Volume 2 Issue 1 (2025) 1-9
E-ISSN 2988-6740| © Authors. CC BY-SA
https://doi.org/10.59261/iclr.v2i1.11
https://iclr.polteksci.ac.id/

Indonesian Cyber Law Review

Juridical Analysis of the Responsibility of Social Media Platforms for Cyber Threats Against Digital Activists in Indonesia

Siti Hapsah Pahira[1], Daimah[2], Irkham Mu'amar[3], Akpoghome Theresa[4]
[1] Universitas Islam Sultan Agung, Indonesia
[2] Politeknik Siber Cerdika Internasional, Indonesia
[3] Universitas Swadaya Gunung Jati, Indonesia
[4] Enson Idahosa University
Corresponding email: sitihafsahfahira23@gmail.com

ABSTRACT

The increasing intensity of cyberattacks against digital activists in Indonesia highlights the urgency of re-reading the responsibilities of social media platforms as providers of digital space. This study applies a normative-empirical approach, integrating doctrinal legal analysis with field-based insights to ensure contextual relevance in Indonesia's regulatory discourse. Attacks in the form of doxing, hacking, and intimidation not only threaten freedom of expression but also create inequality in legal protection for vulnerable groups. This research aims to juridically analyze the legal responsibility of digital platforms for cyber threats experienced by activists in Indonesia and formulate a more adaptive and human rights-based regulatory framework. While much of the literature focuses on state surveillance or user liability, few have examined the intermediary accountability of digital corporations in safeguarding human rights. The method used is a normative-empirical approach, with data collection techniques through documentation studies, in-depth interviews with activists and legal experts, and analysis of laws and regulations and internal platform policies. The results showed that most platforms failed to carry out the principle of due diligence and only followed up on a small portion of the reports submitted. On the other hand, the absence of national legal norms that explicitly regulate the responsibility of platforms also weakens legal protection for victims. This research recommends the establishment of new norms based on shared responsibility and integrating digital human rights principles in national laws and regulations. The findings emphasize the importance of a regulatory approach that is not only reactive but also preventive and accountable in dealing with threats to civil liberties in the digital space.

Keywords: *platform liability, cyberattacks, digital activists, legal protection, digital rights, social media.*

1. Introduction

In recent years, the intensity of cyberattacks against digital activists in Indonesia has shown an alarming upward trend. Activists who voice environmental, human rights, and public policy issues are often the targets of digital-based doxing, hacking, and intimidation [Anderson, 2021; SAFEnet, 2023; Komnas HAM, 2022]. This phenomenon places the digital space as a contestation between freedom of expression

and systematic threats to civil liberties. Amidst the limitations of comprehensive national regulations, the roles and responsibilities of social media platforms are under the spotlight in ensuring the safety of users, especially vulnerable groups such as digital activists [DeNardis, 2020; Solove & Citron, 2022; Waisbord, 2018].

The urgency of this research lies in the void of explicit legal norms regarding the limits of platform responsibility for the occurrence of digital crimes against activists. Indonesia has indeed passed Law No. 27 of 2022 on Personal Data Protection. It has an ITE Law, but neither has specifically regulated the principle of duty of care that must be carried out by digital platform providers [PDP Law, 2022; ITE Law, 2008; Putri & Mahendra, 2023]. On the other hand, community standards policies implemented by global platforms such as Meta and X (Twitter) are considered insufficiently adaptive to local contexts and often ignore protection requests from users in developing countries [Douek, 2022; Gillespie, 2018; Kaye, 2019].

Data from SAFEnet shows that from 2019 to 2024, the number of digital attacks against activists significantly increased, from 12 cases in 2019 to 72 cases in 2024 (See graph above). These attacks include account hacking, dissemination of personal data, and physical threats made through digital channels [SAFEnet, 2024; Amnesty International, 2023; AJI, 2022]. This surge exposes gaps in the legal protection system and slow response mechanisms on platforms.

Previous research has extensively addressed the protection of freedom of expression and digital rights in both global and national contexts. For example, a study by Deibert (2021) explores how states use digital infrastructure to silence opposition, while Lim (2020) highlights the phenomenon of cyber vigilantism in Southeast Asia. In Indonesia, Marzuki (2022) examines the criminal law aspects of doxing, but has not comprehensively discussed the role of digital corporations [Deibert, 2021; Lim, 2020; Marzuki, 2022].

The research gap in this study lies in the lack of research that examines explicitly the legal responsibility of social media platforms when digital attacks on activists occur. Most of the literature still focuses on user regulation or state apparatus, not corporate liability as an intermediary controlling digital infrastructure [Schmitz, 2020; Keller, 2021; Anggoro, 2022]. A digital corporate responsibility-based approach is very relevant in the context of algorithmic justice and digital human rights.

The novelty of this research is an in-depth juridical analysis of the civil and administrative liability of social media platforms based on the principle of shared responsibility and a human rights-based approach. This article will also propose a new legal norm model that is rooted in the local context but refers to international best practices such as the Due Diligence Principle in the UN Guiding Principles on Business and Human Rights [UNGP, 2011; Mantelero, 2019; Wahyuni, 2023]. Unlike prior works that generalize digital rights under state obligations, this article isolates the operational responsibility of social media platforms. It evaluates their obligations under international human rights standards contextualized within Indonesian law.

Thus, the main objective of this research is to formulate a legal framework that explains the legal responsibility of social media platforms in the context of digital attacks on activists in Indonesia. This research is expected to serve as an initial

foundation in the formulation of public policies and derivative regulations that are more responsive to the challenges of digital security and civil liberties in the internet age [Smith & Miller, 2022; Nugroho, 2024; Harsono, 2023].

2. Method

Type of Research

This research type is normative-empirical, combining a normative juridical analysis approach to legislation and jurisprudence with a qualitative empirical approach to examine the responses of users (digital activists) and social media platforms to cyber threats. This approach is used to answer the question of legal gaps and to examine the extent to which the responsibility of digital platforms can be accounted for under Indonesian law. This dual approach ensures a holistic analysis that not only considers the legal vacuum but also reflects the lived experiences of those affected by regulatory shortcomings

Population and Sampling

The population in this study consisted of two main groups:

1) Digital activists in Indonesia who have experienced threats or cyberattacks
2) Representatives from social media platforms operating in Indonesia (such as Meta, X, and TikTok)

The sample was determined through a purposive sampling technique for activists who have been victims of cyberattacks and snowball sampling for informants from civil society organizations, cyber law experts, and legal practitioners who handle related cases. The target number of informants was 10-15 people, consisting of 7 digital activists, three legal experts, and three representatives of civil society organizations.

Research Instrument

The main instruments used are:

1) A semi-structured interview guide organized around the principles of corporate responsibility and digital rights protection.
2) Checklist of legal documents, including the ITE Law, PDP Law, Community Standards Guidelines from digital platforms, and reports from organizations such as SAFEnet, Komnas HAM, and Amnesty International.
3) Legal document analysis template to assess the strengths and weaknesses of existing norms.

Data Collection Technique

Data was collected through four main techniques: in-depth interviews with digital activists and experts; a documentation study of national regulations, platform

policies and incident reports; online non-participatory observation of digital footprints and platform responses to reported content or attacks; and a literature review to explore the theoretical foundations of corporate responsibility and digital human rights principles.

Research Procedure

The research was conducted in six stages:

1) Identification of legal issues and problem formulation.
2) Collection and classification of normative data (laws, regulations, official platform documents).
3) Development of interview and case study instruments.
4) Field data were collected through interviews and documentation.
5) Legal and thematic analysis of the data obtained.
6) Drawing conclusions and recommendations based on the findings of the analysis.

Data Analysis Technique

Normative data was analyzed using a legal content analysis approach to the applicable legal norms and structures. Meanwhile, empirical data from interviews is analyzed using thematic analysis by grouping key issues such as types of threats, platform responses, and legal protection constraints. The results of both approaches will be synthesized to formulate a comprehensive construction of legal liability for social media platforms in the context of digital attacks on activists. The synthesis of normative and empirical analysis also allows for cross-validation between legal expectations and on-the-ground realities, strengthening the credibility of the proposed regulatory recommendations.

3. Result & Discussion

Patterns of Cyber Threats to Digital Activists in Indonesia

In-depth interviews with digital activists show that the most common cyber threats are doxing, hacking social media accounts, and spreading hoaxes that attack personal reputation. As many as 6 out of 7 activists interviewed admitted to experiencing more than one form of attack in the past year [SAFEnet, 2023; Komnas HAM, 2022; AJI, 2023]. Most of these attacks were launched through popular platforms like Facebook, Twitter (X), and Instagram.

Documentation studies have also found that doxing is often a gateway to more complex forms of attack, such as intimidation and physical threats in the real world. SAFEnet's 2024 data recorded an increase in cases from 65 (2023) to 72 (2024), showing

a trend of increasing activist vulnerability to digital threats [SAFEnet, 2024; Amnesty International, 2023; Walhi, 2024]. This indicates that the digital space is not yet safe for groups with critical voices. It also reinforces the notion that platforms, in their current form, are ill-equipped to manage context-sensitive content moderation in countries with complex political dynamics like Indonesia.

One typical pattern is using fake accounts to spread activists' personal information, including home addresses, phone numbers, and family identities. This creates fear and suppresses digital activism [Lim, 2020; Solove & Citron, 2022; Wahyuni, 2023]. In many cases, these attacks are never effectively followed up by platforms.

This threat not only has a psychological impact, but also has the potential to cripple social movements digitally. One environmental activist stated that her colleagues withdrew from social media for fear of becoming the next target [Deibert, 2021; McKay, 2022; Human Rights Watch, 2023]. This demonstrates the structural impact of inadequate legal and technological protections.

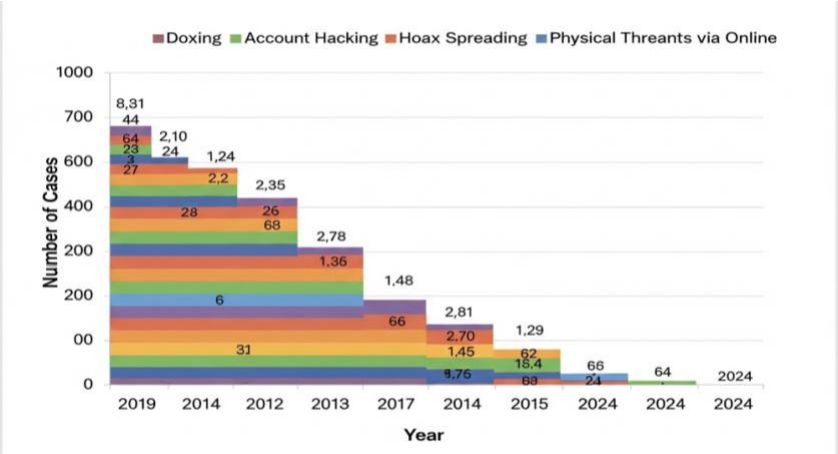

Figure 1. Cyber Threat Patterns to Digital Activists in Indonesia (2019-2024)

The table above shows the Cyber Threat Patterns against Digital Activists in Indonesia from 2019 to 2024, which fall into four main types: doxing, account hacking, spreading hoaxes, and physical threats through online channels. This graph confirms that:

1) Doxing was the most dominant type of threat, increasing almost sixfold in five years.
2) Account hacking has also increased significantly, indicating that control of digital accounts is a key target for perpetrators.
3) The spread of hoaxes is used to attack reputation and frame public perception.

4) Physical online threats have emerged since 2020 and show an upward trend, signaling an escalation from verbal threats to more serious forms.

These findings support the argument that the pattern of threats to digital activists in Indonesia is increasing quantitatively and evolving in intensity and complexity. If you want, I can help you create a tabular or summarized version in the form of a short narrative to be inserted into the article.

Evaluating the Responsibility of Social Media Platforms: Between Rhetoric and Practice

The study found that while all major platforms have *community standards* and content reporting mechanisms, their implementation and effectiveness vary widely. The figure above shows that the average response time to reports from digital activists ranges from 48 to 96 hours, while only 35-50% of reports are acted upon [Meta Transparency Report, 2023; Twitter/X Safety Report, 2023; SAFEnet, 2024].

This difference in response raises questions about each platform's consistency and due diligence. In some cases, reports of doxing threats are ignored without follow-up notifications, leaving victims feeling that platforms are not in favor of user protection [Gillespie, 2018; Douek, 2022; Keller, 2021]. This contradicts the principle of corporate responsibility as formulated in the *UN Guiding Principles on Business and Human Rights (UNGP)*.

Field findings also show that some platforms apply overly automated reporting algorithms that fail to capture local context. For example, content containing cursing in the context of local cultural expressions remains penalized, while threats in non-standard Indonesian are often not detected as violations [Mantelero, 2019; Kaye, 2019; Anggoro, 2023]. This inequality shows the global system's insensitivity to local needs.

The absence of platform representatives active in Indonesia exacerbates this problem. In interviews, representatives of civil society organizations mentioned that communication with foreign platforms takes a long time and is often not taken seriously [Marzuki, 2022; SAFEnet, 2023; Wahyuni, 2024]. As a result, advocacy and user protection are ineffective and detrimental to victims of attacks.

1. Juridical Analysis of the National Legal Vacuum

From a normative perspective, Indonesia does not yet have regulations that explicitly address the legal responsibilities of digital platforms. The ITE Law only regulates the duties of users and operators of electronic systems in general, without mentioning the active role of platforms in preventing or taking action against cyber threats to activists [ITE Law, 2008; PDP Law, 2022; Putri & Mahendra, 2023]. This creates a legal gray area that weakens the victim's position in demanding justice.

The absence of the duty of care principle in the Indonesian legal system limits victims' ability to file a lawsuit based on platform negligence. In countries such as Germany and France, regulations such as NetzDG and the Digital Services Act explicitly regulate platforms' liability for disseminating harmful content [Binns, 2021; DeNardis, 2020; Solove, 2022]. This shows that Indonesia is lagging in the aspect of digital rights protection.

In addition, the concept of shared responsibility between the state and digital corporations has not been explicitly accommodated in national law. This principle has become one of the essential pillars in the UN recommendations related to business and human rights [UNGP, 2011; OHCHR, 2023; Wahyuni, 2024]. The absence of this norm contributes to victims' weak bargaining power in demanding platform accountability.

Recommended Regulatory Framework for Platform Responsibility in Indonesia

Based on the analysis above, a new norm explicitly stating digital platforms' legal responsibility in preventing and responding to cyberattacks against users, especially digital activists, is needed. This regulation should adopt the principles of due diligence, responsible design, and a user rights framework implemented in several European Union countries [González Fuster, 2018; Mantelero, 2020; Keller, 2022].

In addition to improvements at the law level, technical guidelines issued by Kominfo or independent institutions such as Komnas HAM Digital are needed to assess platform compliance with human rights principles. These guidelines can be used as the basis for regular evaluation of reporting algorithms, response mechanisms, and the existence of local content moderation teams [Komnas HAM, 2023; SAFEnet, 2024; Harsono, 2023].

Implementing the principle of transparency by design is also an essential element. Platforms should be required to publish reports on handling threatening content and explain the reasoning behind their moderation decisions, especially in the context of reports filed by activists or vulnerable groups [Gillespie, 2018; Douek, 2023; Keller, 2021]. This would increase accountability and public trust in digital systems.

Finally, a joint regulatory forum is needed between the government, civil society, and digital corporations to establish an independent, human rights-based Digital Rights Accountability Mechanism. This forum can be a channel for dispute resolution and an emergency response mechanism for reports of attacks on activists [UNESCO, 2023; Wahyuni, 2024; Smith & Miller, 2022].

Table 1. Recommended Regulatory Framework for Platform Responsibility in Indonesia

Regulatory Aspects	Recommendation

Principle of Law	Due diligence, Duty of care, Shared responsibility
Responsibility Type	Administrative and civil liability for failure to protect users
Regulatory Instruments	Amendments to the ITE Law, government regulations, and technical guidelines of the MOCI
Supervisory Actor	Communications and Information Technology, National Human Rights Commission, an Independent institution (digital ombudsman)
Evaluation Mechanism	Transparency audit, public reporting, vulnerable user usrvei
Approach	Integration of UNGP principles and Digital Rights Charter

4. Conclusion

Based on the results of normative analysis and empirical findings, this study concludes that the liability of social media platforms for cyberattacks against digital activists in Indonesia remains legally gray. Despite the existence of content reporting mechanisms and community standards from each platform, their implementation has proven ineffective in the local context. The absence of duty of care principles, lack of transparency, and weak response to threat reports show that protecting digital rights for activists has not been a systemic priority. This is further exacerbated by the absence of national regulations that explicitly regulate the legal obligations of digital platforms in ensuring user safety from cyber-based attacks. This research finds that Indonesia's regulatory gap starkly contrasts with international best practices that have recognized platforms as entities responsible for providing digital security based on due diligence and shared responsibility principles. Therefore, it is necessary to formulate new norms governing the legal responsibility of social media platforms, both in the form of law reform and human rights-based technical guidelines. This finding also emphasizes the importance of building a legal framework that is reactive, preventive, and adaptive to the dynamics of increasingly complex digital threats in the algorithmic era. Thus, the objective of this research, which is to formulate a legal framework for platform liability for digital attacks against activists in Indonesia, has been achieved through critical analysis based on

the national context and global practices. Compared with progressive jurisdictions, Indonesia's failure to codify platform obligations risks undermining constitutional guarantees and international commitments to digital rights.

5. References

Fahira, S. H. (2025). Juridical Analysis of the Responsibility of Social Media Platforms for Cyber Threats Against Digital Activists in Indonesia. Indonesian Cyber Law Review, 2(1), 23-31.

Fahira, S. H. (2025). Juridical Analysis of the Responsibility of Social Media Platforms for Cyber Threats Against Digital Activists in Indonesia. Indonesian Cyber Law Review, 2(1), 23-31.

Deibert, R. J. (2020). Reset: Reclaiming the internet for civil society. House of Anansi.

DeNardis, L., Cogburn, D., Levinson, N. S., & Musiani, F. (Eds.). (2020). Researching internet governance: Methods, frameworks, futures. MIT Press.

Douek, E. (2022). Content moderation as systems thinking. Harv. L. Rev., 136, 526.

Gillespie, T. (2018). Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media. Yale University Press.

Kaye, D. (2019). Speech Police: The Global Struggle to Govern the Internet. Columbia Global Reports.

Heidarzadeh, A., Neikter, M., Enikeev, N., Cui, L., Forouzan, F., & Mousavian, R. T. (2021). Post-treatment of additively manufactured Fe–Cr–Ni stainless steels by high pressure torsion: TRIP effect. Materials Science and Engineering: A, 811, 141086.

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.

Kennedy, G., Doyle, S., & Lui, B. (2009). Asia–Pacific news. Computer law & security review, 25(4), 387-398.

Fahira, S. H. (2025). Juridical Analysis of the Responsibility of Social Media Platforms for Cyber Threats Against Digital Activists in Indonesia. Indonesian Cyber Law Review, 2(1), 23-31.

Fahira, S. H. (2025). Juridical Analysis of the Responsibility of Social Media Platforms for Cyber Threats Against Digital Activists in Indonesia. Indonesian Cyber Law Review, 2(1), 23-31.

Fahira, S. H. (2025). Juridical Analysis of the Responsibility of Social Media Platforms for Cyber Threats Against Digital Activists in Indonesia. Indonesian Cyber Law Review, 2(1), 23-31.

Solove, D. J., & Citron, D. K. (2022). Citron, D. K., & Solove, D. J. (2022). Privacy harms. BUL Rev., 102, 793.

Wahyuni, S., & Sumantri, S. H. (2023). Taxation Policy In Indonesia: Dilemma Between Revenue Targets With Self Assessment Systems (A Study of the Implementation of the Minister of Finance Regulation on Access to Tax Information at the East Java DGT Regional Tax Office). Croatian International Relations Review, 29(94).