



The Cyber Footprint of Digital Activism: The Legal Boundary between Freedom of Expression and Hate Speech in Indonesia

Daimah

IAI Bunga Bangsa Cirebon, Indonesia

Corresponding email: marwadaimah@gmail.com

Abstract *The development of digital technology has encouraged the emergence of digital activism as a new means for citizens to voice opinions and fight for socio political rights. However, in Indonesia, this space for digital expression is often faced with repressive regulations, particularly through the use of rubber articles in the Electronic Information and Transaction Law (ITE Law). This research aims to analyze the legal boundary between freedom of expression and hate speech critically in the context of digital activism, identify potential abuses of digital law against activists and citizens, and offer a just and democratic juridical framework. Using a descriptive qualitative approach with a case study method, data were collected through documentation of legal regulations, interviews with key figures, and ethnographic observations on social media. The results show that ITE Law is often used to silence political expression, with the trend of reporting against activists increasing in the last four years. The unclear definition of hate speech and the absence of explicit protection for political expression worsen the conditions for digital democracy. This research recommends digital regulatory reforms, including establishing an independent oversight body and affirming the legal boundary between criticism and hate speech, to guarantee civil rights in the cyber era.*

Keywords: *freedom of expression, hate speech, digital activism, ITE law, cyber regulation, digital democracy bottom*

1. Introduction

Digital transformation has expanded the space for citizen participation in democracy through cyber activism, becoming a tool for resistance to social and political inequality in Indonesia (Paskarina, 2020; Juned et al., 2024; Fernando et al., 2022). However, this freedom is directly confronted with the challenges of hate speech, the spread of disinformation, and digital repression that arise along with the penetration of information technology in the public sphere (Sumartias et al., 2025; Masduki, 2022; Lim, 2017). In this context, there is a tension between the protection of the right to freedom of expression and the need to maintain social order in cyberspace.

In this study, the term *cyber footprint* refers to the digital traces left behind by individuals or groups when engaging in online activism. These traces include, but are not limited to, tweets, digital petitions, viral political memes, hashtag campaigns, user generated videos, and other forms of participatory content shared across digital platforms such as Twitter/X, TikTok, Instagram, and Facebook. These digital artifacts

contribute to a broader digital narrative and become part of the public record of political dissent and engagement in cyberspace.

The legal boundary between freedom of speech and hate speech has become increasingly important amidst social and political polarization, especially ahead of elections and other major political events in Indonesia (Paterson, 2019; Venkiteswaran, 2017; Ghofur, 2024). In practice, articles in the ITE Law are often used to silence citizen criticism of the government or political elites through hate speech reporting (Aditya & Al Fatih, 2021; Julianti & Sugiantari, 2022; Yilmaz et al., 2022). This creates concerns about the misuse of the law as a tool of digital repression.

In the Indonesian context, digital activism has taken diverse and dynamic forms. These include using protest hashtags such as #ReformasiDikorupsi, #PercumaLaporPolisi, and #SaveSambo, which became viral expressions of public discontent. Activists also engage in public criticism of government officials, often through Twitter threads, Instagram infographics, or TikTok videos that highlight corruption, police brutality, or legislative injustice. In more advanced forms, digital activism has included digital

, such as the high profile case of *Bjorka*, who leaked government data to expose cybersecurity vulnerabilities and political hypocrisy. These acts serve as instruments of resistance, enabling marginalized voices to bypass mainstream media gatekeeping and directly challenge power structures. However, these same actions are increasingly framed by authorities as violations of the ITE Law, blurring the line between activism and criminal conduct.

A central issue lies in the vague and overly broad formulation of several key articles in the ITE Law, notably Article 27(3) and Article 28(2). Article 27(3) criminalizes acts deemed to "defame or insult" others through electronic media. In contrast, Article 28(2) prohibits the dissemination of information intended to incite hatred or hostility based on ethnicity, religion, race, or group (SARA). The legal problem stems from the lack of precise definitions for terms like *insult*, *hatred*, and *hostility*. Law enforcement and judiciary bodies tend to adopt subjective interpretations, often influenced by political or ideological leanings. For instance, criticism of government policies or public figures, such as a satirical meme or viral tweet, can be prosecuted under Article 28(2) if deemed by authorities to contain elements of incitement, even when no clear intention or effect exists. This interpretive flexibility has created a climate of legal uncertainty, where digital expressions are increasingly vulnerable to criminalization (Al Fatih, 2021; Ghofur, 2024).

Based on data from SAFEnet and Kominfo, the number of hate speech cases handled by law enforcement officials has significantly increased over the past five years. As seen in the following graph, cases increased from 125 in 2018 to 630 in 2023 (SAFEnet, 2024; Kominfo, 2023; Putra, 2022). This increase illustrates the urgency of reviewing the current legal framework.

Research by Lim (2017) highlights how social media in Indonesia creates "algorithmic enclaves" that fuel tribal nationalism and facilitate hate speech. Meanwhile, studies by Fernando et al. (2022) and Hasfi et al. (2024) show that strict and sometimes ambiguous digital regulations heavily influence online freedom of expression in Indonesia. On the other hand, an international study by Udupa (2021) emphasizes the importance of a differential approach to extreme speech, rather than simply wrapping all controversial expressions in the term "hate speech".

While many studies have examined freedom of expression and hate speech in general, not many have specifically examined how Indonesian law navigates the boundary between digital activism and hate speech in the contemporary socio political context. Moreover, no comprehensive analysis links these dynamics with their impact on Indonesia's digital democratic climate (Paskarina, 2020; Sumartias et al., 2025; Margiansyah, 2025).

This research offers novelty through an interdisciplinary approach that combines aspects of law, digital communication, and politics. This article will also present a mapping of digital narratives through case studies such as the case of digital activist Bjorka and the arrest of several netizens after criticizing the government (Juned et al., 2024; Masduki, 2022; Ghofur, 2024). Thus, this research is normative, empirical, and contextual.

The main objectives of this research are to critically analyze the legal boundary between freedom of expression and hate speech in the context of digital activism in Indonesia, identify potential abuses of digital law against activists and ordinary citizens, and offer a fair juridical framework for regulating online expression without compromising civil rights.

2. Method

Type of Research

This research uses a descriptive qualitative approach with a case study method to deeply understand the dynamics of the legal boundary between freedom of expression and hate speech in Indonesia's digital activism context. This approach was chosen because it allows for the exploration of complex social, legal, and political contexts in a narrative and interpretative manner (Creswell, 2014; Yin, 2018; Neuman, 2011).

Population and Sampling

The population in this study includes all legal documents, court decisions, and digital content (social media, online news, public forums) relevant to cases of hate speech and freedom of expression from 2018 to 2024. The sample in this study was selected using purposive sampling, which included five nationally impactful digital activism cases (e.g., Jerinx SID, Baiq Nuril, and Bjorka), ten relevant legal products

such as the ITE Law, PDP Law, and MOCI Regulations, and semi-structured interviews with at least six informants comprising legal academics, activists, digital lawyers, and journalists.

Research Instrument

The main instrument of this research is the researcher himself as a qualitative instrument supported by:

- 1) Semi structured interview guide.
- 2) Checklist for analyzing legal documents and digital content.
- 3) Coding template for NVivo based thematic narrative analysis.

Data Collection Technique

The data collection techniques used consisted of: (1) In depth interviews with key figures. Documentation: includes laws and regulations, court decisions, NGO reports (such as SAFEnet and Amnesty International), and online news. (2) Netnographic observation on digital platforms (Twitter/X, Instagram, TikTok) to directly observe public discourse.

Research Procedure

- 1) Literature study and digital case identification (May-June 2024).
- 2) Collection of legal documents and online media (July-August 2024).
- 3) Conducting interviews and transcription (September 2024).
- 4) Coding and thematic analysis using NVivo (October 2024).
- 5) Compilation of findings and validation through triangulation (November 2024).
- 6) Finalize report writing (December 2024).

Data Analysis Technique

Data were analyzed using thematic analysis with the procedure:

- (1) Open coding of interview data and documents.
- (2) Categorization is based on broad themes such as "freedom of expression", "hate speech", "abuse of law", and "digital activism".
- (3) In depth interpretation to uncover contextual meanings and patterns of relationships between themes (Braun & Clarke, 2006; Miles & Huberman, 2014).
- (4) Data validation was done through source and method triangulation to maintain the credibility of the research results.

To complement the legal and digital discourse analysis, this study also incorporates political analysis by contextualizing digital activism within Indonesia's broader policy environment and power dynamics. This is done through two strategies: (1) Policy mapping, which examines the evolution and content of key state regulations such as the ITE Law, PDP Law, and ministerial decrees on cyber governance, and their political motivations or implications; and (2) Discourse contextualization, which analyzes how digital expressions intersect with political events such as elections, public protests, or elite controversies. Political positioning of

actors (e.g., state officials, influencers, activist groups) is identified through media narratives and interview data patterns. This dimension helps uncover how public policy, governance culture, and political interests shape the interpretation and enforcement of digital regulations, thereby enriching the legal analysis with institutional and power centered insights.

3. Result & Discussion

Voices from the Field: Perspectives from Key Informants

To deepen the contextual understanding of digital repression, this study draws on semi structured interviews with six key informants: two legal scholars, one digital rights activist, one journalist, and two lawyers specializing in cyber law. These interviews reveal a consistent concern that the vague phrasing of the ITE Law allows for political manipulation of legal instruments. One legal scholar noted: "The term 'hate speech' has become a political tool, and what counts as hate often depends on who is offended and who is in power." A human rights activist described how online reporting functions as a "weaponized" form of silencing: "Even a meme or sarcastic comment can trigger legal threats if it criticizes the ruling elite."

Interviewees also pointed out that the lack of institutional independence in digital law enforcement, particularly within Kominfo, negatively affects online expression. A cyber law attorney stated, "There is no buffer between political pressure and legal action. We have defended multiple clients who reposted criticism and still got summoned." These insights demonstrate how the interplay between legal interpretation, digital surveillance, and political interests creates a fragile ecosystem for free expression in Indonesia.

Digital Activism and the Polarization of Expression in Cyberspace

Digital activism in Indonesia has increased significantly since 2019, and it has been marked by many social campaigns such as #ReformasiDikorupsi, #PercumaLaporPolisi, and #SaveSambo. These campaigns utilize social media as a freer and faster space for resistance (Paskarina, 2020; Lim, 2017; Fernando et al., 2022). However, this space for expression is not entirely free. Activists with critical voices are often labeled as spreading "hate speech" and reported using articles in the ITE Law (Ghofur, 2024; Paterson, 2019; Al Fatih, 2021).

While the percentage of ITE Law reports targeting activists ranges between 3.7% and 4.2%, this seemingly modest proportion obscures its disproportionate political impact. A qualitative review of 26 documented cases by SAFEnet between 2020–2023 shows that most targeted individuals were involved in criticizing state institutions, supporting protest movements, or highlighting corruption and human rights

violations (SAFEnet, 2024). For example, university student activists in Makassar and Medan were summoned for online posters criticizing public infrastructure failures during local elections. A teacher in West Java was charged under Article 27(3) after tweeting criticism about the allocation of education budgets. These cases often occur around politically sensitive periods, such as elections, policy reforms, or national controversies (Paterson, 2019; Fernando et al., 2022). This pattern suggests that the ITE Law has become a flexible instrument of anticipatory governance used to punish dissent and signal the boundaries of “acceptable” discourse in a managed democracy.

Such as Jerinx SID, who was charged with Article 28 paragraph 2 of the ITE Law for criticizing IDI, or Baiq Nuril, who was convicted for disseminating a recording of harassment (Putra, 2022; Masduki, 2022; Venkiteswaran, 2017). This reflects the fundamental tension between the norm of freedom of expression and the protection of honor or public order that is at stake in the digital space.

This phenomenon indicates that digital activism is often misunderstood as radicalism or provocation. In a democracy, critical voices are legitimate expressions in a healthy political life (Hasfi et al., 2024; Yilmaz et al., 2022; Sumartias et al., 2025). However, because regulations are still flexible and interpretative, law enforcers tend to repressively apply "rubber" articles against public expression.

Digital Law Abuse against Activists

SAFEnet data and Kominform reports show an increasing trend of reporting against activists using the ITE Law from 8 cases in 2020 to 19 cases in 2023, with a total percentage ranging from 3.7% to 4.2% of all ITE Law reports (SAFEnet, 2024; Kominform, 2023; Fernando et al., 2022). Although quantitatively small, this value is significant because it targets vulnerable groups such as journalists, students, and human rights defenders.

The figure above shows the trend of an increasing proportion of activists being targeted under the ITE Law, contributing to a decline in the civil liberties index, according to data from Freedom House (2023) and Amnesty International (2023). Paterson (2019) refers to this practice as "digital legalism," using legal tools to silence opposition and control public narratives.

Violations against digital activists are also often accompanied by *doxxing*, hacking, and disinformation that target their credibility (Juned et al., 2024; Margiansyah, 2025; Masduki, 2022). This phenomenon shows that regulations are not just used for order, but also to maintain the status quo of power.

The Legal Boundary Problem between Hate Speech and Political Expression

The main problem is the fine line between *hate speech* and criticism or political expression. The ITE Law, especially Articles 27 and 28, does not provide detailed definitions and leaves too much room for the authorities to interpret them (Al Fatih,

2021; Aditya, 2020; Ghofur, 2024). As a result, expressions on SARA, gender, and political issues are prone to criminalization.

According to Udupa (2021), hate speech should be distinguished from "extreme expressions" that are not necessarily unlawful but can be legitimate political discourse. This approach has not been adopted in Indonesia, so all expressions with controversial sentiments risk being considered illegal. This reinforces Lim's (2017) opinion that Indonesia is caught between digital openness and regulatory repression.

In interviews with SAFEnet activists, it was mentioned that politically affiliated buzzword groups carried out many reports. This emphasizes the importance of revising regulations so that they do not contain elements of "covert silencing" of public expression (Fernando et al., 2022; Yilmaz et al., 2022; Sumartias, 2025).

Alternative Juridical Framework and Digital Law Reform

From the document analysis and interviews, it was found that some countries, such as Germany and Canada, adopt a multilayered approach to regulating hate speech: separating hard hate from soft hate and setting thresholds of evidence and intention (UNESCO, 2023; Udupa, 2021; Paterson, 2019).

This article offers an alternative juridical framework for the Indonesian context: (1) Drafting a special law on digital freedom of expression. (2) The definitional phrases "political criticism" and "civil expression" were added to the revised ITE Law. (3) Increasing the digital literacy capacity of officials and communities to distinguish between criticism and hate speech (Paskarina, 2020; Hasfi, 2024; Ghofur, 2024).

Table.1 The following table shows a comparison of the criminal elements in Indonesia's ITE Law and similar regulations in other democracies:

Country	Definition of Hate Speech	Protection of Political Criticism	Appeal Mechanism
Indonesia	Not explicit, multiple interpretations	Not explicitly protected	Long legal process
Germany	Very detailed and limited	Protected under the Press Law	Independent commission
Canada	There must be evidence of malicious intent	Fully protected	Press Ombudsman

Source: UNESCO (2023), Al Fatih (2021), Ghofur (2024)

The table above shows that Indonesia still faces serious challenges in formulating a firm boundary between hate speech and political expression. The absence of an explicit definition in the ITE Law has led to the criminalization of many public expressions that are a form of democratic participation (Ghofur, 2024; Aditya & Al Fatih, 2021; Sumartias et al., 2025). This contrasts with countries such as Germany and

Canada, which have formulated detailed categories of hate speech based on its context, purpose, and impact on society.

Political criticism does not yet have explicit protection in digital legal regulations, particularly in Indonesia. This causes public discourse that should be allowed to grow in a democratic space to be repressed by ambiguous legal tools (Paterson, 2019; Venkiteswaran, 2017; Lim, 2017). Protection of such expression is important so that not all criticism of state institutions is immediately categorized as defamation or hate speech.

Another difference is seen in the appeal or correction mechanism for digital prosecution. In Indonesia, such mechanisms are still centered on law enforcement officials, and there is no digital independent institution like Germany (Digital Media Ethics Commission) or Canada (national press ombudsman) (UNESCO, 2023; Hasfi, 2024; Margiansyah, 2025). This has resulted in low public trust in procedural justice in digital legal cases.

Therefore, one of the urgent recommendations from this research is to encourage the establishment of an independent oversight institution that specifically handles cases of digital disputes between citizens and the state. This institution can complement conventional courts and strengthen cyber law enforcement's checks and balances system (Paskarina, 2020; Masduki, 2022; Sumartias et al., 2025).

Finally, adjustments to Indonesia's digital legal framework must consider emerging socio political dynamics, while guaranteeing the civil rights of citizens so that freedom of expression is protected, but not allowed to develop into speech that damages everyday life. Law reform is about the content of articles and the democratic spirit behind their implementation (Ghofur, 2024; Fernando et al., 2022; Udupa, 2021).

4. Conclusion

Starting from the aim of this study to analyze the legal boundaries between freedom of expression and hate speech in the context of digital activism in Indonesia, the results show that regulations such as the ITE Law have not clearly distinguished between legitimate political criticism and harmful hate speech. This lack of clarity opens room for abuse of the law, especially against digital activists, journalists, and citizens who express their opinions online. Findings also show that digital law enforcement tends to be repressive, with data showing increased reporting against activists using rubber articles, despite their quantitatively small contribution to total reports.

The research also found structural gaps in the protection of public expression that require substantive and institutional legal reform. Indonesia needs a digital legal framework that protects against hate speech and guarantees civil liberties as part of democracy. Therefore, building a more participatory legal system is important, including redefining hate speech, recognizing criticism as part of freedom of

expression, and establishing independent institutions to resolve digital disputes fairly. These findings support the need for digital law reform based on social justice, deliberative democracy, and human rights protection in Indonesia's cyber ecosystem.

5. References

- Aditya, Z. F., & Al Fatih, S. (2021). Indonesian constitutional rights: Expressing and purposing opinions on the internet. *The International Journal of Human Rights*, 25(10), 1692 1708. <https://doi.org/10.1080/13642987.2020.1826450>
- Fernando, Z. J., Pujiyono, R. U., & Rochaeti, N. (2022). The freedom of expression in Indonesia. *Cogent Social Sciences*. <https://www.researchgate.net/publication/371987265>
- Ghofur, N. (2024). Law, media, and democracy in the digital era: Freedom of expression and ITE regulation in Indonesia. *Al Mazaahib: Journal of Comparative Law*, 12(1), 102 120. <https://ejournal.uin-suka.ac.id/syariah/almazahib/article/view/3703>
- Hamayotsu, K. (2013). The limits of civil society in democratic Indonesia: Media freedom and religious intolerance. *Journal of Contemporary Asia*, 43(4), 658 679. <https://doi.org/10.1080/00472336.2013.780471>
- Julianti, L., & Sugiantari, A. A. P. W. (2022). The State of Digital Freedom in Indonesia: An Assessment of Online Censorship, Privacy, and Free Expression. *Journal of Digital Law and Policy*, 2(2). <https://ejournal.sidyanusa.org/index.php/jdlp/article/view/301>
- Juned, M., Martin, A., & Pratama, N. (2024). Bjorka's hacktivism in Indonesia: The intercourse paradox of cyberdemocracy, cyberactivism, and cybersecurity. *Academic Journal of Digital Politics*. <https://www.researchgate.net/publication/385707043>
- Lim, M. (2017). Freedom to hate: Social media, algorithmic enclaves, and the rise of tribal nationalism in Indonesia. *Critical Asian Studies*, 49(3), 411 427. <https://doi.org/10.1080/14672715.2017.1341188>
- Lim, M. (2018). Disciplining dissent: Freedom, control, and digital activism in Southeast Asia. In *Routledge Handbook of Urbanization in Southeast Asia* (pp. 529 545). https://www.taylorfrancis.com/chapters/edit/10.4324/9781315562889_44
- Margiansyah, D. (2025). Digitalizing authoritarianism in Indonesia: Exploring the intersection of civic activism, digital repression, and democratic erosion. In *Indonesian Perspectives on Democracy* (pp. 153 176). Springer. https://link.springer.com/chapter/10.1007/978_981_96_3137_7_11
- Masduki. (2022). Cyber troops, digital attacks, and media freedom in Indonesia. *Asian Journal of Communication*, 32(4), 297 312. <https://doi.org/10.1080/01292986.2022.2062609>
- Paskarina, C. (2020). Digital activism and democracy in Indonesia. *Indonesian Journal of Political Research*, 2(1), 1 12. <https://www.academia.edu/download/65369336/Paskarina2020.pdf>
- Paterson, T. (2019). Indonesian cyberspace expansion: A double edged sword. *Journal of Cyber Policy*, 4(2), 220 237. <https://doi.org/10.1080/23738871.2019.1627476>

- Putra, E. N. (2022). Criminal defamation and freedom of speech in the internet age: A study for Indonesian democratic values. CORE. <https://core.ac.uk/download/pdf/541436183.pdf>
- Udupa, S. (2021). Extreme speech and digital constitutionalism. *International Journal of Communication*, 15, 2287–2305. <https://ijoc.org/index.php/ijoc/article/view/16770>
- Yilmaz, I., Saleem, R. M. A., Pargoo, M., & Shukri, S. (2022). Religious populism, cyberspace and digital authoritarianism in Asia: India, Indonesia, Malaysia, Pakistan, and Turkey. Deakin University Report. <https://dro.deakin.edu.au/articles/report/20621349>