's intent and the message's impact on recipients (Shuy, 2001). In hate speech analysis, this involves determining whether a statement constitutes defamation, incitement, or marginalization of targeted groups.

On Instagram, hate speech transcends explicit text, manifesting through:

1. Ambiguous captions
2. Irony-laden Reels
3. Symbolically offensive imagery in Stories
4. Coded emoji usage

Unlike purely semantic analysis, forensic linguistics interrogates social function and legal implications, assessing how language weaponizes discrimination while evading platform moderation.

During election periods, political hate speech increasingly adopts covert linguistic strategies utilizing euphemisms, irony, sarcasm, and metaphors to circumvent algorithmic detection and platform moderation. Olsson (2008) conceptualizes such discourse as a form of symbolic violence that systematically excludes and stigmatizes political opponents. This study bridges Critical Discourse Analysis (CDA) and forensic linguistics to create an integrated analytical framework. Where CDA reveals the ideological structures and power dynamics embedded in hate speech, forensic linguistics provides systematic methods for examining linguistic evidence, including trigger words, syntactic patterns of threats, and pragmatic intent. Together, these approaches offer both theoretical depth and practical utility for legal interventions and digital policy formulation (Shuy, 2010).

The combined CDA-forensic linguistics approach proves particularly vital for analyzing Instagram's unique ecosystem. As a multimodal platform characterized by rapid virality and algorithmic fragmentation, Instagram hosts hate speech that operates through symbolic narratives (e.g., nationalist memes concealing xenophobia), manipulative captions with covert calls to action, and decentralized yet collectively harmful comment threads. CDA deconstructs how these elements form ideological discourses, while forensic linguistics examines their capacity to inflict verbal harm and potentially violate legal boundaries. This dual perspective not only advances academic understanding of digital political communication but also informs more nuanced platform governance approaches capable of detecting and addressing both explicit and implicit forms of hate speech.

## METHODOLOGY

This study employs a qualitative descriptive approach with a case study design, selected for its capacity to provide in-depth, contextual examination of hate speech communication patterns on social media platforms, particularly Instagram, during the 2024 political year. The qualitative methodology enables researchers to explore the underlying meanings of hate speech and reveal how such discourse is constructed, disseminated, and perpetuated through various social media features. Theoretically, the study is grounded in Critical Discourse Analysis (CDA) as developed by Teun A. van Dijk and Norman Fairclough, complemented by Forensic Linguistics, particularly the frameworks of Olsson (2008) and Shuy (2001) regarding violent language as a form of discriminatory speech and verbal aggression.

The research is conducted online, with Instagram serving as the primary case study platform. Instagram was chosen due to its status as one of Indonesia's most popular social media platforms, especially during political campaigns and electoral contests such as the 2024 General Elections. The research timeframe spans January to June 2024, covering the campaign period, election day, and post-election phase. All data are collected from publicly available posts on the platform, with no researcher intervention or content modification.

This study utilizes two main types of data. Primary data consists of public Instagram content—including captions, comments, Stories, and Reels—containing indicators of hate speech or verbal aggression, with a focus on politically influential accounts such as public figures, politicians, political parties, digital media outlets, and high-engagement buzzers. Secondary data comprises news articles, academic literature, and digital hate speech regulations to provide broader socio-political context.

Data collection follows three systematic phases. First, account and content identification through tracking political hashtags and key events like presidential debates, election results, or trending issues. Second, manual data capture via screenshots, comment transcripts, and relevant Reel/Story downloads. Third, metadata documentation including upload timestamps, like/comment counts, and supporting narratives for in-depth analysis. All data is systematically archived using only public content without user interaction, ensuring research ethics compliance.

The collected data was analyzed using two primary approaches. First, Critical Discourse Analysis (CDA) was employed to examine the structure of hate speech, discourse strategies, and the relationship between language and power evident in digital texts. In this approach, van Dijk's model was used to analyze micro and macro discourse structures, while Fairclough's framework was applied to investigate the connections between text, discursive practices, and social practices. Second, Forensic Linguistics was utilized to detect both explicit

and implicit elements of verbal violence within hate speech. This analysis focused on word choice, speech intent, sentence structure, and the social and symbolic impact of the utterances, following the methodologies recommended by Olsson and Shuy.

## RESULTS AND DISCUSSION

### 1. Hate Speech Communication Patterns on Instagram During the 2024 Political Year

Analysis reveals that hate speech on Instagram during the 2024 political year exhibited recurring communication patterns, linguistically and pragmatically reflecting users' discursive strategies and ideological frameworks. Examination of 15 public accounts and 300 political posts (January–March 2024) identified four dominant communication patterns:

#### a. Linear Pattern

Hate speech is delivered unidirectionally, typically as direct comments on political posts. For example, explicit attacks like "*Stupid cebong [pro-government supporters] – penghianat negara!*" emerge without discursive context. Van Dijk's (1993) framework classifies this as a discourse macrostructure that reinforces an "*kami*" (*ingroup*) vs. "*mereka*" (*outgroup*) dichotomy. **Description:** One-way verbal attacks, often manifesting as crude comments or insults on political content.

**Data (Instagram Comment):** "*Dasar cebong otak udang! Hidup makan bansos, nyembah pejabat korup!*"

**Analysis:** This comment exemplifies explicit stereotyping through Derogatory labels: "*cebong*" (pro-government supporters) and "*otak udang*" reflect ideological dehumanization. Moral bifurcation: Van Dijk's CDA reveals how such rhetoric constructs '*kami*' ("moral citizens") versus '*mereka*' ("ignorant masses").

#### b. Spiral Pattern

Hate speech triggers reciprocal responses, generating increasingly hostile comment chains that intensify political identity conflicts. This discourse demonstrates meaning escalation and the expansion of discriminatory narratives, revealing how hate speech becomes a recurrent social practice that is continuously reinterpreted.

**Description:** Occurs when hate speech prompts counter-responses from other users, creating aggressive debate threads that exacerbate socio-political divisions.

**Data (Thread comment):**

1. **User A:** "*Mampuslah kalian para kadrin, nggak akan pernah menang di republik ini!*"
2. **User B:** "*Bilang ke cebong penjilat asing, kalian yg jual negara!!*"
3. **User C:** "*Sama aja dua-duanya, cuma bisa ribut doang tiap pemilu!*"

**Analysis:** The spiral pattern in hate speech demonstrates an escalation of hostility that occurs through reciprocal attacks between users, where each interaction progressively intensifies the level of animosity. From a forensic pragmatic perspective, this interaction pattern reveals three key characteristics:

First, there is emotional amplification characterized by increasing intensity of aggressive language, ranging from ordinary criticism to explicit threats, along with progressively harsher dehumanizing metaphors in each response. Second, symbolic threat construction emerges through the transformation of individual slurs into collective identity attacks, including the development of coded violence such as the phrase "They deserve to be taught a lesson" which carries implicit threats. Third, collective hate normalization occurs where initially individual expressions evolve into sustained social practices. This chain reaction is reinforced by algorithmic visibility that continually sharpens the boundaries between "us" and "them" groups.

Theoretically (Shuy, 2002; Olsson, 2008), this pattern proves that hate speech constitutes a form of interactional violence - where meaning and damaging impact are collectively constructed through networked participation, rather than stemming from singular utterances. These findings highlight the importance of content moderation approaches that consider conversational contexts holistically, rather than focusing solely on individual comments.

### c. Circular Pattern

The content is redistributed through Instagram's and features, frequently accompanied by additional narratives or symbols that reinforce the ideological stance of the sharer. This representation not only expands the reach of hate speech but also strengthens digital echo chambers that entrench intergroup hostility.

#### Description

Hate speech is recirculated via or features, supplemented with antagonistic narratives that either vilify opponents or bolster the ideological position of one's own group.

#### Data (Story Instagram):

One of the account's Story @gerakpolitika2024 showed caption the replied:

*"Kalau bukan kita yang ubah, siapa lagi?"*

Reposted by an account @anti\_kadrun\_id with additional text:

*"Iya, ubah jadi negara khilafah kalian maksudnya? #TolakPengkhianatNegara"*

**Analysis:** Hate speech is strategically manipulated through the reframing of political opponents' captions. By adding new contextual layers, the original posts are transformed into identity-based attacks. This pattern reveals the strategic function of reposting as a tool for *reproducing hate speech* in viral, collective formats

d. **Covert Pattern**

Hate communication is conveyed through political memes, ambiguous captions, and degrading symbols or emojis (e.g., combinations of pig emojis + party flags). Within the forensic linguistics framework, this pattern constitutes speech containing euphemistic or discriminatory metaphors, which are legally challenging to prosecute yet carry symbolic violent intent (Olsson, 2008).

**Description:** A form of hate speech disguised through sarcastic tones, humor, metaphors, or visual symbols (emojis, stickers, memes), making it difficult to categorize as explicit speech.

**Data (Instagram Meme Caption):**

*Meme:* Photo of a presidential candidate laughing

*Caption:* "When the president is a clown, don't be surprised if the country becomes a circus 🤡🐷 ID #RIPlogic"

**Analysis:** This caption employs offensive metaphors and symbols (clown and pig emojis) that carry culturally derogatory meanings. From a forensic linguistics perspective, it constitutes covert speech containing symbolic violent intent. This strategy evades automated detection while effectively conveying degrading messages about specific individuals or groups.

These patterns demonstrate that social media hate speech has evolved beyond explicit text into systematic communication practices integrated with platform features. Theoretically, these findings reinforce CDA's position that hate speech operates as Power discourse embedded in specific social contexts and Ideological tools for group representation and marginalization (Fairclough, 1995; van Dijk, 1998)

2. **Manifestations of Hate Speech Across Instagram Features**

The findings demonstrate how various Instagram features facilitate the dissemination of hate speech through distinct discursive patterns. Comments emerge as the most direct channel, characterized by brief yet virulent expressions laden with stereotypes and derogatory labels like "kadrung anti-Pancasila" (a slur targeting conservative Muslims) or "nasbung penjilat asing" (pejorative for nationalist elites). These exemplify classic markers of verbal violence - labeling, dehumanization, and provocation - often meeting forensic criteria for harassment or threats.

Captions adopt more sophisticated rhetorical strategies, employing extended narratives that manipulate discourse through ideological framing. Typical techniques include rhetorical questions questioning opponents' legitimacy ("Do you still trust this person?"), sovereignty-themed alarmism ("threat to national sovereignty"), and conspiratorial metaphors ("neo-colonialism"). Such constructions reflect deliberate attempts to shape public perception through discourse-level persuasion rather than overt slurs.

The ephemeral Story feature has evolved into a potent tool for viral shaming and conflict escalation. Users commonly repurpose opponents' posts by adding inflammatory stickers, emojis, or captions ("so everyone sees who you really are"), or share screenshots of contentious comments with amplifying commentary. While temporary, these curated Story narratives effectively reinforce in-group solidarity while broadening the reach of interpersonal conflicts. This feature-specific analysis reveals how platform architecture influences hate speech modalities from comments' immediacy to captions' persuasive framing and Stories' viral shaming dynamics. The patterns underscore the need for equally nuanced moderation approaches tailored to each feature's communicative affordances and potential harms.

Reels, represent a multimodal form of hate speech, combining short video content with visual, audio, and textual elements. Examples identified in this study include edited political debate clips paired with sarcastic background music or derogatory captions targeting specific candidates. Both multimodal Critical Discourse Analysis (CDA) and forensic linguistic approaches can be synergistically applied to decode hidden messages and potential manipulation within such content.

These findings confirm that hate speech on Instagram has evolved beyond spontaneous expressions into a strategic form of political communication. It systematically leverages the platform's full range of features to construct dominant narratives, reinforce ideological positions, and attack opponents through both linguistic and visual means. The study highlights the complementary strengths of CDA and forensic linguistics. While CDA reveals the embedded power structures and ideologies within discourse, forensic linguistics provides a framework for analyzing the linguistic and pragmatic markers of verbal violence. In the digital political sphere, this dual approach not only identifies hate speech but also offers new insights into its construction, dissemination, and societal impact.

Ultimately, the research demonstrates that hate speech constitutes a complex socio-linguistic phenomenon requiring: Interdisciplinary analysis combining discourse and forensic methods, Regulatory measures tailored to platform-specific communication patterns, and Critical media literacy initiatives to empower users against manipulation

## **CONCLUSIONS AND RECOMMENDATIONS**

This study reveals that hate speech circulating on Instagram during Indonesia's 2024 political year cannot be dismissed as mere spontaneous outbursts or individual expressions of emotion. The findings demonstrate that such hate speech follows distinct, recurring communication patterns that reflect deeper ideological divisions and political polarization in digital spaces. Specifically, four key patterns emerged: linear patterns involving direct personal attacks, spiral patterns marked by escalating comment threads, circular patterns characterized by reposted content with amplified narratives, and covert patterns employing symbolic, sarcastic, or metaphorical language. These patterns collectively indicate that hate speech on social media operates as a deliberate discursive strategy—systematically produced and reinforced within digital ecosystems, especially during periods of political competition.

Applying Critical Discourse Analysis (CDA), this research frames these patterns as manifestations of power dynamics that perpetuate identity-based conflicts, social exclusion, and ideological manipulation of public perception. Hate speech functions not only as a tool for personal attacks against political adversaries but also to construct exclusionary nationalist narratives that marginalize certain groups. Complementing this perspective, Forensic Linguistics analysis uncovered the linguistic markers of verbal violence embedded within these communications, including dehumanizing labels, harmful stereotypes, implicit threats, and other forms of derogatory speech.

The integration of CDA and Forensic Linguistics provides a comprehensive framework for analyzing hate speech across Instagram's various features—comments, captions, Stories, and Reels while also illuminating its broader societal consequences. These findings underscore the urgent need for multi-faceted interventions, including platform-specific content moderation strategies that account for the unique ways hate speech manifests in different features, legal reforms that address evolving digital hate tactics, and media literacy initiatives to help users recognize and resist manipulative discourse.

Ultimately, this study positions political hate speech as a complex, networked phenomenon rather than isolated incidents, demanding interdisciplinary solutions that bridge linguistic analysis, platform governance, and policy development to mitigate its harmful effects on democratic discourse. Overall, this study confirms that digital hate speech during political periods is no longer just a matter of communication ethics, but rather a form of power discourse with linguistic, ideological, and even forensic dimensions. The strategic use of social media features to spread hate speech demonstrates that platforms like Instagram are not neutral spaces, but can become mediums for social conflict production if not wisely regulated and controlled. Therefore,

understanding these hate speech patterns is crucial not only for academic purposes but also for public policy, legal protection, and strengthening society's digital literacy.

For academic recommendations, hate speech research on social media should be expanded through interdisciplinary approaches combining applied linguistics, digital communication, political sociology, and information technology. Future studies should develop multimodal analyses examining textual, visual, audio, and symbolic elements in hate speech while mapping their correlation with social media algorithms. Longitudinal research is particularly important to track the evolution of hate speech over time, especially during Indonesia's political cycles, to contribute to a national digital hate speech discourse mapping.

For policy recommendations, these findings highlight the need for: Strengthening digital regulations that address not just explicit violations but also covert, symbolic, and ideologically charged hate speech, revising digital content laws to incorporate linguistic analysis and social context as violation indicators, and enhancing collaboration between government, the Ministry of Communication (Kominfo), social media platforms, and independent institutions to develop more adaptive monitoring systems based on forensic linguistic research

Third, for platform providers such as Instagram, YouTube, and other social media companies, it is crucial to develop more sophisticated content moderation systems. This includes implementing advanced natural language processing (NLP) and artificial intelligence (AI) technologies capable of detecting hate speech not only in explicit forms but also within narrative structures and symbolic expressions. Strengthening community guidelines and actively engaging users in reporting hate content are equally critical measures. By doing so, these platforms can evolve beyond being mere technology service providers to becoming stewards of healthy, democratic digital public spaces.

Fourth, for the public, particularly active social media users, there is an urgent need to enhance critical and reflective digital literacy. Users must develop the ability to recognize hate speech, understand the dangers of reproducing harmful content, and practice inclusive digital communication ethics. Digital literacy education should extend beyond technical safety measures to include discourse analysis, ideological awareness, and social responsibility in online spaces. Through such comprehensive education, digital citizens can become proactive agents in countering the spread of hate speech - both during political cycles and in everyday online interactions.

## **FUTURE RESEARCH**

This study was limited to examining Instagram during the early stages of Indonesia's 2024 political year, with inherent constraints regarding timeframe and data sources. Additionally, it did not comprehensively investigate the potential relationship between digital hate speech and real-world (offline) behavioral consequences.

To develop a more comprehensive and systematic understanding of hate speech dissemination, several critical research expansions are necessary. First, studies should broaden their platform coverage beyond Instagram to encompass other major social media ecosystems including TikTok, X (Twitter), and YouTube, as each platform features distinct communication patterns and user demographics that may influence hate speech propagation. Methodologically, researchers should incorporate quantitative approaches such as social network analysis to precisely map how hate speech spreads through digital networks and identify key amplification nodes.

A particularly urgent research need involves investigating platform mechanisms through two parallel tracks: evaluating the effectiveness of different content moderation policies across platforms, and conducting rigorous analyses of how recommendation algorithms influence hate speech dynamics through content prioritization and amplification.

Three key research gaps require prioritized attention: comparative cross-platform studies examining how hate speech manifests differently across various media environments; longitudinal studies tracking the evolution of hate speech through complete electoral cycles; and empirical studies measuring the concrete linkages between online hate speech and real-world violent incidents. Addressing these gaps would significantly advance both theoretical understanding and practical interventions for mitigating digital hate speech's societal impacts.

## REFERENCES

- Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. London: Longman.
- Fairclough, N. (2001). *Language and power* (2nd ed.). London: Longman.
- Gabur, I. (2023). *Hate speech and digital power: A discourse approach to online political conflict*. Routledge.
- Hanum, F. (2024). *Wacana politik dalam media sosial dan strategi kampanye digital*. Jakarta: LP3M Press.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). London: Routledge.
- Kusumasari, B., & Arifianto, R. (2020). Meme sebagai bentuk ujaran kebencian: Analisis wacana dalam media digital. *Jurnal Komunikasi dan Media Digital*, 5(2), 123–137. <https://doi.org/10.25077/jkmd.v5i2.123>
- Lia, A. F. (2021). Politik identitas dan ujaran kebencian dalam media sosial. *Jurnal Ilmu Sosial Politik dan Pemerintahan*, 8(1), 34–47. <https://doi.org/10.22219/jispg.v8i1.11876>
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- Mustamin, A., Rahmawati, D., & Kartika, H. (2024). Praktik ujaran kebencian dalam fitur story Instagram: Studi kasus akun buzzer politik. *Jurnal Sosiolinguistik Digital*, 2(1), 45–61.
- Olsson, J. (2008). *Forensic linguistics: An introduction to language, crime and the law* (2nd ed.). London: Continuum.
- Papacharissi, Z. (2010). *A private sphere: Democracy in a digital age*. Polity Press.
- Putri, R. A., Dhamayanti, F., & Nurhasanah, S. (2024). Kampanye digital dan konten ideologis dalam Instagram politik. *Jurnal Politik dan Teknologi*, 6(1), 19–34.
- Rahmawati, N. (2020). Representasi ujaran kebencian berbasis agama di media sosial: Analisis kritis pada konten Instagram. *Jurnal Studi Media dan Agama*, 4(2), 75–89.

- Rizky, A., Prabowo, D., & Yuliana, D. (2023). Polarisasi politik dalam komentar Instagram: Kajian pragmatik ujaran kebencian. *Jurnal Bahasa dan Politik*, 3(1), 56–68.
- Sari, M. I., & Prasetya, H. (2022). Interaksi pengguna dan distribusi kebencian dalam platform digital: Studi kasus Instagram. *Jurnal Komunikasi Online*, 9(3), 188–204.
- Shuy, R. W. (2001). *Linguistic battles in trademark disputes*. New York: Palgrave Macmillan.
- Shuy, R. W. (2010). *The language of defamation cases*. Oxford: Oxford University Press.
- Silverman, D. (2020). *Interpreting qualitative data* (6th ed.). London: Sage.
- Sumpter, R. S. (2021). *Understanding hate speech: A multidimensional approach to speech that divides*. Palgrave Macmillan.
- Utami, W., & Darmaiza, E. (2020). Literasi digital dan ujaran kebencian pada masa pemilu: Sebuah studi naratif. *Jurnal Komunikasi Sosial*, 7(1), 1–15. <https://doi.org/10.21009/jks.v7i1.234>
- van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & Society*, 4(2), 249–283. <https://doi.org/10.1177/0957926593004002006>
- van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. London: Sage Publications.
- van Dijk, T. A. (2006). Discourse, context and cognition. *Discourse Studies*, 8(1), 159–177. <https://doi.org/10.1177/1461445606059566>
- Widyatnyana, K. M., Susanti, E., & Amalia, N. (2023). Ujaran kebencian dan identitas politik dalam pemilu digital: Pendekatan analisis wacana kritis. *Jurnal Linguistik Terapan*, 9(2), 101–115.
- Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: Sage Publications.