

Criminal Liability Of Perpetrators Of Sara Related Cyber Crime Following The Entry Into Force Of The National Criminal Code

Adhi Luhur Wicaksono^{1*}, Muhammad Rifzal Alief Ramadhan², Farrah Eka Suryani³

^{1,2,3}Bussines Law Department, Sugeng Hartono University, Sukoharjo, Indonesia

Email: adhiluhurwicaksono@gmail.com

Abstract. Digital expansion has broadened expression while heightening the risk of hate speech based on ethnicity, religion, race, and inter-group relations (SARA). The issue has grown more complex since the National Criminal Code (Law Number 1 of 2023) took effect on 2 January 2026, coexisting with the 2024 amendment to the Electronic Information and Transactions (ITE) Law. This article analyses the construction of the offence and criminal liability for SARA-related cyber crime, the relationship between Article 28(2) jo. Article 45A(2) of the ITE Law and Articles 242–243 of the National Criminal Code, and the limits set by Constitutional Court Decisions Number 105/PUU-XXII/2024 and 115/PUU-XXII/2024. Using normative juridical research with statutory, conceptual, and case approaches, the study draws on legislation, Constitutional Court decisions, and recent reputable journals. Findings show Article 28(2) is no longer complete upon mere distribution; Decision 105 requires substantive identity-based hatred, intent, public conduct, and a genuine risk of discrimination, hostility, or violence, while Article 243 governs technology-mediated statements against a group. Liability thus demands proof of intent, targeted identity, context, and genuine risk. The article proposes a harmonisation model based on narrow *lex specialis*, constitutional avoidance, and risk-based thresholds to prevent overcriminalisation while protecting targeted groups.

Keywords: *Criminal Liability; Cyber Crime; Hate Speech; SARA; National Criminal Code*

INTRODUCTION

Digital transformation has turned the public sphere from a purely physical space into a hybrid space in which the production, distribution, and reproduction of information occur at a scale and speed unmatched by conventional communication. In this context, identity-based hate speech is not merely a matter of uncivil communication; it can function as an instrument for constructing

stigma, discrimination, hostility, and even violence against particular groups. A study by Ibrohim and Budi shows that detecting hate speech on Indonesian social media faces problems of contextuality, linguistic variation, and ambiguity between merely offensive and genuinely dangerous speech [1]. A parallel problem arises in criminal law: the state must protect groups from hatred, yet it must not turn every instance of criticism, satire, controversy, or religious expression into a criminal offence.

Within the Indonesian legal system, the legal regime governing digital hate speech has undergone a significant change. Law Number 1 of 2024 amended Article 28(2) of the ITE Law by defining the offence as the distribution or transmission of electronic information that incites, invites, or influences others in a manner that gives rise to hatred or hostility toward individuals and/or groups based on a particular identity. The corresponding penalty is set out in Article 45A(2), namely imprisonment for a maximum of six years and/or a fine of up to one billion rupiah [2]. Article II of the 2024 ITE Law, however, expressly provides that a number of its penal provisions, including Article 28(2) and Article 45A(2), remain in force only until Law Number 1 of 2023 on the Criminal Code takes effect [2]. This means that, since 2 January 2026, the application of these norms must be read within the context of the transition to, and harmonisation with, the National Criminal Code.

The National Criminal Code introduces a different construction. Article 242 criminalises public statements expressing hostility, hatred, or contempt toward groups of the Indonesian population on the basis of race, nationality, ethnicity, skin colour, religion, sex, or mental or physical disability. Article 243 extends this to statements made through information technology with the intent that they become known, or more widely known, to the public. Article 158, meanwhile, defines ‘in public’ to include spaces accessible through electronic media. This formulation shows that the National Criminal Code does not disregard cyberspace but instead integrates it into the general concept of offences against population groups [3].

The difficulty is that the coexistence of these norms does not automatically produce harmonisation. Putri, examining twenty-seven Indonesian court decisions, found that, prior to its amendment, Article 28(2) of the ITE Law lacked an adequate objective threshold, allowing courts to construe the elements of hatred and resulting harm inconsistently [4]. This finding matters because it shows that the central problem is not simply whether hate speech should be criminalised, but how the law draws the line between protected expression and expression that warrants punishment.

This problem acquired a constitutional dimension through Constitutional Court Decision Number 105/PUU-XXII/2024. The Court held that Article 28(2) and Article 45A(2) of the ITE Law are conditionally unconstitutional unless construed to cover only electronic information or documents that substantively contain acts or the dissemination of hatred based on a particular identity, committed intentionally and in public, and giving rise to a genuine risk of discrimination, hostility, or violence [5]. This decision materially changes how the offence must be read: mere distribution of content is no longer sufficient; there must be a substantive quality of hatred and a legally provable genuine risk.

Decision Number 115/PUU-XXII/2024 is likewise relevant, although its main object differs. The Court gave a constitutional interpretation of the word 'unrest' (kerusuhan) in Article 28(3) and Article 45A(3) of the ITE Law, confining it to conditions that disturb public order in the physical world rather than purely digital conditions [6]. This decision reinforces the principle that the element of consequence in cyber criminal law may not be interpreted abstractly or without limit. Within criminal policy, such a limit is necessary to prevent the expansion of criminal law to cover every controversial form of digital communication.

Recent literature also shows that digital hate speech is a cross-jurisdictional problem. Butler and Turenne emphasise that the regulation of online hate speech must take account of variations in criminal standards, enforcement mechanisms, and freedom-of-expression protections [7]. Enarsson argues that content moderation requires a balance between protection from hate speech and the right to freedom of expression [8], while Alkiviadou shows that heightened platform liability can create a risk of over-removal and undue restriction of expression where legal obligations are not accompanied by adequate safeguards [9]. These findings are relevant for Indonesia now that the National Criminal Code has entered into force and corporations are recognised as subjects of criminal law.

Prior research in Indonesia has largely examined hate speech from the perspective of law enforcement, individual liability, or the effectiveness of the ITE Law [10][11]. In the offline domain, research on criminal law policy toward racial discrimination from a Pancasila perspective shows that Law Number 40 of 2008 on the Elimination of Racial and Ethnic Discrimination had, in fact, already formulated a criminal policy consistent with the spirit of Pancasila, yet its implementation at the level of law enforcement still leaves much to be desired, calling for further dialogue and legal reform [21]. Research on corporate liability under the ITE Law has likewise demonstrated a doctrinal basis for treating corporations as subjects of criminal law [12]. Nevertheless, a gap remains in the

analysis of the concrete relationship between the 2024 ITE Law, the now-effective National Criminal Code, recent Constitutional Court decisions, and the construction of criminal liability for both individual and corporate perpetrators in the context of SARA-related cyber crime.

Building on this gap, the article addresses three questions. First, how should SARA-related cyber crime be juridically qualified following the entry into force of the National Criminal Code? Second, how should criminal liability be constructed for both individual and corporate perpetrators in such cases? Third, how do Constitutional Court Decisions Number 105/PUU-XXII/2024 and 115/PUU-XXII/2024 affect the limits of criminalisation and the standard of proof? The novelty of this article lies in bringing these three dimensions together within a single framework: normative harmonisation, criminal liability, and the constitutional threshold of genuine risk.

METHODS

This is normative legal research of a prescriptive character. Law is treated as a body of norms whose coherence must be analysed through the relationships among rules, principles, doctrines, and court decisions. The study employs a statute approach, a conceptual approach, and a case approach. The statute approach is used to examine Law Number 1 of 2023 on the Criminal Code, Law Number 1 of 2024 on the Second Amendment to the ITE Law, Law Number 40 of 2008 on the Elimination of Racial and Ethnic Discrimination, and the 1945 Constitution of the Republic of Indonesia. The case approach is applied chiefly to Constitutional Court Decisions Number 78/PUU-XXI/2023, 105/PUU-XXII/2024, and 115/PUU-XXII/2024.

Secondary legal materials were selected with particular attention to articles addressing hate speech, freedom of expression, cyber law, content moderation, and corporate criminal liability. The analysis proceeds in three stages: (1) inventory and mapping of the elements of the offence; (2) systematic and constitutional interpretation of the relationship between the ITE Law and the National Criminal Code; and (3) formulation of a construction of criminal liability that satisfies the principles of legality, culpability, proportionality, and the protection of freedom of expression. This approach modifies classical normative legal research by incorporating a risk-based threshold analysis drawn from Constitutional Court decisions and recent hate-speech literature.

RESULT AND DISCUSSION

1. Reconstructing the Juridical Qualification of SARA-Related Cyber Crime

SARA-related cyber crime must be distinguished from cyber-dependent crime. In cyber-dependent crime, technology is itself the primary object of attack, as in malware, unauthorised access, or system interference. In a cyber-enabled content offence, by contrast, technology serves merely as the means of extending the reach of conduct that fundamentally attacks human interests or social order. SARA-related hate speech falls within the latter category. Accordingly, the ‘cyber’ character of an act should never, by itself, be the basis for criminalisation; what matters is the substance of the expression, the manner of dissemination, the targeted identity, the context, and the risk created.

Article 28(2) of the 2024 ITE Law initially placed distribution/transmission at the core of the offence. Following Constitutional Court Decision Number 105/PUU-XXII/2024, however, this formulation must now be read against a constitutional threshold. The Court requires that the information substantively contain acts or the dissemination of hatred based on a particular identity, be committed intentionally and in public, and give rise to a genuine risk of discrimination, hostility, or violence [5]. The offence can therefore no longer be understood as a purely formal one that is complete merely because the ‘upload’ or ‘send’ button has been pressed.

This shift is important for distinguishing hate speech from merely offensive speech. Not every insult, harsh criticism, religious statement, joke, or offensive comment amounts to hate speech. Putri shows that the previous lack of an objective threshold produced inconsistent assessments of harm [4]. Conceptually, an expression becomes criminally relevant only where a demonstrable link exists between the expression, the identity of the targeted group, the perpetrator’s intent or awareness, and a genuine risk of discrimination, hostility, or violence. This model aligns with a stricter, harm-based approach rather than a broader offence-based one.

Articles 242 and 243 of the National Criminal Code are structured differently. Article 242 is oriented toward statements of hostility, hatred, or contempt against particular population groups, whereas Article 243 governs the form of such statements when communicated through information technology. Because the National Criminal Code is now in force, these provisions must serve as the starting point for determining the substantive criminal law applicable to conduct occurring after 2 January 2026. The 2024 ITE Law itself limits the validity of a number of its penal provisions until the National Criminal Code takes effect [2], reflecting a transitional legislative design rather than an intention to permanently maintain two identical penal norms.

For this reason, the principle of *lex specialis derogat legi generali* cannot be applied mechanically to always place the ITE Law above the National Criminal Code. Once the transitional period has ended, it must be determined whether the relevant ITE Law norm remains in force for a given object, whether its penal provisions have lapsed under Article II of the 2024 ITE Law, and whether any relevant transitional or special provisions apply. From a harmonisation perspective, cyber law should function primarily to regulate electronic systems and evidence, while the criminalisation of hate-speech content should rest on whichever penal norm remains in force.

2. Elements of the Offence and the Standard of Proof Following Constitutional Court Decision Number 105/PUU-XXII/2024

Constitutional Court Decision Number 105/PUU-XXII/2024 marks a turning point by introducing the element of ‘genuine risk’ into the interpretation of Article 28(2) and Article 45A(2) of the ITE Law [5]. Methodologically, this element can be operationalised through five indicators: (a) the identity of the targeted group must be clear; (b) the substance of the content must contain advocacy or incitement of hatred, rather than mere criticism; (c) there must be intent, or at least awareness, of the hateful character of the content; (d) the communication must occur in a space accessible to the public; and (e) there must be a contextually explicable genuine risk of discrimination, hostility, or violence.

These indicators do not mean that the public prosecutor must prove that actual violence has occurred. ‘Genuine risk’ operates as a threshold between an abstract offence and an actual consequence. What must be proven is a sufficient probability of, and connection to, potential harm. In this context, audience size, the character of the account, the socio-political moment, the form of the appeal, repetition, the use of demeaning symbols against a group, and the ensuing response may all serve as contextual facts. Sheer viewer numbers, however, must never be treated as automatic proof that an expression constitutes a criminal offence.

International literature supports this caution. Hassett-Walker shows that applying the incitement doctrine to online hate speech is problematic where expression spreads rapidly but its causal link to violent acts is not always direct [13]. Farrand likewise shows that coded language, memes, and online community context make it difficult to identify whether particular content constitutes hate speech or ordinary communication [14]. Indonesian law enforcement therefore

needs to develop contextual proof, rather than relying solely on a lexical analysis of words.

This approach is also consistent with the requirements of digital evidence. Screenshots stripped of context are not sufficient evidence. Investigators must verify account authenticity, the time of posting, the link between the account and the perpetrator, the conversational context, any alteration or deletion of content, and available metadata. In cases turning on linguistic meaning, forensic linguistic analysis can help distinguish descriptive statements, criticism, insult, satire, incitement, and appeal. The standard of proof must accordingly be multi-layered, combining digital forensic evidence, linguistic context, social context, and evidence of intent.

3. Individual Criminal Liability: Actus Reus, Mens Rea, and Capacity for Responsibility

Individual criminal liability does not arise merely because a person can be identified as the one who uploaded certain content. Under the National Criminal Code, the principle of culpability forms the basis of criminal liability. Because the hate-speech offence essentially requires intentional conduct, the public prosecutor must prove a link between the perpetrator's intent and the substance of the disseminated content. Intent may be inferred from word choice, editing or producing the content, repeated distribution, captions, hashtags, accompanying comments, and the perpetrator's response after being informed of the content's character.

In practice, the most difficult issue is distinguishing the original creator of content from those who merely pass it on. Reposting, sharing, retweeting, quoting, duetting, stitching, and forwarding do not all carry the same weight of mens rea. A person who redistributes content in order to criticise or rebut it cannot automatically be equated with one who redistributes it to affirm an incitement to hatred. Communicative context must therefore form part of the proof of intent.

Research by Marbun and Hartanto shows that criminal liability for hate speech on social media is closely tied to freedom of opinion and the limits of an individual's duty to answer for their own expression [11]. Putri's research, meanwhile, shows that inconsistency in assessing harm can widen judicial discretion [4]. These findings reinforce the need for judicial guidelines so that mens rea and harm are not determined solely on the basis of subjective sensitivity to a given statement.

In practical terms, individual liability may be constructed as follows: first, there must be an act of digital communication satisfying the objective elements of the offence; second, the perpetrator must possess the capacity for responsibility; third, the perpetrator must have known and intended, or at least been aware of, the character of the act as defined by the offence; fourth, there must be a targeted group based on a protected identity; and fifth, where required by the applicable formulation, there must be a genuine risk of discrimination, hostility, or violence. If any of these core elements is not proven, the application of criminal punishment must be discontinued in accordance with the principle of *in dubio pro reo* and the principle of legality.

4. Corporate Liability and the Limits of Digital Platform Responsibility

The National Criminal Code affirms corporations as subjects of criminal law and regulates corporate liability in Articles 45 to 50. This change matters for technology-based crime because the distribution of digital content is not always the act of a lone individual. Platforms shape system design, recommendation algorithms, monetisation policy, moderation mechanisms, and complaint-handling procedures. Recognising corporations as criminal subjects, however, does not mean that every piece of illegal content appearing on a platform automatically constitutes a corporate offence.

Rahayu and Lukitasari find that the ITE Law regime has opened the door to corporate liability through an identification approach [12]. Research on the evolving models of corporate criminal liability under the National Criminal Code likewise points to the growing use of identification, corporate-culture, management-failure, and strict/vicarious liability models within Indonesian criminal law discourse [15]. These developments provide a conceptual basis for assessing corporate fault, yet their application must still be tailored to the formulation of the offence and the principle of culpability.

For digital platforms, a management-failure or corporate-culture model is more appropriate than strict liability. Corporate liability should only be considered where it can be shown that a decision or omission at the organisational level bears a sufficient connection to the commission of the offence for instance, an internal policy that knowingly disregards a known risk, a systemic failure to meet legal obligations, or a business decision that promotes the distribution of hateful content without reasonable mitigation. Conversely, a platform should not be punished merely for its inability to remove all harmful content in real time.

International scholarship on platform liability reveals the same dilemma. Alkiviadou warns that overly aggressive content-removal obligations can threaten freedom of expression [9], while Frosio highlights the importance of due process and the protection of fundamental rights within platform liability regimes [16]. If Indonesian law is to develop corporate liability in the context of content moderation, the applicable standard should therefore be based on knowledge, control, foreseeability, and failure to act, rather than automatic liability.

Within the National Criminal Code framework, proof of corporate fault may be directed at four questions: who made the decision; whether that decision fell within the scope of the corporation's business activities; whether the organisation knew or ought to have known of the risk of an offence; and whether reasonable preventive measures were available but not taken. This model preserves the balance between effective law enforcement and the principle of personal culpability.

5. Implications of Constitutional Court Decisions Number 78/PUU-XXI/2023 and 115/PUU-XXII/2024

Constitutional Court Decision Number 78/PUU-XXI/2023 held that Articles 14 and 15 of Law Number 1 of 1946 conflict with the 1945 Constitution and are no longer legally binding [18]. This decision matters as constitutional background, establishing that penal norms concerning false information may not be formulated or applied so broadly as to threaten freedom of expression. Although its object was not Article 28(2) of the ITE Law, its reasoning reinforces the principle of caution in criminalising expression.

Constitutional Court Decision Number 115/PUU-XXII/2024 has a more specific effect on the relationship between digital space and criminal consequence. The Court interpreted 'unrest' (kerusuhan) in Article 28(3) and Article 45A(3) as referring solely to conditions that disturb public order in the physical world, not merely conditions in digital space [6]. This decision teaches that any term with the potential to expand criminalisation must be given an objectively testable limit.

Read together, these three decisions establish a constitutional threshold within cyber criminal law: a penal norm must have a clear target, an identifiable act, a sufficient link between expression and harm, and protection for legitimate criticism and opinion. Consequently, it is not enough for law enforcement officials merely to show that content went viral or gave offence to a particular

group; they must demonstrate why the content satisfies every element of the offence and why the resulting restriction on expression is proportionate.

This framework is also consistent with the need of democratic states to distinguish hate speech from dissent. Paulos and Çelik emphasise that the regulation of hate speech on social media is always in tension with freedom of expression [17]. In the Indonesian context, this tension should be resolved not by abolishing criminalisation altogether, but by tightening the legal threshold and the quality of proof required.

6. A Model for Harmonising the ITE Law and the National Criminal Code

Table 1. Harmonisation Matrix for Norms Governing SARA-Related Cyber Crime

Aspect	2024 ITE Law	National Criminal Code	Enforcement Implications
Object	Electronic information/documents	Statements against a group; Article 243 for IT-based means	After 2026, the primary focus shifts to the National Criminal Code once the relevant ITE norm has lapsed
Mens rea	Intentional and unauthorised conduct; following Decision 105, must be read as intentional incitement of hatred	Intent is inherent in the statement of hostility/hatred/contempt	Proof must not stop at mere account ownership or distribution
Harm	Decision 105 requires a genuine risk of discrimination, hostility, or violence	Articles 242–243 focus on statements against a group	Context, target, and risk become the distinguishing factors
Digital space	Explicitly covered through	Article 243 and the definition of “in	Technology does not automatically determine the

	distribution/transmission	public” accommodate electronic media	applicable norm; the substance of the act remains decisive
Corporations	Basis for liability found in the ITE Law and the corporate liability regime	Articles 45–50 of the Criminal Code regulate corporations generally	Proof of organisational fault is required; not automatic platform liability

The matrix shows that harmonisation must not be achieved simply by selecting whichever norm carries the heavier penalty. Once the National Criminal Code is in force, the first step is to determine the applicable law based on the *tempus delicti* and the transitional provisions. Where the conduct occurs after 2 January 2026 and the relevant penal provision of the ITE Law has lapsed under Article II of the 2024 ITE Law, Articles 242–243 of the National Criminal Code become the primary point of reference. The ITE Law remains relevant for matters concerning electronic systems, evidence, the obligations of electronic system operators, and any non-penal provisions that remain in force.

The proposed harmonisation model consists of three layers. The first is a legality filter: whether the norm relied upon remains in force and whether the conduct falls within the formulation of the offence. The second is a constitutional filter: whether the application of the norm satisfies the Constitutional Court’s interpretation, particularly regarding group identity, intent, public space, and genuine risk. The third is a proportionality filter: whether resorting to criminal law is a proportionate response compared with non-penal mechanisms such as clarification, a right of reply, moderation, or restorative measures. These three layers may guide investigators, prosecutors, and judges alike.

This model also addresses concerns about overly elastic (‘rubber’) provisions. Consistent law enforcement does not mean expanding criminalisation but improving the quality of case selection. Only content that passes all three filters should be processed as a criminal offence. In this way, the state remains able to protect society from hate speech without turning criminal law into an instrument for controlling public opinion.

7. Criminal Policy and the Unity of Indonesia

Criminal policy toward hate speech must be situated within a framework that protects both society and individual rights. Unity of Indonesia (*Persatuan Indonesia*), as a constitutional and philosophical value, provides the state with grounds to prevent communication that genuinely drives discrimination,

hostility, or violence. This perspective is consistent with scholarship that positions Pancasila particularly its second principle, ‘a just and civilised humanity’ as the grundnorm underlying criminal law policy on racial discrimination, such that the success of criminal policy should be measured by the extent to which Pancasila’s values are realised in law-enforcement practice, rather than merely by the existence of a penal norm [21]. Unity, however, is not synonymous with uniformity of opinion. Democracy, indeed, presupposes protected differences of view, provided that such differences do not turn into attacks on the rights and dignity of other groups.

For this reason, an ideal criminal policy is not simply about imposing harsher sanctions but about building a precise ecosystem of prevention and enforcement. Law enforcement officials need guidelines on proof; courts need a framework for assessing harm; platforms need transparent moderation procedures; and society needs digital literacy. Studies on online hate speech likewise show that effective control depends not only on criminal law but on a combination of regulation, moderation, and public participation [8][16].

From a criminal-law perspective, the principle of *ultimum remedium* does not mean that hate speech must always be tolerated until it results in actual violence. Prevention is warranted once a genuine risk can be demonstrated. The use of criminal law must nevertheless remain bounded by the principles of legality, culpability, proportionality, and fair proof. This is the point of balance between protecting unity and protecting freedom of expression.

CONCLUSION

The entry into force of the National Criminal Code on 2 January 2026 has reshaped the architecture of criminal liability for SARA-related cyber crime. Digital hate speech can no longer properly be understood merely as a matter of distributing electronic information; it must instead be examined as an offence to be tested against the substance of the expression, the identity of the targeted group, intent, the public-space context, and the genuine risk of discrimination, hostility, or violence. Constitutional Court Decision Number 105/PUU-XXII/2024 provides an essential foundation by narrowing Article 28(2) of the ITE Law through constitutional interpretation, while Decision Number 115/PUU-XXII/2024 reinforces the need for an objective limit on the element of consequence in digital space. Following the entry into force of the National Criminal Code, Articles 242–243 must serve as the primary point of reference in accordance with the *tempus delicti* and transitional provisions, while the ITE Law remains relevant for matters concerning electronic systems and any non-penal provisions that remain in force. Individual liability requires proof of *actus reus*, *mens rea*,

and capacity for responsibility; corporate liability requires proof of organisational fault and cannot rest merely on the presence of illegal content. The principal contribution of this article is a three-layer harmonisation model — a legality filter, a constitutional filter, and a proportionality filter — intended as a guide for preventing overcriminalisation while preserving legal protection for groups targeted by hatred. This model also carries forward the agenda of strengthening Pancasila values within anti-discrimination criminal policy first advanced in the offline domain through Law Number 40 of 2008 on the Elimination of Racial and Ethnic Discrimination [21], so that a Pancasila-based approach to SARA-related conduct can be applied consistently in both physical and digital spaces. The model may serve as a basis for developing law-enforcement guidelines, platform moderation policy, and empirical research on the application of the National Criminal Code to cases of SARA-related cyber crime.

REFERENCES

- [1] Ibrohim, M. O., & Budi, I. (2023). Hate speech and abusive language detection in Indonesian social media: Progress and challenges. *Heliyon*, 9(8), e18647. <https://doi.org/10.1016/j.heliyon.2023.e18647>
- [2] Government of the Republic of Indonesia. (2024). Law Number 1 of 2024 on the Second Amendment to Law Number 11 of 2008 on Electronic Information and Transactions [Undang-Undang Nomor 1 Tahun 2024 tentang Perubahan Kedua atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik]. https://jdih.komdigi.go.id/produk_hukum/view/id/884/t/undangundang%20nomor%201%20tahun%202024
- [3] Government of the Republic of Indonesia. (2023). Law Number 1 of 2023 on the Criminal Code [Undang-Undang Nomor 1 Tahun 2023 tentang Kitab Undang-Undang Hukum Pidana]. State Gazette of the Republic of Indonesia of 2023 Number 1.
- [4] Putri, D. K. (2023). Hate speech and the harm in Indonesian judicial decisions. *Cogent Social Sciences*, 9(2), 2274430. <https://doi.org/10.1080/23311886.2023.2274430>
- [5] Constitutional Court of the Republic of Indonesia. (2025). Decision Number 105/PUU-XXII/2024. <https://www.mkri.id/perkara/persidangan/putusan>
- [6] Constitutional Court of the Republic of Indonesia. (2025). Decision Number 115/PUU-XXII/2024. <https://www.mkri.id/perkara/persidangan/putusan>

- [7] Butler, O., & Turenne, S. (2022). The regulation of hate speech online and its enforcement – a comparative outlook. *Journal of Media Law*, 14(1), 20–24. <https://doi.org/10.1080/17577632.2022.2092261>
- [8] Enarsson, T. (2024). Navigating hate speech and content moderation under the DSA: insights from ECtHR case law. *Information & Communications Technology Law*, 33(3), 384–401. <https://doi.org/10.1080/13600834.2024.2395579>
- [9] Alkiviadou, N. (2025). Platform liability, hate speech and the fundamental right to free speech. *Information & Communications Technology Law*, 34(2), 207–217. <https://doi.org/10.1080/13600834.2024.2411799>
- [10] Ramadhani, W., & Hasibuan, Z. L. M. (2024). Tinjauan hukum atas perbuatan ujaran kebencian [A legal review of hate-speech conduct]. *Indonesian Journal of Criminal Law and Criminology*, 5(2). <https://doi.org/10.18196/ijclc.v5i2.21915>
- [11] Marbun, T. N. W., & Hartanto. (2024). Criminal responsibility of criminal acts of hate speech through social media. *JILPR Journal Indonesia Law and Policy Review*, 5(2), 409–422. <https://doi.org/10.56371/jirpl.v5i2.230>
- [12] Rahayu, H. S., & Lukitasari, D. (2021). The concept of corporate criminal liability in the Law on Information and Electronic Transactions. *Indonesian Journal of Criminal Law Studies*, 6(1). <https://doi.org/10.15294/ijcls.v6i1.38400>
- [13] Hassett-Walker, C. (2022). Does Brandenburg v. Ohio still hold in the social media era? Racist (and other) online hate speech and the First Amendment. *Cogent Social Sciences*, 8(1), 2038848. <https://doi.org/10.1080/23311886.2022.2038848>
- [14] Farrand, B. (2023). ‘Is This a Hate Speech?’ The Difficulty in Combating Radicalisation in Coded Communications on Social media Platforms. *European Journal on Criminal Policy and Research*, 29, 477–493. <https://doi.org/10.1007/s10610-023-09543-z>
- [15] Alfakar, M. W., Masyhar, A., Wulandari, C., & Nte, N. D. (2024). Evolution of corporate criminal liability models and theories under Indonesian New Criminal Code. *Indonesian Journal of Criminal Law Studies*, 8(2). <https://doi.org/10.15294/ijcls.v8i2.36741>
- [16] Frosio, G. (2023). Taking fundamental rights seriously in the Digital Services Act's platform liability regime. *European Law Journal*. <https://doi.org/10.1111/eulj.12475>
- [17] Paulos, B., & Çelik, S. (2021). The challenges of regulating hate speech on social media in light of the theory of freedom of expression. *Süleyman Demirel Üniversitesi Hukuk Fakültesi Dergisi*, 11(1), 97–134. <https://doi.org/10.52273/sduhfd.922588>

- [18] Constitutional Court of the Republic of Indonesia. (2024). Decision Number 78/PUU-XXI/2023. <https://www.mkri.id/perkara/persidangan/putusan>
- [19] Government of the Republic of Indonesia. (1945). The 1945 Constitution of the Republic of Indonesia [Undang-Undang Dasar Negara Republik Indonesia Tahun 1945].
- [20] Government of the Republic of Indonesia. (2008). Law Number 40 of 2008 on the Elimination of Racial and Ethnic Discrimination [Undang-Undang Nomor 40 Tahun 2008 tentang Penghapusan Diskriminasi Ras dan Etnis].
- [21] Wicaksono, A. L., & Soponyono, E. (2022). Kebijakan Hukum Pidana dalam Upaya Penanganan Diskriminasi Rasial dengan Perspektif Pancasila [Criminal law policy in addressing racial discrimination from a Pancasila perspective]. *Jurnal Belo*, 8(1), 48–66. <https://doi.org/10.30598/belovol8issue1page48-66>