



Human Rights Enforcement in the Post-Truth Era: Digital Discourse of #IndonesiaHumanRightsSOS on Twitter

Michael Angello Qadosy Riyadi

Direktorat Kampus Surabaya, Telkom University, Indonesia

Email: michaelangelo@student.telkomuniversity.ac.id

A'isyah Belqis Febi Aulia

Direktorat Kampus Surabaya, Telkom University, Indonesia

Email: aisyahbelqis@student.telkomuniversity.ac.id

M. Asif Nur Fauzi

STEBI Syaikhona Kholil Sidogiri, Indonesia

Email: asif.elek24@gmail.com

**Corresponding Author*

Article History

Submitted : 2026-04-02

Accepted : 2026-04-16

Revised : 2026-04-15

Published : 2026-06-18

DOI: <https://doi.org/10.31571/jpkn.v10i1.10681>

Abstract

The proliferation of social media has transformed how Human Rights (HR) issues are articulated, contested, and politicized within digital public spheres. This study examines the discourse surrounding the hashtag #IndonesiaHumanRightsSOS on Twitter in December 2020 as a public response to the alleged extrajudicial killing of six members of the Front Pembela Islam (FPI) by Indonesian police at KM50 of the Jakarta–Cikampek Toll Road. By analyzing 106,903 tweets from more than 45,000 accounts within a 72-hour period, this qualitative-descriptive study employs digital discourse analysis (DDA) to identify discourse forms, human rights enforcement challenges, and post-truth characteristics. The findings reveal the dominance of demands for justice and accountability, expressions of public solidarity, and condemnation of state actors. Key challenges include impunity, procedural injustice, institutional trust crisis, and politicization. This study contributes to the advancement of digital discourse analysis in the post-truth context in Indonesia and underscores the importance of strengthening critical digital literacy in Civic Education.

Keyword: Human Rights Discourse, Digital Discourse Analysis, Post-Truth Era, Twitter, Indonesia

Abstrak

Perkembangan media sosial telah mengubah cara isu hak asasi manusia (HAM) diartikulasikan, diperdebatkan, dan dipolitisasi dalam ruang publik digital. Penelitian ini mengkaji diskursus tagar #IndonesiaHumanRightsSOS di Twitter pada Desember 2020 sebagai respons atas dugaan pembunuhan di luar proses hukum terhadap enam anggota Front Pembela Islam (FPI) oleh aparat kepolisian di KM50 Tol Jakarta–Cikampek. Dengan menganalisis 106.903 tweet dari lebih dari 45.000 akun dalam 72 jam, studi kualitatif-deskriptif ini menggunakan digital discourse analysis (DDA) untuk mengidentifikasi bentuk diskursus, tantangan penegakan HAM, serta karakteristik post-truth. Hasil menunjukkan dominasi tuntutan keadilan dan akuntabilitas, solidaritas publik, serta kecaman terhadap aktor negara. Tantangan utama meliputi impunitas, ketidakadilan prosedural, krisis kepercayaan institusional, dan politisasi. Penelitian ini berkontribusi pada pengembangan analisis diskursus digital dalam konteks post-truth di Indonesia serta menegaskan pentingnya penguatan literasi digital kritis dalam pembelajaran Pendidikan Kewarganegaraan.

Kata Kunci: Diskursus Hak Asasi Manusia, Analisis Wacana Digital, Era Post-Truth, Twitter, Indonesia



INTRODUCTION

Human Rights (HR) constitute the normative foundation of democratic governance and the rule of law. In Indonesia, the constitutional commitment to HR protection is enshrined in Articles 28A–28J of the 1945 Constitution and operationalized through Law No. 39 of 1999 on Human Rights. Yet persistent structural obstacles—including institutional impunity, weak accountability mechanisms, and an eroding public trust in law enforcement—continue to frustrate effective HR enforcement. These challenges are inseparable from the broader dynamics of public opinion, political polarization, and the evolving media ecology that characterizes contemporary Indonesian democracy (Mietzner and Muhtadi 2022; Hadiz 2021).

The digital revolution has irrevocably altered the architecture of public discourse. Twitter/X, as a globally significant microblogging platform, has become a primary arena for opinion formation, political contestation, and social mobilization (González-Bailón et al. 2023; Freelon 2021). In Indonesia, politically charged hashtags frequently serve as real-time barometers of public sentiment toward high-profile controversies involving state actors. These digital spaces both reflect and amplify pre-existing societal tensions, making them fertile sites for examining the intersection of HR discourse, democratic participation, and the contemporary information ecology (Lim 2021).

December 2020 witnessed the eruption of #IndonesiaHumanRightsSOS as one of the most significant HR-related digital mobilizations in Indonesian Twitter history. The precipitating event was the fatal shooting of six FPI paramilitary members (Laskar FPI) by the National Police (Polri) at KM50 of the Jakarta - Cikampek Toll Road on the night of December 6–7, 2020. Police maintained the shootings were acts of lawful self-defense; FPI leadership and substantial public segments alleged extrajudicial execution (Noeyislearning 2021). Within 72 hours, the hashtag generated over 106,000 tweets, trending globally and attracting engagement from international HR organizations.

This phenomenon unfolds against the backdrop of the 'post-truth era'- a condition in which public opinion formation is more powerfully driven by emotional appeals, moral intuitions, and viral narratives than by evidence appraisal (Lewandowsky and van der Linden 2021; Ecker et al. 2022). In this epistemic environment, HR issues—which normatively demand careful deliberation and evidence evaluation—become susceptible to simplification, polarization, and political instrumentalization. Viral mechanisms reward emotional resonance over factual precision (Pennycook and Rand 2021; Pennycook et al. 2021), creating conditions in which post-truth dynamics may simultaneously amplify and distort legitimate HR grievances.

A burgeoning literature addresses social media misinformation and political polarization (Lorenz-Spreen et al. 2023; Nyhan et al. 2023; Allen et al. 2021), and a parallel scholarship

addresses Indonesian digital politics (Lim 2021; Mietzner and Muhtadi 2022). However, existing studies largely emphasize general content dynamics and broad patterns of online behavior, with limited attention to human rights discourse analyzed through a post-truth lens. In particular, the intersection between human rights enforcement challenges and post-truth communicative dynamics remains underexplored, especially within the context of Twitter-based digital mobilization in Global South settings. Addressing this gap, the present study analyzes the #IndonesiaHumanRightsSOS discourse using a Digital Discourse Analysis (DDA) approach to systematically examine how human rights narratives are constructed, contested, and amplified within a high-intensity post-truth digital environment.

This study is guided by three research questions (RQs): (RQ1) How is public discourse about human rights represented within the #IndonesiaHumanRightsSOS corpus? (RQ2) What HR enforcement challenges are manifested in digital public opinion in the post-truth era? (RQ3) How do post-truth characteristics manifest in public discourse about HR on social media? The study's objectives are correspondingly threefold: to describe dominant HR discourse forms; to identify enforcement challenges reflected in public opinion; and to analyze the relationship between HR discourse and post-truth dynamics. Theoretically, the study contributes a post-truth analytical framework applicable to HR discourse in Global South digital spheres; practically, it provides an evidence-based foundation for integrating critical digital literacy, discourse evaluation skills, and post-truth awareness into Civic Education curricula at the higher education level, thereby supporting the development of students as critical, informed, and responsible “Smart and Good Citizens” in Indonesia’s digital democracy.

METHOD

Research Design and Epistemological Stance

This study employs a qualitative-descriptive design grounded in constructivist epistemology. The analytical focus is on meaning-making-how actors construct, contest, and circulate interpretations of HR events-rather than on causal inference or hypothesis testing. Digital Discourse Analysis (DDA) serves as the primary methodological framework, combining systematic textual analysis with contextual interpretation of the socio-political conditions shaping the discourse (Freelon 2021). This design is appropriate because the research questions concern the character, thematic structure, and ideological dimensions of discourse, not its statistical distribution.

Data Source and Corpus Characteristics

The dataset was sourced from the publicly available Kaggle repository 'Twitter Data on IndonesiaHumanRightsSOS' (Noeyislearning 2021). Data were collected via the Twitter

Academic Research API employing keyword filtering for #IndonesiaHumanRightsSOS across December 18–20, 2020 (UTC+7). Table 1 summarizes dataset characteristics.

The use of secondary Twitter data is methodologically appropriate given the study’s focus on a discrete, temporally bounded episode of digital HR discourse. As noted by Freelon (2021), hashtag-based corpora represent purposive, event-contingent communicative assemblages that are analytically self-contained and do not require external representativeness.

Table 1. Characteristics of the #IndonesiaHumanRightsSOS Dataset

Parameter	Value / Description
Total tweets	106,903 tweets
Collection window	December 18–20, 2020 (72 hours, UTC+7)
Platform	Twitter/X — Academic Research API
Unique accounts	> 45,000 individual user accounts
Ethical classification	Publicly posted content; no identifiable private data

Data Preprocessing

Raw tweet text underwent a four-stage preprocessing pipeline: (1) text normalization lowercasing, removal of URLs, @mentions, and non-alphanumeric characters except apostrophes; (2) whitespace normalization and de-duplication of verbatim retweets; (3) stopword removal using a combined Indonesian–English stopword lexicon (1,247 tokens); and (4) tokenization and frequency ranking to identify the 100 most common substantive terms. After preprocessing, the effective analytical corpus comprised 78,412 content-bearing tweets. In this study, individual tweets are treated as the primary units of analysis, representing discrete expressions of public opinion within the digital discourse.

Analytical Framework and Coding

The analysis employed four complementary layers (Table 2), each addressing one or more RQs and collectively ensuring systematic coverage of analytical objectives. Post-truth indicator scoring followed operationalizations derived from Lewandowsky and van der Linden (2021) and Ecker et al. (2022).

The analytical process was conducted iteratively across the four layers. Lexical analysis was first used to identify dominant vocabularies, followed by inductive thematic coding to group recurring patterns into coherent discourse categories (Ahmed et al. 2025). Each tweet in the subsample was then evaluated using the post-truth indicator rubric (0–4 scale), and scores were aggregated to construct composite indicator profiles (Krippendorff 2019). Finally, close reading was applied to selected high-engagement tweets to provide contextualized interpretation and to capture the nuanced interplay between language, emotion, and political meaning.

Table 2. Four-Layer Analytical Framework

Layer	Technique	RQ Addressed	Key Output
1	Lexical frequency analysis	RQ1	Dominant terms; thematic vocabulary map
2	Inductive thematic clustering (manual coding)	RQ1, RQ2	Six discourse form categories; enforcement challenge taxonomy
3	Post-truth indicator scoring (0–4 rubric per tweet)	RQ3	Composite indicator radar profile (0–100 scale)
4	Substantive discourse analysis (close reading)	RQ1–RQ3	Qualitative interpretation via representative tweets

For Layers 2 and 4, a purposive subsample of 500 tweets was selected based on the highest engagement scores (user interaction metrics such as likes, retweets, and replies) (combined likes and retweets), as these tweets represent the most visible and influential discourse within the hashtag stream (Zeng and Schäfer 2021). This engagement-based sampling strategy enables the analysis to capture high-impact narratives that are most likely to shape public opinion in a post-truth digital environment (Cinelli et al. 2021). Two independent coders applied the coding schema, achieving inter-rater reliability of $\kappa = 0.81$, with disagreements resolved through structured discussion.

Both coders have academic backgrounds in Civic Education, Law, and Political Studies, with prior experience in qualitative discourse analysis. A calibration process was conducted to ensure consistent interpretation of coding categories, and each post-truth indicator was operationalized using clearly defined criteria and observable discourse markers (see Table 3).

Table 3. Post-Truth Indicator Operationalization

Indicator	Operationalization	Example Discourse Markers
Emotional Dominance	Proportion of tweets where emotional language supersedes factual claims	Exclamation clusters, grief emoji chains, capitals, repetitive 'Oh...'
Moral Simplification	Binary good/evil framing; absence of moral ambiguity	'Murderers in uniform' vs. 'martyrs'; hero/villain dichotomy
Institutional Delegitimization	Categorical rejection of state institution legitimacy	'Law belongs to the ruler'; 'complicit police'; 'failed state'
Virality & Trending Behavior	Engagement-driven amplification over content depth	Hashtag $\times 11$ in one tweet; 'copy-paste and share' imperatives
Narrative Polarization	Binary ingroup–outgroup identity construction	'We/us' vs. 'regime/them'; islamophobic or counter-islamist slurs
Unverified Claims	Strong factual assertions without source attribution	'They were tortured first'; stated as established fact without evidence

Validity and Research Ethics

Validity was strengthened through methodological triangulation (lexical, thematic, and close-reading analysis) and data triangulation (cross-referencing discourse patterns with contemporaneous news reports and Komnas HAM official records). All analyzed content was publicly posted and accessible without authentication, meeting standard ethical criteria for social media research (Altay, de Araujo, and Acerbi 2022). Usernames in illustrative examples are partially anonymized to minimize reidentification risk.

In addition, inter-coder reliability (Cohen's $\kappa = 0.81$) was employed to ensure consistency in thematic classification, indicating substantial agreement. The combination of methodological triangulation, systematic coding procedures, and reliability testing enhances the robustness and credibility of the analytical findings.

RESULTS AND DISCUSSION

Overall Discourse Landscape: Scale, Intensity, and Lexical Profile

The #IndonesiaHumanRightsSOS corpus represents an extraordinary episode of mass digital mobilization around an HR event. The 106,903 tweets were generated within 72 hours at an average posting rate of approximately 1,485 tweets per hour underscoring acute emotional intensity. Figure 1 presents the twenty most frequent substantive words. Notably, 'indonesia' (1,978), 'fpi' (1,371), 'justice' (1,360), 'world' (1,401), 'people' (1,286), 'law' (1,010), and 'rights' (966) define the semantic core. The co-occurrence of 'justice' and 'law' alongside 'police' and 'state' reveals that the discourse frames the KM50 incident as a crisis of institutional legitimacy and legal accountability rather than a discrete security operation.

Taken together, these patterns indicate that the discourse is not merely reactive but structurally oriented toward amplifying legitimacy claims and delegitimizing institutional authority. The convergence of high-frequency legal terminology with emotionally charged expressions suggests a hybrid discourse that frames the issue simultaneously as a moral crisis and a legal failure. At the same time, the prominence of terms such as 'world' (1,401) and 'hear' (579) signals an explicit internationalization strategy, whereby users seek to draw global attention to what is framed as a domestic HR emergency consistent with findings that high-salience political hashtags function as both mobilization tools and international advocacy signals (González-Bailón et al. 2023). This dynamic is further reinforced by a highly skewed contribution structure, with top users posting at extreme frequencies (e.g., 308 tweets in 72 hours), indicating super-spreader behavior that blurs the distinction between high visibility and broad public representation in post-truth communication environments.

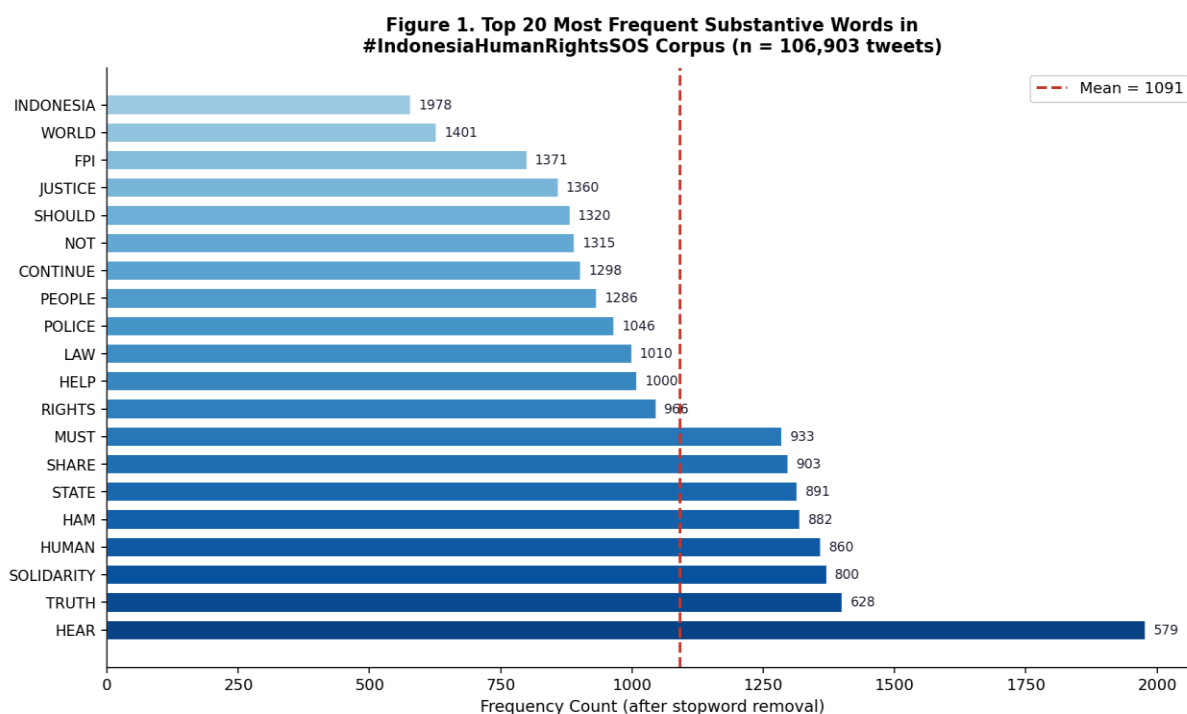


Figure 1. Top 20 Most Frequent Substantive Words (Content Words Representing Main Meaning) in #IndonesiaHumanRightsSOS (n = 106,903 tweets)

RQ1: Forms of Human Rights Discourse

Thematic coding yielded six discourse form categories (Figure 2). The dominant category Demands for Justice and Accountability (28.4%) encompasses tweets explicitly calling for criminal investigation, independent fact-finding, legal prosecution, and accountability to international HR standards. Solidarity and Victim Support (22.1%) includes expressions of grief, condolence, and collective identification with victims. Condemnation of State/Security Forces (19.3%) comprises direct delegitimizing statements targeting the National Police and broader state apparatus.

Collectively, these discourse forms reveal a layered communicative structure in which informational, emotional, and strategic elements operate simultaneously. Rather than functioning as isolated categories, these forms interact to reinforce message salience, indicating that digital HR discourse is shaped by both normative advocacy and platform-driven visibility dynamics.

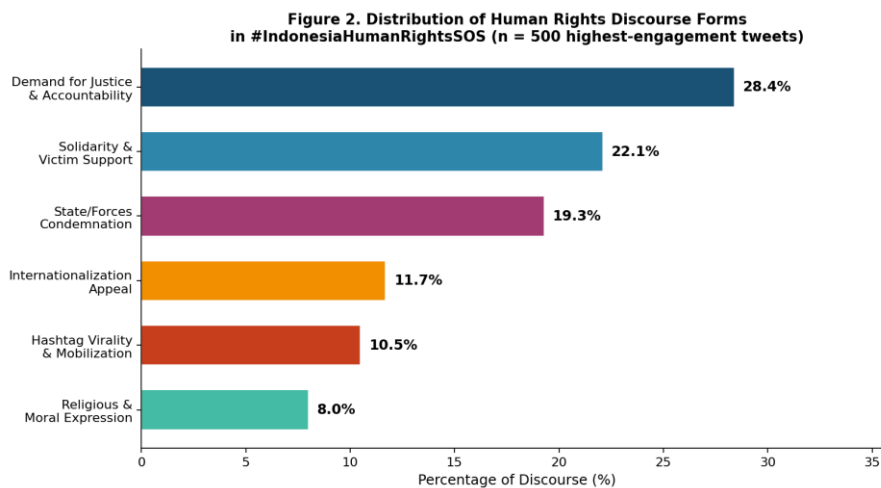


Figure 2. Distribution of Human Rights Discourse Forms in #IndonesiaHumanRightsSOS

The Internationalization Appeal category (11.7%) reflects a deliberate strategic choice: users directly tagged @UNHumanRights, @hrw, @IntlCrimCourt, @AJEnglish, and heads of state, constituting a form of 'networked advocacy' that leverages global connectivity to bypass domestic institutional gatekeepers. One representative tweet (paraphrased and anonymized) read: 'International attention to the extrajudicial killing of 6 martyrs by the Indonesian Police will have an impact on Indonesia's reputation in the international world.' The Hashtag Mobilization category (10.5%) reveals engagement-optimization behaviors some tweets repeating #IndonesiaHumanRightsSOS eleven times in a single post alongside 'copas dan share' (copy and share) imperatives. Pennycook et al. (2021) demonstrate experimentally that such metric-gaming behaviors decouple virality from content accuracy, lowering audiences' epistemic standards. To provide a more concrete illustration of how these discourse categories are articulated in practice, Table 4 presents representative tweets for each category along with their key discursive features.

Table 4. Representative Tweets by Discourse Form Category (Paraphrased & Anonymized)

Category (%)	Representative Tweet	Key Discursive Features
Justice & Accountability (28.4%)	"We will never stop speaking out. Silence is complicity in injustice. Uphold law equally, without discrimination or selectivity."	Religious legitimation; egalitarian legalism; imperative mood; moral urgency
Solidarity & Victim Support (22.1%)	"How many more victims must fall? 6 shot dead by police. When will this stop?"	Victim enumeration; rhetorical escalation; emotional appeal; collective identification
State Condemnation (19.3%)	"LAW BELONGS TO THE RULERS. Has the people's right to protection been stolen?"	Capitalization emphasis; rhetorical question; institutional rejection; power critique

Internationalization (11.7%)	"INTERNATIONAL PLEASE HELP!!! @UNHumanRights @hrw @AJEnglish @IntlCrimCourt"	Foreign address; multi-hashtag bundling; urgency markers; bypassing domestic channels
Hashtag Mobilization (10.5%)	"#IndonesiaHumanRightsSOS × 11 repetitions. Copas n Share"	Metric inflation; zero substantive content; viral trigger
Religious & Moral (8.0%)	"Ya Allah, reveal the slander destroying this nation. Slander is worse than murder. Allahu Akbar."	Divine address; religious moral authority; collective outrage framing

RQ2: Human Rights Enforcement Challenges

Figure 3 presents the distribution of HR enforcement challenge themes, with Impunity and Procedural Injustice as the most prevalent (31.2%), followed by Institutional Trust Crisis (25.7%) and Politicization of HR Cases (18.4%). Together, these three themes account for 75.3% of the discourse, indicating a mutually reinforcing triple deficit in accountability, trust, and political neutrality. This pattern suggests that public perceptions of HR enforcement are shaped not by isolated institutional failures but by an interconnected set of structural deficits, reflecting a broader legitimacy crisis in which legal processes are perceived as contingent upon political power rather than governed by the rule of law.

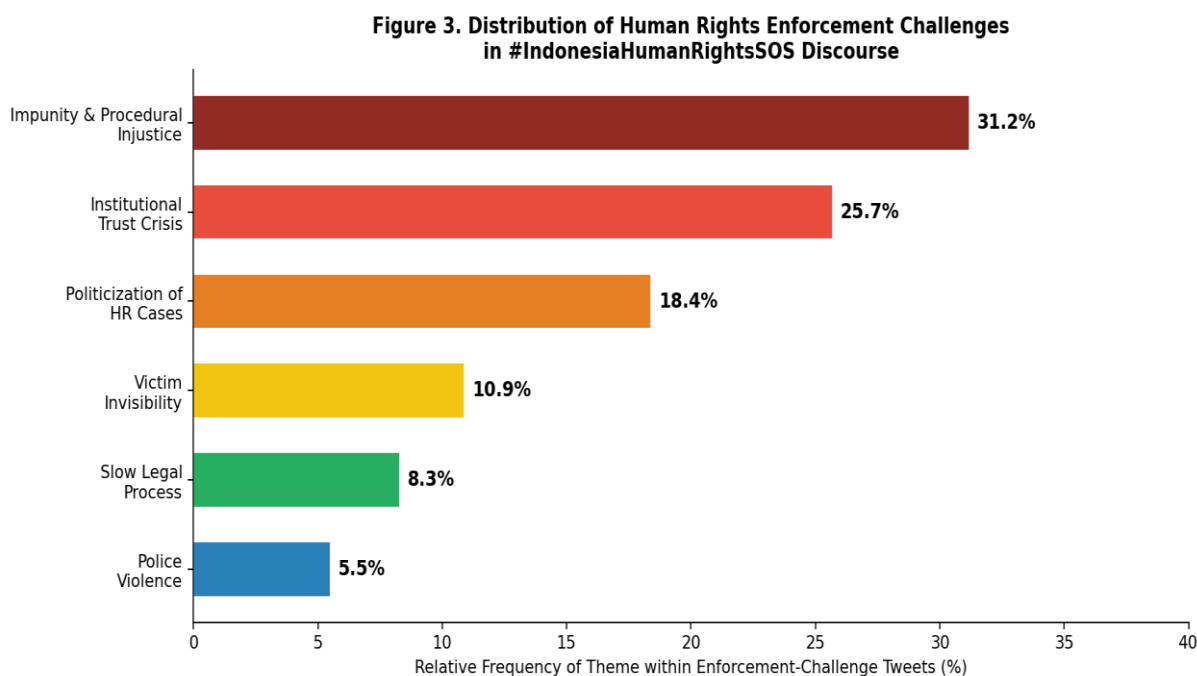


Figure 3. Distribution of HR Enforcement Challenge Themes in #IndonesiaHumanRightsSOS

The Impunity and Procedural Injustice theme (31.2%) manifested through a rhetorical strategy we term 'procedural irony' juxtaposing victims' pre-mortem legal status with the fact of their deaths to expose the gap between legal procedure and political reality. A widely retweeted example (paraphrased): 'Still classified as SUSPECT, already SHOT DEAD—Police: Status of 6 FPI Laskar Still Terlapor, Not Yet Formally Charged.' This framing exploits

procedural detail to mount a moral argument about impunity. The Institutional Trust Crisis (25.7%) was evidenced by affective delegitimization: phrases such as 'HUKUM MILIK PENGUASA' ('Law belongs to the rulers') indicate a collapse of perceived institutional legitimacy. This aligns with Lim's (2021) documentation of how platform algorithms intensify ideological closure, creating communicative environments in which official narratives are systematically discounted. Moreover, Ecker et al. (2022) identify institutional trust collapse as a primary driver of misinformation susceptibility, predisposing distrustful audiences to accept counter-narratives reinforcing prior beliefs.

Politicization of HR Cases (18.4%) emerged in tweets linking KM50 to broader partisan contests foregrounding political actors and framing HR demands within an opposition government binary. Such politicization exemplifies what Hadiz (2021) describes as Indonesia's 'dysfunctional democratic' tendency, wherein HR violations are incorporated into larger narratives of regime illegitimacy. Critically, this simultaneously amplifies HR visibility while fragmenting deliberative conditions necessary for credible fact-finding. Victim Invisibility (10.9%) appeared in tweets either extending the category of HR victims beyond FPI members to include unrelated policy failures, or conversely contesting 'martyr' framing reflecting the deep symbolic stakes of HR discourse in polarized contexts.

RQ3: Post-Truth Characteristics in HR Discourse

Figure 4 profiles six post-truth indicators, revealing a discourse profoundly shaped by post-truth dynamics, with three indicators exceeding 80/100. Virality and Trending Behavior scored highest (92/100), reflecting a structural orientation toward metric maximization rather than informational completeness. Consistent with Altay et al. (2023), highly shareable content is driven less by factual accuracy than by emotional and identity-based resonance, indicating the dominance of viral communication logic in HR discourse.

Emotional Dominance (87/100) was evident in intensive capitalization, repetitive expressions, emoji amplification, and morally charged language. In many cases, emotional intensity substituted for evidential argumentation, with highly engaged tweets relying primarily on affective appeals. This pattern aligns with the “affective cascade” mechanism and findings by Pennycook and Rand (2021) that emotional arousal reduces accuracy evaluation.

Institutional Delegitimization (81/100) manifested in categorical rejection of state legitimacy, extending beyond critique into systemic distrust. This reflects broader patterns identified by Lorenz-Spreen et al. (2023), in which declining institutional trust undermines the epistemic foundations of democratic accountability.

Figure 4. Post-Truth Indicator Profile of #IndonesiaHumanRightsSOS Discourse (Composite Score 0-100)

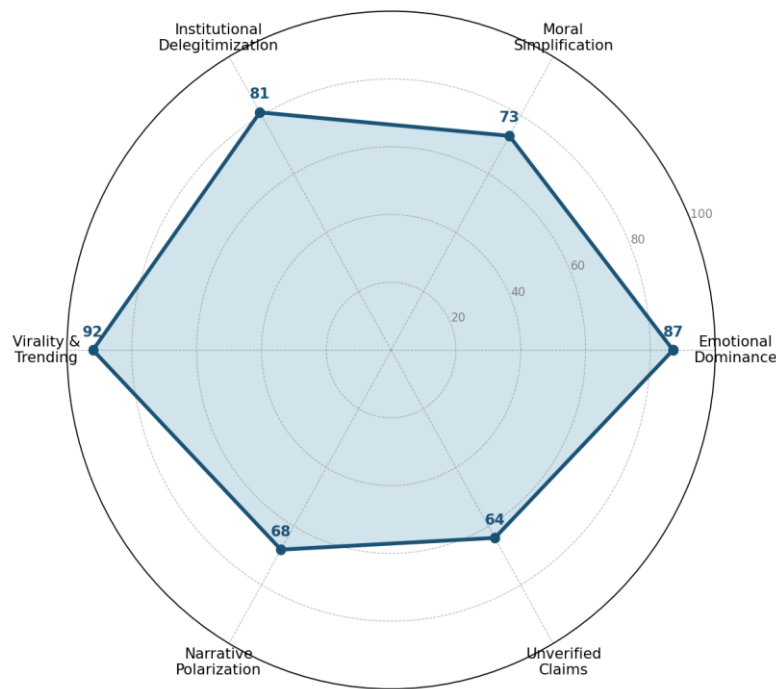


Figure 4. Post-Truth Indicator Profile of #IndonesiaHumanRightsSOS Discourse (Score 0–100)

Moral Simplification (73/100) appeared through binary moral framings that reduced complex events into clear victim perpetrator dichotomies, consistent with “simplicity bias” (Lewandowsky and van der Linden 2021). Narrative Polarization (68/100) was reflected in persistent ingroup–outgroup constructions, reinforcing ideological divisions and limiting cross-partisan engagement. Unverified Claims (64/100) further indicated the presence of assertive factual statements without supporting evidence, illustrating how uncertainty is often resolved through narrative rather than verification (Hameleers 2022).

Overall, these findings demonstrate that HR discourse is not only shaped by emotional and moral factors but is structurally conditioned by platform logics that prioritize engagement over epistemic quality. This configuration reinforces the dominance of simplified, polarized, and highly shareable narratives within the discourse ecosystem. To synthesize the key findings across analytical dimensions, Table 5 provides a concise summary of dominant patterns and their interpretations.

Table 5. Summary of Key Findings in #IndonesiaHumanRightsSOS Discourse

Analytical Dimension	Key Findings	Interpretation
Discourse Forms	Dominance of justice demands, solidarity, and state condemnation	Indicates strong normative orientation and emotional mobilization
Enforcement Challenges	Impunity, institutional distrust, politicization	Reflects systemic legitimacy crisis in HR enforcement

Post-Truth Dynamics	High virality, emotional dominance, delegitimization	Shows platform-driven amplification of affective and polarized narratives
---------------------	--	---

These findings collectively reveal a post-truth paradox in HR advocacy: the very communicative features generating unprecedented mobilization emotional intensity, moral simplification, viral mechanics simultaneously undermine the deliberative conditions necessary for credible fact-finding and durable accountability. This finding resonates with Altay et al.'s (2022) demand-side misinformation research: audiences actively seek content confirming prior moral commitments, creating structural incentives that reward affective over epistemic quality. Addressing this paradox requires platform-level design interventions that reward evidential quality over emotional virality, alongside prebunking-based civic education programs that inoculate citizens against post-truth manipulation (Lewandowsky and van der Linden 2021; Roozenbeek, Traber, and van der Linden 2022).

Implications for Civic Education and HR Governance

The pervasive post-truth characteristics identified in this study highlight the need to integrate critical digital literacy into civic education curricula. Students must develop competencies to distinguish between legitimate emotional advocacy and manipulative affective mobilization, evaluate source credibility, and recognize identity-based polarization as a barrier to HR accountability. Prebunking approaches proactively inoculating individuals against post-truth strategies offer a more effective alternative to reactive fact-checking and represent a key direction for civic and HR education.

For HR institutions such as Komnas HAM, the findings underscore the importance of rapid, transparent, and evidence-based communication during high-velocity events. The documented institutional trust crisis (25.7%) indicates a credibility gap that predisposes the public to reject official narratives. Strengthening institutional independence and procedural transparency is therefore essential for restoring discourse legitimacy.

These findings align with prior research on emotionally driven and identity-based digital communication (Pennycook and Rand 2021; Pennycook et al. 2021; Ecker et al. 2022), while extending the literature by demonstrating how post-truth dynamics reshape human rights advocacy in the Indonesian context. Specifically, the study shows that even legitimate HR concerns can be reframed within post-truth communicative structures, complicating efforts toward evidence-based public deliberation.

CONCLUSION

This study examined #IndonesiaHumanRightsSOS as a case of human rights discourse in the post-truth era, based on 106,903 tweets collected within a 72-hour period. The findings yield

three principal conclusions aligned with the research questions. *First*, in response to RQ1, the discourse is dominated by demands for justice and accountability (28.4%), solidarity expressions (22.1%), and condemnation of state actors (19.3%), indicating a communicative structure that is simultaneously normative, emotional, and strategically oriented toward international visibility. A notable proportion (10.5%) reflects engagement-optimization practices with minimal substantive content, suggesting the influence of platform-driven amplification logics. *Second*, in response to RQ2, human rights enforcement challenges are primarily characterized by impunity and procedural injustice (31.2%), institutional trust crisis (25.7%), and politicization (18.4%). These patterns indicate a structurally embedded legitimacy deficit in which accountability, trust, and political neutrality are mutually reinforcing constraints on effective human rights enforcement. *Third*, in response to RQ3, post-truth characteristics are strongly evident, particularly in virality (92/100), emotional dominance (87/100), and institutional delegitimization (81/100). The discourse reflects a post-truth epistemic configuration in which emotional resonance, moral simplification, and metric amplification often supersede evidential reasoning, giving rise to what this study conceptualizes as a post-truth paradox in human rights advocacy.

These findings carry important implications. Substantively, they demonstrate that human rights discourse in digital environments is not only shaped by normative concerns but also structurally conditioned by platform logics that privilege visibility and emotional engagement. Practically, this underscores the need to integrate prebunking-based critical digital literacy into civic education, while strengthening rapid, transparent, and evidence-based communication strategies among human rights institutions to rebuild public trust.

This study is limited by its focus on a single high-intensity event and a temporally bounded dataset, which may not capture longer-term discourse evolution. Future research should adopt longitudinal and comparative designs, as well as computational approaches such as sentiment analysis and network modeling, to further examine the dynamics and diffusion of digital human rights discourse across contexts.

REFERENCES

- Ahmed, S. K., R. A. Mohammed, A. J. Nashwan, R. H. Ibrahim, A. Q. Abdalla, B. M. M. Ameen, and R. M. Khdir. (2025). Using Thematic Analysis in Qualitative Research. *Journal of Medicine, Surgery, and Public Health* 6: 100198. <https://doi.org/https://doi.org/10.1016/j.glmedi.2025.100198>.
- Allen, J., A. A. Arechar, G. Pennycook, and D. G. Rand. (2021). Scaling up Fact-Checking Using the Wisdom of Crowds. *Science Advances* 7 (36): eabf4393. <https://doi.org/10.1126/sciadv.abf4393>.
- Altay, S., E. de Araujo, and A. Acerbi. (2022). Quantifying the Appetite for Misleading News: Demand-Side Evidence from a Representative Doorstep Survey. *Harvard Kennedy*

School Misinformation Review 3 (2). <https://doi.org/10.37016/mr-2020-104>.

- Cinelli, M., A. Pelicon, I. Mozetič, and et al. (2021). Dynamics of Online Hate and Misinformation. *Scientific Reports* 11 (1): 22083. <https://doi.org/10.1038/s41598-021-01487-w>.
- Ecker, U. K. H., S. Lewandowsky, J. Cook, P. Schmid, L. H. Fledderus, K. A. Hatch, and M. S. Zaragoza. (2022). The Psychological Drivers of Misinformation Belief and Its Resistance to Correction. *Nature Reviews Psychology* 1: 13–29. <https://doi.org/10.1038/s44159-021-00006-y>.
- Freelon, D. (2021). Computational Research in the Post-API Age. *Political Communication* 38 (1–2): 1–14. <https://doi.org/10.1080/10584609.2020.1722149>.
- González-Bailón, S., D. Lazer, P. Barberá, M. Zhang, B. Nyhan, J. Settle, and J. A. Tucker. (2023). Asymmetric Ideological Segregation in Exposure to Political News on Facebook. *Science* 381 (6656): 392–98. <https://doi.org/10.1126/science.ade7138>.
- Hadiz, V. R. (2021). The ‘end of Liberalism’ and Southeast Asian Dysfunctional Democracies. *The Pacific Review* 34 (6): 1057–76. <https://doi.org/10.1080/09512748.2021.1915965>.
- Hameleers, M. (2022). Separating the Wheat from the Chaff: A Conceptual, Experimental, and Meta-Analytic Review of the Research on Misinformation and Its Content-Related Correction. *Social Media + Society* 8 (2): 20563051221137604. <https://doi.org/10.1177/20563051221137603>.
- Krippendorff, K. (2019). *Content Analysis: An Introduction to Its Methodology*. Fourth Edition. Thousand Oaks, California: SAGE Publications, Inc. <https://doi.org/10.4135/9781071878781>.
- Lewandowsky, S., and S. van der Linden. (2021). Countering Misinformation and Fake News through Inoculation and Prebunking. *European Review of Social Psychology* 32 (2): 348–84. <https://doi.org/10.1080/10463283.2021.1876983>.
- Lim, M. (2021). Claiming the Public Sphere on Digital Platforms: Twitter and the Politics of Visibility in Indonesia. *Asian Journal of Communication* 31 (4): 259–75. <https://doi.org/10.1080/01292986.2021.1896553>.
- Lorenz-Spreen, P., L. Oswald, S. Lewandowsky, and R. Hertwig. (2023). A Systematic Review of Worldwide Causal and Correlational Evidence on Digital Media and Democracy. *Nature Human Behaviour* 7: 74–101. <https://doi.org/10.1038/s41562-022-01460-1>.
- Mietzner, M., and B. Muhtadi. (2022). The Populist Challenge to Liberal Democracy in Indonesia. *Democratization* 29 (4): 662–79. Doi: <https://doi.org/10.1080/13510347.2021.2011754>.
- Noeyislearning. (2021). Twitter Data on IndonesiaHumanRightsSOS [Dataset]. <https://www.kaggle.com/datasets/noeyislearning/twitter-data-on-indonesiahumanrightssos>.
- Nyhan, B., J. Settle, E. Thorson, M. Wojcieszak, P. Barberá, A. Y. Chen, and J. A. Tucker. (2023). Like-Minded Sources on Facebook Are Prevalent but Not Polarizing. *Nature* 620 (7972): 137–44. <https://doi.org/10.1038/s41586-023-06297-w>.
- Pennycook, G., Z. Epstein, M. Mosleh, A. A. Arechar, D. Eckles, and D. G. Rand. (2021). Shifting Attention to Accuracy Can Reduce Misinformation Online. *Nature* 592: 590–95. <https://doi.org/10.1038/s41586-021-03344-2>.
- Pennycook, G., and D. G. Rand. (2021). The Psychology of Fake News. *Trends in Cognitive Sciences* 25 (5): 388–402. <https://doi.org/10.1016/j.tics.2021.02.007>.

- Roozenbeek, J., C. S. Traberg, and S. van der Linden. (2022). Techniques of Science Denial and Their Relationship to COVID-19 Conspiracy Theories. *Royal Society Open Science* 9 (3): 211412. <https://doi.org/10.1098/rsos.211412>.
- Zeng, J., and M. S. Schäfer. (2021). Conceptualizing ‘Dark Platforms’. Covid-19-Related Conspiracy Theories on 8kun and Gab. *Digital Journalism* 9 (9): 1321–43. <https://doi.org/10.1080/21670811.2021.1938165>.