



Research Article



## Combating Digital Defamation: Regulations, Challenges and Protecting Reputation

Anis Mashdurohatusun <sup>1,\*</sup>, Bambang Sugihartono <sup>1</sup>, Nurul Masrifah <sup>1</sup>, Arief Indra Kusuma Adhi <sup>1</sup>, Henning Glaser <sup>2</sup>

<sup>1</sup> Faculty of Law, Universitas Islam Sultan Agung, Semarang, Indonesia

<sup>2</sup> German-Southeast Asian Center of Excellence for Public Policy and Good Governance (CPG), Bangkok, Thailand.

\* Correspondence: [anism@unissula.ac.id](mailto:anism@unissula.ac.id)

Received: March 11, 2025 / Accepted: September 10, 2025 / Published: September 16, 2025

**Abstract:** Indonesia regulates digital defamation through the ITE Law; however, the existing framework remains inadequate because it fails to achieve a balanced integration of reputation protection, freedom of expression, legal coherence, institutional coordination, and public awareness. This research aims to analyze regulatory inadequacies on digital defamation, identify weaknesses in the existing framework, and propose a justice-based reconstruction for Indonesia's legal system. Employing a constructivist paradigm, this study adopts a socio-legal research approach with a descriptive type of analysis. Data were derived from secondary sources, including primary, secondary, and tertiary legal materials, collected through library research and analyzed qualitatively. The research demonstrates that Indonesia's regulation of digital defamation remains unjust because it fails to adequately protect citizens' right to express opinions without surveillance, restriction, or silencing. Substantively, lawmakers must critically reconsider Article 45 paragraph (4) of the ITE Law to ensure fair and consistent enforcement. Structurally, law enforcement institutions must strengthen coordination between the police, prosecutors, and judiciary to address weak institutional synergy. Culturally, policymakers and stakeholders must expand public dissemination of the ITE Law to improve legal understanding across society. Based on these findings, the study concludes that reconstructing the regulation of digital defamation requires embedding justice-oriented values into both legal norms and enforcement mechanisms, particularly through reforming Article 27A and Article 45 paragraph (5) of the ITE Law.

**Keywords:** Criminal Act; Defamation; Protection; Regulation.



This is an open-access article under the [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/) license

## INTRODUCTION

Indonesia explicitly regulates digital defamation in Law Number 1 of 2024 concerning the Second Amendment to Law Number 11 of 2008 on Electronic Information and Transactions (ITE Law). Article 27A stipulates that any individual who intentionally attacks another person's honor or reputation by making an allegation, with the intent that it becomes publicly known, when conducted through electronic information and/or electronic documents via an electronic system, may be subject to criminal liability. Article 45 paragraph (4) further prescribes criminal sanctions for violations of this provision, imposing a maximum imprisonment of two years and/or a fine of up to IDR 400,000,000.<sup>1</sup>

Normatively, the provision seeks to protect individual dignity and honor in the digital sphere. However, its implementation reveals fundamental challenges. Critics

<sup>1</sup> Sue Spaid, 'Chapter 18 - On Deterring Hate Speech, While Maximizing Security and Privacy', in *Ethics in Online AI-Based Systems*, ed. by Santi Caballé, Joan Casas-Roma, and Jordi Conesa, Intelligent Data-Centric Systems (Academic Press, 2024), pp. 365–81 <https://doi.org/10.1016/B978-0-443-18851-0.00021-4>

often highlight four recurring issues: first, the use of the article as a tool for personal revenge; second, its application as a shock therapy instrument; third, its function in silencing criticism; and fourth, its use as a bargaining tool in legal settlements. These issues generate ongoing debates regarding the balance between protecting individual honor and safeguarding freedom of expression.<sup>2</sup> The case of Prita Mulyasari illustrates this tension. She faced charges of digital defamation after expressing dissatisfaction with hospital services through an email distributed on a mailing list. Nevertheless, in its Judicial Review Decision Number 22 PK/Pid.Sus/2011, the Supreme Court ruled that her actions did not constitute criminal defamation. This landmark ruling set an important precedent affirming that criticism, particularly concerning public services, cannot automatically be classified as a criminal offense.<sup>3</sup>

Over the past two decades, the advancement of information and communication technology in Indonesia has fundamentally transformed the dynamics of social, political, and economic interaction. Social media now serves as an active and indispensable element of the digital ecosystem, functioning as the primary mode of communication in contemporary society. The Indonesian Internet Service Providers Association (APJII) reported that in 2024, 221,563,479 individuals out of a total population of 278,696,200 accessed the internet, producing a national penetration rate of 79.5 percent. This data indicates not only the extensive integration of Indonesian society into the digital sphere but also the necessity for the state to exercise comprehensive regulatory authority over online activities. Consequently, the urgency of legal governance in the digital domain intensifies to ensure that technological progress remains consistent with the principles of rights protection, the fulfillment of civic obligations, and the maintenance of legal certainty for all citizens.<sup>4</sup>

One issue that the government is currently addressing concerns the proposed policy requiring all social media accounts to be registered with official identification. Through this policy, the government seeks to mitigate the risks of digital crime, particularly the dissemination of hoaxes, hate speech, and defamation. This initiative is highly relevant because social media today functions not only as a medium of personal interaction but also as a communication and public service instrument for both governmental and private institutions. From a legal perspective, defamation has long been a subject of contention. Historically, insults directed at the President or Vice President were regulated under Articles 134, 136 bis, and 137 of the Criminal Code. However, the Constitutional Court, through Decision Number 013-022/PUU-IV/2006, invalidated these provisions on the grounds that they posed a threat to freedom of expression. Despite this ruling, law enforcement practices continue to demonstrate a tendency to apply defamation provisions in controversial contexts, thereby raising

---

<sup>2</sup> Anna Demenko, 'Jurisdiction in Online Speech Crime Cases in the European Union', *International Journal of Law, Crime and Justice*, 80 (2025), 100718 <https://doi.org/https://doi.org/10.1016/j.ijlcj.2024.100718>

<sup>3</sup> Aloysius Wisnubroto and Hilaire Tegnau, 'Preventing AI Crime Towards A New Legal Paradigm: Lessons From United States', *Journal of Human Rights, Culture and Legal System*, 5.2 (2025), 630–58 <https://doi.org/10.53955/jhcls.v5i2.606>

<sup>4</sup> Musarat Yasmin, Tahira Jabeen and Sania Noor, 'Examining Digital Hate: Gendered Discursive Strategies in Online Harassment of Pakistani Influencers', *Computers in Human Behavior Reports*, 19 (2025), 100759 <https://doi.org/https://doi.org/10.1016/j.chbr.2025.100759>

persistent debates over the balance between protecting reputation and safeguarding democratic freedoms.<sup>5</sup>

Several prominent cases concretely illustrate the ongoing debate over the application of defamation laws in the digital sphere. The Prita Mulyasari case exemplifies how expressions of dissatisfaction with hospital services through digital media escalated into criminal proceedings, although the Supreme Court ultimately acquitted Prita through a judicial review. In a similar vein, the case of Haris Azhar and Fatia Maulidiyanti highlighted the judiciary's recognition of freedom of expression, as the panel of judges concluded that the phrase "Lord Luhut" in a podcast did not constitute defamation but instead represented legitimate public discourse. This decision affirmed that criticism directed toward public officials must be understood within the framework of democratic freedoms. Conversely, the Adam Deni case reflects a more repressive application of defamation provisions. The court found him guilty and imposed a prison sentence for remarks deemed defamatory toward a member of the House of Representatives in mass media. A comparable judicial stance emerged in the case of Jovi Andrea Bachtiar, a prosecutor convicted under Article 45 paragraph (4) of the ITE Law for criticizing a colleague on social media. These cases collectively demonstrate the persistent and robust enforcement of defamation provisions across diverse forms of digital interaction, thereby revealing the tension between legal protection of reputation and the preservation of expressive freedoms.<sup>6</sup>

This phenomenon underscores a fundamental legal dilemma between, on the one hand, the necessity of protecting individual honor and maintaining a healthy public sphere, and, on the other hand, the state's constitutional obligation to safeguard freedom of expression. From the standpoint of democratic theory, public criticism of government officials should function as a legitimate instrument of social control rather than be construed as a personal affront. Accordingly, a comprehensive reassessment of defamation regulations, particularly those contained in the Electronic Information and Transactions (ITE) Law, has become imperative. Policy reform must be undertaken to ensure that defamation provisions operate not as tools of repression but as balanced legal instruments that simultaneously uphold individual dignity and protect citizens' expressive freedoms in a proportional and just manner.<sup>7</sup>

The Indonesian legal system explicitly regulates the concept of defamation through the Criminal Code (KUHP), particularly in Articles 310 to 321. Article 310 paragraph (1) defines defamation as an act that unlawfully generates fear or hatred toward an individual or institution. Article 310 paragraph (2) further specifies written defamation, namely the act of producing or disseminating written material that damages another party's reputation. Article 311 addresses slander, which encompasses the deliberate

---

<sup>5</sup> Gloria del Valle-Cano and others, 'SocialHaterBERT: A Dichotomous Approach for Automatically Detecting Hate Speech on Twitter through Textual Analysis and User Profiles', *Expert Systems with Applications*, 216 (2023), 119446 <https://doi.org/https://doi.org/10.1016/j.eswa.2022.119446>

<sup>6</sup> Rustamaji Muhammad and others, 'The Reduction of Criminal Justice Policy in Indonesia: Justice versus Virality', *Journal of Human Rights, Culture and Legal System*, 5.2 (2025), 442–72 <https://doi.org/10.53955/jhcls.v5i2.637>

<sup>7</sup> Olha Kaminska, Chris Cornelis and Veronique Hoste, 'Fuzzy Rough Nearest Neighbour Methods for Detecting Emotions, Hate Speech and Irony', *Information Sciences*, 625 (2023), 521–35 <https://doi.org/https://doi.org/10.1016/j.ins.2023.01.054>

spread of false accusations intended to harm a person's image.<sup>8</sup> Article 315 regulates minor insults that are carried out publicly and result in reputational harm. In addition, Article 317 prohibits the filing of false reports, while Article 318 criminalizes false accusations made with the intent of discrediting others. Articles 320 and 321 extend legal protection to deceased persons by criminalizing insults directed at the dead. Collectively, these provisions underscore the seriousness with which Indonesian criminal law protects human dignity, honor, and reputation. Nevertheless, the emergence of digital platforms has introduced significant challenges. In the digital age, defamation frequently occurs through social media, which possesses a reach and impact far greater than that of conventional media, thereby complicating the enforcement of existing legal provisions.<sup>9</sup>

The new Criminal Code (KUHP) actively regulates insult and defamation under Articles 433 to 450, emphasizing the dual objectives of safeguarding individual dignity and maintaining a balance with freedom of expression. Article 433 defines defamation as the act of attributing a statement or allegation to another person in a manner that damages their honor or reputation, while Article 434 extends this provision to written defamation. Article 435 prescribes sanctions for spreading slanderous accusations, and Article 436 expands the scope of regulation by addressing insults committed in public spaces, including digital platforms, thereby adapting conventional provisions to technological developments. Furthermore, Articles 437 to 439 establish aggravating circumstances for defamation committed against public officials, state institutions, or through mass media.<sup>10</sup>

The new KUHP also introduces a significant provision in Articles 240 and 241, which reinstate criminal sanctions for insults against the President and Vice President, despite the Constitutional Court's 2006 ruling that invalidated similar provisions in the old KUHP. Legislators justify these provisions as necessary for protecting state institutions, yet critics contend that they threaten freedom of expression and risk reviving authoritarian legal practices. A similar concern arises with Article 263, which criminalizes insults against the government or state authorities and raises the possibility of misuse to suppress dissent and legitimate criticism. Normatively, the new KUHP demonstrates the legislature's effort to codify a comprehensive framework of dignity protection, extending its reach to individuals, state officials, and institutions. However, this regulatory approach has triggered debate in legal scholarship. Proponents argue that the provisions reflect the state's obligation to uphold human dignity and preserve public order, while opponents question their constitutional

---

<sup>8</sup> Eva Lievens, 'Bullying and Sexting in Social Networks: Protecting Minors from Criminal Acts or Empowering Minors to Cope with Risky Behaviour?', *International Journal of Law, Crime and Justice*, 42.3 (2014), 251–70 <https://doi.org/https://doi.org/10.1016/j.ijlcrj.2014.02.001>

<sup>9</sup> Matt DeLisi and Michael G Vaughn, 'Foundation for a Temperament-Based Theory of Antisocial Behavior and Criminal Justice System Involvement', *Journal of Criminal Justice*, 42.1 (2014), 10–25 <https://doi.org/https://doi.org/10.1016/j.jcrimjus.2013.11.001>

<sup>10</sup> Brent E Turvey and Aurelio Coronado, 'Chapter 2 - Victimity: Entering the Criminal Justice System', in *Forensic Victimology (Third Edition)*, ed. by Brent E Turvey, Third Edition (San Diego: Academic Press, 2023), pp. 33–76 <https://doi.org/https://doi.org/10.1016/B978-0-12-821768-9.00002-1>



compatibility with Article 28E of the 1945 Constitution, which guarantees the right to freedom of opinion and expression.<sup>11</sup>

The rapid expansion of digital communication further complicates the implementation of defamation provisions. The overlap between the new KUHP and the Electronic Information and Transactions Law (ITE Law) creates the potential for concurrent sanctions. For instance, defamatory acts committed through social media may simultaneously fall under Article 433 of the new KUHP and Article 27A of the ITE Law, thereby generating risks of double jeopardy and legal uncertainty. This situation underscores the urgency of harmonizing the new KUHP with sectoral legislation governing cyberspace.<sup>12</sup>

From the perspective of Pancasila, particularly the second principle, Just and Civilized Humanity, acts of defamation are not merely legal violations but also forms of conduct that fail to embody the values of humanity, justice, and civility. The practice of digital etiquette on social media represents a tangible manifestation of this principle, as users are expected to uphold tolerance, respect diversity, and refrain from disseminating misinformation or false accusations that may harm others. The circulation of hoaxes and slander not only undermines individual reputation but also generates broader societal consequences, including social fragmentation, the weakening of national unity, and threats to political stability. As the philosophical foundation of the state, the values of Pancasila establish a necessary balance between the right to freedom of expression and the responsibility to respect human dignity. Neglecting these values risks precipitating moral decline, ethical degradation, and disruption of social harmony.<sup>13</sup>

In practice, the utilization of social media frequently gives rise to negative behaviors, such as insults, defamation, blasphemy, provocation, and hate speech, all of which stand in fundamental contradiction to the ethical framework of Pancasila and the principles of a national life founded on humanity and justice. For this reason, the internalization of Pancasila values in digital interaction constitutes an urgent necessity to ensure that the advancement of technology does not erode the moral, legal, and cultural foundations of Indonesian society.<sup>14</sup>

Recent studies actively shape the discourse on digital defamation by examining regulatory frameworks, intermediary liability, and strategies to protect online reputation. Aloysius Wisnubroto analyzes the liability of internet intermediaries and demonstrates that inconsistent legal standards across jurisdictions create uncertainty

---

<sup>11</sup> Naganna Chetty and Sreejith Alathur, 'Hate Speech Review in the Context of Online Social Networks', *Aggression and Violent Behavior*, 40 (2018), 108–18 <https://doi.org/https://doi.org/10.1016/j.avb.2018.05.003>

<sup>12</sup> Kate Barber, 'Legal Cynicism in Men's Rights Discourses: Using Corpus Linguistics to Investigate How Distrust in the Legal System Excuses and Perpetuates Sexual Violence against Women', *Applied Corpus Linguistics*, 5.3 (2025), 100148 <https://doi.org/https://doi.org/10.1016/j.acorp.2025.100148>

<sup>13</sup> Fatima Zahrah, Jason R C Nurse and Michael Goldsmith, 'Unmasking Hate in the Pandemic: A Cross-Platform Study of the COVID-19 Infodemic', *Big Data Research*, 37 (2024), 100481 <https://doi.org/https://doi.org/10.1016/j.bdr.2024.100481>

<sup>14</sup> Claire Dillon, Vikramaditya Joshi and Amra Sabic-El-Rayess, 'Addressing Extremist Abuses of Medieval Past: A Connection-First Approach to Narratives of Hate', *International Journal of Educational Development*, 111 (2024), 103160 <https://doi.org/https://doi.org/10.1016/j.ijedudev.2024.103160>

for platforms while simultaneously restricting victims' ability to pursue effective redress. His research emphasizes that scholars and policymakers must design coherent legal approaches that balance innovation, platform autonomy, and the protection of individual reputation.<sup>15</sup> Sundari this discussion by investigating how intermediary responsibilities influence victims' access to justice. He shows that unclear procedural safeguards erode trust in the regulatory system and weaken efforts to protect reputation without unduly restricting freedom of expression.<sup>16</sup> Rustamadji contributes through a systematic review of online hate speech and identifies how blurred definitions between hate speech, harassment, and defamation produce inconsistencies in content moderation. He argues that such inconsistencies reduce victim protection and compromise the legitimacy of digital governance, making standardized moderation policies across platforms essential.<sup>17</sup> Collectively, these three strands of research demonstrate that scholars consistently emphasize the need for an integrated and transparent regulatory architecture. They show that combating digital defamation requires not only defining intermediary responsibilities but also ensuring procedural fairness, cross-jurisdictional harmonization, and a careful balance between freedom of expression and the protection of individual dignity. By articulating these challenges in detail, researchers actively guide policymakers, regulators, and technology companies toward building governance systems that restore trust in digital communication while safeguarding fundamental rights.

This research critically examines the phenomenon of digital defamation by analyzing the adequacy and effectiveness of existing regulatory frameworks, identifying the multidimensional challenges in their implementation, and exploring mechanisms to safeguard both individual and institutional reputations within the dynamics of rapid technological advancement. The study aims to integrate legal, technological, and sociocultural instruments in a balanced manner to address online defamatory practices while simultaneously upholding freedom of expression and privacy rights. By doing so, this research contributes to theoretical discourse and offers practical recommendations to refine public policies that ensure effective regulation, enforcement, and protection of reputational interests in the digital environment.

## METHOD

This study employed a qualitative approach with descriptive methods to facilitate an interpretive analysis of law as it operates in practice. The methodological foundation rests on sociological legal research, also referred to as socio-legal or sociological juridical research, which conceptualizes law not merely as a set of normative provisions but as an empirical phenomenon embedded in social reality. By adopting this approach, the study actively examines how legal norms function, interact, and encounter contestation within the broader dynamics of society.<sup>18</sup> The

---

<sup>15</sup> Wisnubroto and Hilaire Tegan.

<sup>16</sup> Elisabeth Sundari, Helidorus Chandera Halim and Ousu Mendy, 'Should Indonesia Adopt Legal Representation in Civil Cases?', *Journal of Human Rights, Culture and Legal System*, 5.2 (2025), 554–80 <https://doi.org/10.53955/jhcls.v5i2.604>

<sup>17</sup> Muhammad and others.

<sup>18</sup> Anis Mashdurohatun and others, 'Rethinking Palm Oil Plastic Regulations for Sustainable and Ecological Justice', *Journal of Human Rights, Culture and Legal System*, 5.2 (2025), 500–530 <https://doi.org/10.53955/jhcls.v5i2.681>

research design followed a descriptive-analytical orientation and utilized juridical methods based on systematic examination of secondary legal materials, such as statutory regulations, judicial decisions, and scholarly works. These sources were critically reviewed to establish a strong theoretical and doctrinal foundation for the analysis. In parallel, the study collected primary data through direct fieldwork. Primary data encompassed factual information, case documentation, and statements obtained from respondents closely connected to the research object, as well as observable practices relevant to the identified legal issues. Secondary data, meanwhile, provided conceptual grounding and served as an analytical framework to interpret and contextualize the empirical findings.

Data analysis relied on a qualitative technique aimed at producing descriptive-analytical insights. The researcher synthesized respondents' oral and written testimonies together with observed behaviors, subjecting them to a comprehensive and contextual examination. This process enabled the identification of recurring patterns, the recognition of internal contradictions, and the exploration of underlying meanings within the interaction between law and society. Through this approach, the study not only described legal practices as they unfold in reality but also critically interpreted their implications for the development, coherence, and effectiveness of Indonesia's broader legal system.

## RESULT AND DISCUSSION

### *Challenges and Strategies in Combatting Digital Defamation*

The elements constituting an act of defamation under the Electronic Information and Transactions Law are formulated in a more specific manner compared to the general provisions of the Criminal Code. He emphasizes that defamation must be distinguished from statements expressed in the form of opinions, since opinions are not categorized as criminal acts of defamation as long as they are not intended to deliberately undermine an individual's dignity or honor. This distinction reflects the principle that criminal liability in defamation cases requires the existence of specific intent directed at damaging another person's reputation. In the domain of civil law, the regulation of defamation in Indonesia is contained in the Civil Code, particularly Articles 1372 to 1380. These provisions conceptualize defamation as an unlawful act as stipulated in Article 1365 of the Civil Code.<sup>19</sup>

The primary purpose of civil claims in cases of insult or defamation is to compensate material and immaterial losses resulting from such acts, as well as to restore the honor and good name of the injured party, as stipulated in Article 1372. Furthermore, Article 1373 grants the defamed party the right to petition the court to declare that the act constitutes defamation, thereby strengthening their claim for damages. However, judges retain discretion to dismiss such claims if it is proven that the act did not contain defamatory intent and was instead carried out in the public interest or as an act of self-defense. This indicates that civil law recognizes a more

---

<sup>19</sup> Collins Udanor and Chinatu C Anyanwu, 'Combating the Challenges of Social Media Hate Speech in a Polarized Society', *Data Technologies and Applications*, 53.4 (2019), 501–27 <https://doi.org/https://doi.org/10.1108/DTA-01-2019-0007>

restorative and corrective function, rather than purely punitive objectives, in addressing defamation.<sup>20</sup>

From the perspective of criminal law, defamation is explicitly regulated in Article 310 of the Criminal Code and further elaborated in Article 27A of the Electronic Information and Transactions Law. Both provisions classify defamation as a complaint based offense. Consequently, prosecution of defamation is conditional upon a formal complaint lodged by the injured party. This procedural requirement underscores the private nature of the interest protected in defamation cases, distinguishing it from public offenses that can be pursued ex officio by the state. Moreover, Article 75 of the Criminal Code stipulates that a complaint can be withdrawn within three months of its submission, leading to the termination of the criminal process. In practice, reconciliation between the alleged offender and the injured party frequently serves as the basis for withdrawing such complaints, reflecting the socio legal reality that defamation disputes are often resolved through non litigation mechanisms rather than prolonged criminal proceedings.<sup>21</sup>

This dual regulation across both criminal and civil legal frameworks demonstrates the complex positioning of defamation within Indonesian law. On the one hand, the law seeks to protect individual dignity and reputation. On the other hand, it provides procedural safeguards to prevent the excessive criminalization of speech, particularly when such speech serves the public interest.<sup>22</sup> The overlapping yet distinct treatment in civil and criminal law illustrates the tension between retributive, restorative, and preventive functions of legal regulation in balancing the protection of personal honor with the constitutional guarantee of freedom of expression. The development of Indonesian criminal law demonstrates a gradual shift from a retributive to a restorative paradigm, as reflected in both statutory reforms and institutional practices. Restorative justice has increasingly been adopted as an alternative settlement mechanism for alleged criminal acts, with the primary objective of restoring harmony between the parties involved rather than imposing punitive sanctions. This approach emphasizes dialogue and reconciliation among the perpetrator, the victim, their families, and other relevant stakeholders, thereby aiming to repair the harm caused and to restore social equilibrium.<sup>23</sup>

Initially, the application of restorative justice was limited to minor offenses punishable by a maximum of three months' imprisonment or fines of limited amounts, as provided in various institutional instruments, including the Memorandum of Understanding between the Chief Justice of the Supreme Court, the Minister of Law

---

<sup>20</sup> Shivang Agarwal and C Ravindranath Chowdary, 'Combating Hate Speech Using an Adaptive Ensemble Learning Model with a Case Study on COVID-19', *Expert Systems with Applications*, 185 (2021), 115632 <https://doi.org/https://doi.org/10.1016/j.eswa.2021.115632>

<sup>21</sup> Sandip Modha and others, 'Detecting and Visualizing Hate Speech in Social Media: A Cyber Watchdog for Surveillance', *Expert Systems with Applications*, 161 (2020), 113725 <https://doi.org/https://doi.org/10.1016/j.eswa.2020.113725>

<sup>22</sup> Jack Levin and Jack McDevitt, 'Hate Crimes', in *Encyclopedia of Violence, Peace, & Conflict (Third Edition)*, ed. by Lester R Kurtz, Third Edition (Oxford: Academic Press, 2022), pp. 28–34 <https://doi.org/https://doi.org/10.1016/B978-0-12-820195-4.00040-6>

<sup>23</sup> Prashant Kapil and Asif Ekbal, 'A Deep Neural Network Based Multi-Task Learning Approach to Hate Speech Detection', *Knowledge-Based Systems*, 210 (2020), 106458 <https://doi.org/https://doi.org/10.1016/j.knosys.2020.106458>



and Human Rights, the Attorney General, and the Chief of the Indonesian National Police. This development was later reinforced through the Regulation of the Chief of Police Number 8 of 2021 concerning the Handling of Criminal Acts Based on Restorative Justice, which required law enforcement authorities to prioritize mediation and reconciliation in resolving specific criminal cases. The codification of restorative justice in the new Indonesian Criminal Code (Law No. 1 of 2023) marks a significant step forward in institutionalizing this principle.<sup>24</sup>

The new KUHP explicitly acknowledges restorative justice as a fundamental consideration in criminal law enforcement. Provisions such as Articles 51 and 52 introduce the possibility of settlement through restorative mechanisms, particularly for offenses categorized as minor crimes or those involving a maximum penalty of up to five years of imprisonment. These provisions affirm that the resolution of certain criminal cases should prioritize restoration and reconciliation, thereby positioning restorative justice not as a mere policy instrument but as a codified legal norm. This represents a paradigm shift in the philosophy of Indonesian criminal law, aligning it more closely with contemporary global movements that emphasize proportionality, human rights, and rehabilitation.<sup>25v</sup>

In the context of defamation, the new KUHP provides a comprehensive framework that both expands and restructures existing provisions. Articles 433 to 450 regulate insults and defamation, underscoring the dual objectives of protecting individual dignity and ensuring a balance with freedom of expression. For instance, Article 433 defines general defamation as attributing something to another person in a manner that could harm their honor or reputation, while Article 434 extends these provisions to written defamation. Article 435 criminalizes slanderous accusations, and Article 436 broadens the regulation of insults committed in public spaces, including digital platforms, reflecting the growing influence of technology in contemporary communication. Articles 437 to 439 introduce aggravating circumstances for defamation against public officials, state institutions, or through mass media, thereby enhancing the protection of state symbols and authorities.<sup>26</sup>

A controversial dimension of the new KUHP lies in the reintroduction of provisions on insults against the President and Vice President in Articles 218 to 220, as well as against the government and state institutions in Article 240 to 241. Although these provisions are justified by legislators as measures to protect state dignity, they have generated constitutional debates, as similar provisions in the old KUHP were previously annulled by the Constitutional Court in 2006.<sup>27</sup> Critics argue that such articles may restrict democratic freedoms and reintroduce authoritarian legal remnants,

---

<sup>24</sup> Zewdie Mossie and Jenq-Haur Wang, 'Vulnerable Community Identification Using Hate Speech Detection on Social Media', *Information Processing & Management*, 57.3 (2020), 102087 <https://doi.org/https://doi.org/10.1016/j.ipm.2019.102087>

<sup>25</sup> Aidan McKearney, 'Changing Contexts: From Criminal to Citizen', *Journal of Organizational Change Management*, 33.3 (2019), 515–26 <https://doi.org/https://doi.org/10.1108/JOCM-11-2018-0331>

<sup>26</sup> Ronald Miller and Brent E Turvey, 'Chapter 9 - Ethical Issues for Criminal Prosecutors', in *Ethical Justice*, ed. by Brent E Turvey and Stan Crowder (San Diego: Academic Press, 2013), pp. 257–316 <https://doi.org/https://doi.org/10.1016/B978-0-12-404597-2.00009-7>

<sup>27</sup> Stan Crowder, Shawn Mikulay and Brent E Turvey, 'Chapter 2 - Ethical Issues for Students of Criminal Justice', in *Ethical Justice*, ed. by Brent E Turvey and Stan Crowder (San Diego: Academic Press, 2013), pp. 21–62 <https://doi.org/https://doi.org/10.1016/B978-0-12-404597-2.00002-4>

potentially silencing dissent and criticism. From the perspective of restorative justice, the handling of defamation cases under the new KUHP is particularly significant.<sup>28</sup> Because defamation is classified as a complaint-based offense (*klacht delict*), prosecution depends on the submission of a complaint by the injured party. This procedural requirement aligns with the restorative justice framework, as it opens space for mediation, reconciliation, and withdrawal of complaints before formal prosecution. Articles 223 to 225 of the new KUHP reinforce this mechanism by allowing for the termination of prosecution if the parties reach a settlement, provided the victim's dignity and rights are restored. Such provisions underscore the legislator's intention to reduce the over-criminalization of speech-related offenses and to strengthen restorative approaches as a practical alternative to imprisonment.<sup>29</sup>

The convergence of restorative justice provisions and defamation regulations under the new KUHP reflects the state's effort to harmonize legal protection of individual dignity with constitutional guarantees of freedom of expression. Nevertheless, challenges remain in practice. The overlapping application of the KUHP and the Electronic Information and Transactions (ITE) Law may result in double jeopardy and legal uncertainty, especially in cases involving online defamation. For example, a defamatory statement posted on social media could simultaneously fall within the ambit of Article 433 of the new KUHP and Article 27A of the ITE Law. Without harmonization, this overlap risks undermining the restorative justice philosophy by subjecting individuals to multiple layers of criminal liability.<sup>30</sup>

The regulation of criminal acts of defamation via the internet in Indonesia remains problematic from both a philosophical and legal perspective. Ideally, citizens should be able to exercise their right to express opinions in the digital sphere without fear of surveillance, restriction, or criminal prosecution. However, Article 27A of the Electronic Information and Transactions Law (UU ITE) introduces provisions that criminalize the act of attacking the honor or good name of another person through electronic means. This article allows social media users to face imprisonment merely for expressing opinions online, thereby raising concerns about the state's inability to fully protect freedom of expression.<sup>31</sup>

Philosophically, the provision contained in Article 27A contradicts the spirit of reform (*semangat reformasi*), which emphasized democratization and the protection of human rights following the 1998 constitutional transition. Normatively, it conflicts with Article 28E paragraph (3) and Article 28F of the 1945 Constitution, which

---

<sup>28</sup> Daniyar Sultan and others, 'Cyberbullying-Related Hate Speech Detection Using Shallow-to-Deep Learning', *Computers, Materials and Continua*, 74.1 (2022), 2115–31 <https://doi.org/https://doi.org/10.32604/cmc.2023.032993>

<sup>29</sup> Amy M Donley, James D Wright and Jerian Benwell-Lybarger, 'Violence, Criminalization, and the Homeless', in *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*, ed. by James D Wright, Second Edition (Oxford: Elsevier, 2015), pp. 111–15 <https://doi.org/https://doi.org/10.1016/B978-0-08-097086-8.64101-6>

<sup>30</sup> Muhammad Okky Ibrohim and Indra Budi, 'Hate Speech and Abusive Language Detection in Indonesian Social Media: Progress and Challenges', *Heliyon*, 9.8 (2023), e18647 <https://doi.org/https://doi.org/10.1016/j.heliyon.2023.e18647>

<sup>31</sup> Jacqueline B Helfgott, 'Criminal Behavior and the Copycat Effect: Literature Review and Theoretical Framework for Empirical Investigation', *Aggression and Violent Behavior*, 22 (2015), 46–64 <https://doi.org/https://doi.org/10.1016/j.avb.2015.02.002>

explicitly guarantee the right of every individual to freedom of expression and the right to seek, obtain, and disseminate information through any available medium.<sup>32</sup> Moreover, Article 27A overlaps with several statutory provisions, including Law Number 14 of 2008 on Public Information Disclosure, Law Number 40 of 1999 on the Press, and Law Number 8 of 1999 on Consumer Protection. It also creates potential inconsistency with Articles 310 and 311 of the Indonesian Criminal Code (KUHP), which already regulate defamation in a conventional framework. The cumulative effect of these overlaps is the emergence of legal uncertainty regarding how to adjudicate online expressions considered insulting or defamatory.<sup>33</sup>

Within the Indonesian legal system, the crime of defamation has long been recognized under the KUHP, which functions as the general criminal law framework. Chapter XVI of Book II of the KUHP, specifically Articles 310, 311, 315, 317, and 318, defines defamation as an act of insulting or slandering another individual. These provisions primarily cover conventional acts of defamation carried out orally or in writing. By contrast, the UU ITE specifically regulates online defamation through Article 27A, which stipulates that any person who intentionally attacks the honor or good name of another person by making accusations accessible to the public through electronic information, electronic documents, or electronic systems, may be held criminally liable. Article 45 paragraph (4) further provides that such acts are punishable by up to two years of imprisonment and/or a fine of up to four hundred million rupiah.<sup>34</sup>

From a doctrinal standpoint, the coexistence of the KUHP and the UU ITE in regulating defamation has resulted in a dualistic legal framework, where acts of online defamation may fall simultaneously under both Article 310 of the KUHP and Article 27A of the UU ITE. This dualism produces the risk of over-criminalization and double jeopardy, as well as disproportionate restrictions on freedom of expression. Consequently, the enforcement of Article 27A not only undermines constitutional guarantees but also fails to reflect the principle of legal certainty, which constitutes a core element of the rule of law (*rechtsstaat*) in Indonesia. Referring to Article 27 paragraph (4) of the new Indonesian Criminal Code, the crime of threats constitutes a complaint-based offense (*delik aduan*) that can only be prosecuted if the victim files a formal complaint.<sup>35</sup> Consequently, law enforcement officials should not initiate prosecution in the absence of such a complaint. Misapplication, as reflected in the three judicial decisions under review, creates legal uncertainty and risks violating fundamental human rights. In particular, the rights of the accused may be compromised because no individual should be convicted without proven error or

---

<sup>32</sup> Anderson Almeida Firmino, Cláudio de Souza Baptista and Anselmo Cardoso de Paiva, 'Improving Hate Speech Detection Using Cross-Lingual Learning', *Expert Systems with Applications*, 235 (2024), 121115 <https://doi.org/https://doi.org/10.1016/j.eswa.2023.121115>

<sup>33</sup> Andrés Zapata Roza and others, 'Cyber Democracy in the Digital Age: Characterizing Hate Networks in the 2022 US Midterm Elections', *Information Fusion*, 110 (2024), 102459 <https://doi.org/https://doi.org/10.1016/j.inffus.2024.102459>

<sup>34</sup> Vahid Parvaresh, 'Covertly Communicated Hate Speech: A Corpus-Assisted Pragmatic Study', *Journal of Pragmatics*, 205 (2023), 63–77 <https://doi.org/https://doi.org/10.1016/j.pragma.2022.12.009>

<sup>35</sup> August F Y Chao and others, 'From Hate to Harmony: Leveraging Large Language Models for Safer Speech in Times of COVID-19 Crisis', *Heliyon*, 10.16 (2024), e35468 <https://doi.org/https://doi.org/10.1016/j.heliyon.2024.e35468>

fault within a lawful criminal process. The continued reliance on Article 45 paragraph (4) of the former Electronic Information and Transactions (EIT) Law in these cases requires comprehensive re-examination by law enforcement authorities to ensure alignment with the principle of legality and the formulation of sound positive law that provides clear guidance for sentencing.<sup>36</sup>

This study therefore investigates whether the formulation of Article 45 paragraph (4) in conjunction with Article 27 paragraph (4) of the former EIT Law meets the objectives of criminal law policy, particularly with regard to legal certainty, justice, and human rights protection. It further explores the implications of applying these provisions in judicial practice, especially in relation to safeguarding the constitutional rights of defendants. The substantive weakness lies in the lack of harmony between the old EIT Law and the new Criminal Code, which emphasizes the principle that complaint-based offenses must respect the victim's right to determine whether prosecution should proceed.<sup>37</sup>

To address this, institutional reform becomes essential. The establishment of a cybercrime task force not only involves the Indonesian National Police but also integrates Civil Servant Investigators (PPNS), prosecutors, and judges, whose coordination extends from central to regional levels. These law enforcement institutions must also collaborate with independent experts in information technology, ensuring that such expertise meets both academic rigor and professional accountability standards. Moreover, law enforcement requires reinforcement through public legal education. Citizens must be taught to exercise their right to express criticism in ways consistent with Pancasila values, while refraining from conduct that constitutes unlawful defamation. At the same time, law enforcement officials must strengthen their readiness to respond promptly, proportionately, and fairly to online defamation cases, ensuring compliance with the principles of justice enshrined in the new Criminal Code.<sup>38</sup>

A structural weakness in the Indonesian legal system lies in the insufficient synergy among law enforcement officials, which necessitates enhanced coordination among the police, prosecutors, and judiciary to ensure effective law enforcement and uphold legal certainty. The Criminal Code establishes distinct provisions for the protection of kings, heads of friendly states, and representatives of foreign countries in Indonesia, differentiating their status from that of ordinary citizens. This distinction manifests in separate articles that provide heightened legal protection for these figures. Specifically, individuals who insult a king, head of a friendly state, or a foreign representative face longer imprisonment terms compared to those who insult ordinary citizens. Furthermore, while defamation against ordinary citizens requires a formal complaint to initiate prosecution, offenses targeting kings or foreign dignitaries are classified as

---

<sup>36</sup> Rachel Griffin, 'New School Speech Regulation as a Regulatory Strategy against Hate Speech on Social Media: The Case of Germany's NetzDG', *Telecommunications Policy*, 46.9 (2022), 102411 <https://doi.org/https://doi.org/10.1016/j.telpol.2022.102411>

<sup>37</sup> Ehtesham Hashmi and others, 'Self-Supervised Hate Speech Detection in Norwegian Texts with Lexical and Semantic Augmentations', *Expert Systems with Applications*, 264 (2025), 125843 <https://doi.org/https://doi.org/10.1016/j.eswa.2024.125843>

<sup>38</sup> Giulia Lanza, 'The Fundamental Role of International (Criminal) Law in the War in Ukraine', *Orbis*, 66.3 (2022), 424–35 <https://doi.org/https://doi.org/10.1016/j.orbis.2022.05.010>



common offenses subject to immediate legal action. This distinction reflects Indonesia's diplomatic obligations and the special status afforded to heads of state and foreign representatives, ensuring protection of international relations and mutual respect between countries.<sup>39</sup>

The application of these legal norms, however, raises potential constitutional concerns, particularly regarding Articles 28 and 28E paragraphs (2) and (3) of the 1945 Constitution, which guarantee freedom of expression. For instance, a situation may arise in which an individual's criticism of the President is interpreted by investigators or public prosecutors as an insult, potentially infringing on constitutional rights. It is important to note that such violations relate to the implementation of the norm rather than its constitutionality. Errors in the interpretation or application of legal provisions may indeed infringe on rights, but these practical misapplications are distinct from the inherent unconstitutionality of the law itself.<sup>40</sup>

Historically, Article 134 of the former Criminal Code was declared unconstitutional by the Constitutional Court in Decision No. 013-022/PUU-IV/2006. The Court found that the provision restricted the fundamental right of every citizen to express opinions freely and created legal uncertainty (*rechtsonzekerheid*). The law was overly susceptible to subjective interpretation, as statements of protest or criticism could be mischaracterized as insults to the President or Vice President. Moreover, the provision's original intent in the WvS (Article 111) was to protect kings, and there was no indication that this protection could be extended to the President and Vice President. The annulment of Article 134 highlights the critical need to harmonize legal provisions with constitutional guarantees, ensuring that protection of high-ranking officials does not unjustly curtail citizens' rights to free expression.<sup>41</sup>

### ***Legal Framework for Combating Digital Defamation in Germany and Inggris***

Germany has established stringent standards to combat hate content on social media platforms through the *Netzwerkdurchsetzungsgesetz* (NetzDG), which was enacted in 2017 and became effective in January 2018. This legislation obliges social media platforms with more than two million users to remove clearly illegal content within 24 hours and other illegal content within seven days after notification, imposing fines of up to €50 million for non-compliance. Over time, the regulation has undergone several amendments to strengthen law enforcement and enhance user rights. A notable amendment was introduced in 2021 through the *Gesetz zur Bekämpfung des Rechtsextremismus und der Hasskriminalität* (Law on Combating Right-Wing Extremism and Hate Crime), which came into effect on April 3, 2021. This amendment requires social media providers to proactively report potentially criminal

---

<sup>39</sup> Giacomo Giorgini Pignatiello, 'Countering Anti-Lgbti+ Bias in the European Union. A Comparative Analysis of Criminal Policies and Constitutional Issues in Italian, Spanish and French Legislation', *Women's Studies International Forum*, 86 (2021), 102466 <https://doi.org/https://doi.org/10.1016/j.wsif.2021.102466>

<sup>40</sup> Yasmine M Ibrahim, Reem Essameldin and Saad M Darwish, 'An Adaptive Hate Speech Detection Approach Using Neutrosophic Neural Networks for Social Media Forensics', *Computers, Materials and Continua*, 79.1 (2024), 243–62 <https://doi.org/https://doi.org/10.32604/cmc.2024.047840>

<sup>41</sup> Aldo Zammit Borda, 'International Criminal Courts and Tribunals', in *Encyclopedia of Violence, Peace, & Conflict (Third Edition)*, ed. by Lester R Kurtz, Third Edition (Oxford: Academic Press, 2022), pp. 627–33 <https://doi.org/https://doi.org/10.1016/B978-0-12-820195-4.00155-2>

content to the Bundeskriminalamt (BKA) and reinforces user rights, including the ability to file complaints against content removal decisions and to access explanations regarding the reasons for removal.<sup>42</sup>

The implementation of the NetzDG has produced notable outcomes in mitigating the dissemination of hate content on social media platforms. Nevertheless, enforcement challenges persist. German government surveys indicate that social media companies' compliance with the regulation remains relatively low; for instance, Facebook removed approximately 39 percent of reported illegal content, while Twitter removed only 1 percent. These figures suggest that, despite the signing of a voluntary code of conduct in 2015 requiring the removal of hate content within 24 hours, the regulation's effectiveness still requires substantial improvement. Comparative analysis with other countries highlights that Germany adopts a more stringent approach to regulating hate content online. In contrast, the United States lacks federal legislation that explicitly obliges social media platforms to remove hate speech, leaving companies largely self-regulated. In Germany, however, social media providers are legally accountable for the content published on their platforms and face significant sanctions if they fail to comply with existing regulations, demonstrating a proactive legal framework aimed at curbing online hate and protecting public discourse.<sup>43</sup>

Nevertheless, this stringent approach has attracted criticism from various stakeholders. Human rights organizations and academics have highlighted the risk of overblocking, whereby lawful content is removed preemptively to avoid potential fines. This issue creates a persistent dilemma between protecting the public from hate speech and preserving freedom of expression in digital spaces. In terms of law enforcement, Germany demonstrates a strong commitment to combating digital hate crimes. For instance, on 25 June 2025, the German government conducted a large-scale enforcement operation involving raids at 65 locations and the arrest of 140 suspects accused of disseminating hate content online. These measures indicate that violations of the NetzDG carry not only administrative penalties, such as fines, but may also result in serious criminal sanctions.<sup>44</sup>

The NetzDG represents one of Germany's most important legal instruments for addressing the spread of hate content and disinformation on digital platforms. The 2021 amendment expanded the regulation's scope by introducing proactive reporting obligations and reinforcing user rights, while the imposition of fines up to €50 million underscores the government's seriousness in enforcing compliance. Nonetheless, significant challenges remain in implementing the law, including low compliance rates among social media companies and the risk of excessive content removal. Effective enforcement that is both rigorous and proportional is essential to ensure that digital spaces remain safe without undermining freedom of expression. By maintaining this

---

<sup>42</sup> Griffin.

<sup>43</sup> Sabrina Maaß, Jil Wortelker and Armin Rott, 'Evaluating the Regulation of Social Media: An Empirical Study of the German NetzDG and Facebook', *Telecommunications Policy*, 48.5 (2024), 102719 <https://doi.org/https://doi.org/10.1016/j.telpol.2024.102719>

<sup>44</sup> Ben Wagner and others, 'Mapping Interpretations of the Law in Online Content Moderation in Germany', *Computer Law & Security Review*, 55 (2024), 106054 <https://doi.org/https://doi.org/10.1016/j.clsr.2024.106054>

balance, Germany sustains a high regulatory standard that simultaneously upholds legal protection and democratic values.<sup>45</sup>

Germany adopts a rigorous approach to regulating defamation and hate content on social media through the NetzDG, enacted in 2017 and implemented in 2018, with subsequent amendments in 2021 under the Gesetz zur Bekämpfung des Rechtsextremismus und der Hasskriminalität. This legislation obliges social media companies with more than two million users to remove illegal content within 24 hours and resolve other reported content within seven days, imposing fines of up to €50 million on companies and €5 million on their local directors. Additionally, courts possess the authority to order the disclosure of user identities responsible for spreading hate content. Germany's emphasis on platform accountability and digital content oversight reflects the state's commitment to preventing disinformation and hate speech that could influence public opinion, particularly during election periods. Government surveys indicate that compliance remains limited; for example, Facebook removed only 39 percent of reported illegal content, while Twitter removed merely 1 percent, despite both platforms having signed a voluntary code of conduct in 2015. This approach underscores the necessity of clear and consistent enforcement mechanisms while maintaining a balance between user privacy and freedom of expression.<sup>46</sup>

In the United Kingdom, defamation law specifically regulates the protection of individual and organizational reputations while maintaining a balance with freedom of expression. A significant reform in this legal framework was the enactment of the Defamation Act 2013, which came into force on 1 January 2014. The legislation aims to address the imbalance between reputational protection and freedom of speech and to reduce the number of unmeritorious defamation claims. Prior to the Defamation Act 2013, defamation law was governed by the Defamation Act 1996, which provided the legal basis for individuals or entities to initiate claims if their reputations were harmed by false statements. However, the 1996 Act was increasingly considered insufficient to address contemporary challenges, particularly in the context of social media and online publications. The Defamation Act 2013 introduced substantial changes, most notably the "serious harm" threshold under Section 1, which requires that a statement must have caused or be likely to cause serious harm to the claimant's reputation in order to be actionable. This threshold aims to prevent frivolous lawsuits and ensures that only claims with demonstrable reputational damage proceed to court.<sup>47</sup>

---

<sup>45</sup> Purbani Kar and Swapan Debbarma, 'Sentimental Analysis & Hate Speech Detection on English and German Text Collected from Social Media Platforms Using Optimal Feature Extraction and Hybrid Diagonal Gated Recurrent Neural Network', *Engineering Applications of Artificial Intelligence*, 126 (2023), 107143 <https://doi.org/https://doi.org/10.1016/j.engappai.2023.107143>

<sup>46</sup> Robert Gorwa, 'Elections, Institutions, and the Regulatory Politics of Platform Governance: The Case of the German NetzDG', *Telecommunications Policy*, 45.6 (2021), 102145 <https://doi.org/https://doi.org/10.1016/j.telpol.2021.102145>

<sup>47</sup> Amira Ghenai and others, 'Exploring Hate Speech Dynamics: The Emotional, Linguistic, and Thematic Impact on Social Media Users', *Information Processing & Management*, 62.3 (2025), 104079 <https://doi.org/https://doi.org/10.1016/j.ipm.2025.104079>

The Act also eliminated the presumption that defamation cases must be tried by jury, granting courts the discretion to determine whether a jury trial is appropriate. This change increases procedural flexibility, accelerates legal proceedings, and reduces litigation costs. Additionally, the Act introduces several new defenses for defendants, including truth, honest opinion, and publication on a matter of public interest. These defenses protect individuals who make statements that are factually accurate, express genuine opinion, or address issues of public concern. By establishing these provisions, the Defamation Act 2013 reflects the United Kingdom's effort to create a defamation framework that balances the protection of reputation with the necessity of free expression, particularly in an era dominated by digital communication and social media platforms. It underscores the importance of procedural efficiency, legal certainty, and the safeguarding of public discourse while preventing abuse of the legal system through unsubstantiated claims.<sup>48</sup>

The Defamation Act 2013 in the United Kingdom provides particular attention to online publications. Under Section 5, website operators may be exempt from liability if they comply with specific procedures, such as allowing claimants to contact the authors of allegedly defamatory material and offering the opportunity to remove or correct such content. This provision addresses challenges arising from the rapid and widespread dissemination of information on the internet. Since the Act's implementation, significant changes have occurred in the number of defamation cases filed in court. Judicial statistics indicate that claims filed in the High Court of England and Wales decreased from 323 cases in 2019 to 152 cases in 2020, demonstrating that the legislation effectively reduced unmeritorious claims and encouraged out-of-court dispute resolution.<sup>49</sup>

Despite this overall decline, cases involving social media have increased. In 2015, 13 defamation cases filed in court were related to social media comments, up from 11 cases the previous year. This trend illustrates that, although the new law has been enacted, challenges associated with online defamation persist and require ongoing attention. The Defamation Act 2013 introduced several critical reforms: it established the "serious harm" threshold, revised procedural rules, and incorporated new defenses, including truth, honest opinion, and publication on matters of public interest. These reforms aim to balance reputational protection with freedom of expression while ensuring procedural efficiency and legal certainty.<sup>50</sup>

Defamation in the UK is primarily a civil matter and has been decriminalized. The Act differentiates between libel (written or permanent defamation) and slander (spoken or temporary defamation), encompassing print media, broadcasts, film, online content, and public statements during theatrical performances. The legislation also introduces restorative mechanisms, allowing defendants to issue written apologies

---

<sup>48</sup> Eva Nave and Lottie Lane, 'Countering Online Hate Speech: How Does Human Rights Due Diligence Impact Terms of Service?', *Computer Law & Security Review*, 51 (2023), 105884 <https://doi.org/https://doi.org/10.1016/j.clsr.2023.105884>

<sup>49</sup> Atiyeh Sadeghi, Sebastian Pape and David Harborth, 'The Impact of Individuals' Social Environments on Contact Tracing App Use: Survey Study', *JMIR Human Factors*, 10 (2023) <https://doi.org/https://doi.org/10.2196/45825>

<sup>50</sup> Julia Jorgensen, 'The Functions of Sarcastic Irony in Speech', *Journal of Pragmatics*, 26.5 (1996), 613–34 [https://doi.org/https://doi.org/10.1016/0378-2166\(95\)00067-4](https://doi.org/https://doi.org/10.1016/0378-2166(95)00067-4)



and provide agreed-upon compensation to resolve disputes without protracted litigation. This approach encourages amicable resolution, reduces unnecessary court proceedings, and strengthens legal certainty. Judicial statistics reflect a notable reduction in traditional defamation cases since the enactment of the Act, while cases associated with digital platforms continue to rise. These patterns underscore the need for legal adaptation in response to technological and social media developments. The Defamation Act 2013 thus represents a significant milestone in the reform of defamation law in the UK, aiming to protect reputations, maintain freedom of expression, and adapt to the evolving digital landscape.<sup>51</sup>

Indonesia regulates defamation through Articles 310–321 of the Criminal Code (KUHP) and Article 27(3) of the Electronic Information and Transactions Law (UU ITE), combining criminal and civil sanctions. Penalties include fines and imprisonment, applicable to both traditional and digital defamation cases. However, law enforcement faces significant challenges, including ambiguous interpretations of the UU ITE, disharmony between law enforcement authorities and the judiciary, and the absence of restorative mechanisms that facilitate amicable dispute resolution. Consequently, litigation processes are often prolonged, susceptible to abuse, and generate legal uncertainty for the public. Additional challenges arise from the rapid and extensive dissemination of online content, necessitating a more structured and platform-specific approach to enforcement, similar to practices in Germany.

Comparative analysis with Germany and the United Kingdom highlights several lessons for Indonesian legal reform. First, Indonesia could impose obligations on digital platforms to remove illegal content within a specified timeframe and enforce administrative penalties or fines against companies and their management for non-compliance, following the German model. Second, restorative mechanisms, such as written apologies and agreed-upon compensation adopted in the UK, could be incorporated as alternative dispute resolution methods, reducing court burdens and preventing protracted conflicts. Third, establishing a “serious harm” threshold would ensure that cases lacking substantial reputational damage are not brought to court, thereby preventing misinterpretation and overcriminalization. Fourth, regulations should explicitly address digital media, incorporating transparent procedures and content-handling deadlines while safeguarding freedom of expression. Finally, enhancing public legal and digital literacy is essential to ensure that citizens understand their rights and responsibilities regarding content dissemination, thereby supporting effective and equitable law enforcement.

By integrating these approaches, Indonesian legal reform can achieve a balance between rigorous enforcement, protection of victims’ rights, and the preservation of freedom of expression. Such a framework would modernize Indonesia’s defamation law, rendering it more responsive to contemporary digital challenges while promoting legal certainty and societal trust in the judicial process. Implementing structured, platform-specific obligations, restorative practices, and clear thresholds for actionable harm can transform the defamation regime into a system that is both protective and

---

<sup>51</sup> Kerry Bray, Nils Braakmann and John Wildman, ‘Austerity, Welfare Cuts and Hate Crime: Evidence from the UK’s Age of Austerity’, *Journal of Urban Economics*, 141 (2024), 103439 <https://doi.org/https://doi.org/10.1016/j.jue.2022.103439>

adaptable, aligning national law with international best practices in digital governance and human rights protection.

### ***Legal Framework for Combating Digital Defamation***

Indonesia regulates online defamation through Law No. 11 of 2008 concerning Electronic Information and Transactions (ITE Law), as amended by Law No. 19 of 2016, specifically Article 27, Paragraphs (1) and (3). Article 27(1) prohibits any individual from intentionally and without authorization disseminating, transmitting, or making available electronic data containing unethical content. Article 27(3) further specifies that individuals are prohibited from intentionally and without authority sending, transmitting, or producing electronic information containing insulting or defamatory material. These provisions integrate criminal and civil dimensions to address both traditional and digital defamation, reflecting Indonesia's efforts to regulate online behavior and protect public and private reputations.<sup>52</sup>

Although these regulations are derived from legal rules and principles that predate Indonesia's Proclamation of Independence and remain applicable under the Transitional Provisions, Article II of the 1945 Constitution, their implementation has been adapted to post-independence circumstances. Scholars such as Notonegoro argue that such adaptations involve interpreting the values embedded in colonial-era law rather than unconditionally preserving them. Article II of the Transitional Provisions stipulates that pre-existing regulations continue to apply only insofar as no new provisions have been enacted under the 1945 Constitution. Consequently, the validity of colonial legal provisions is conditional, and they remain operative solely until replaced or reformed under national law, without a fixed expiration period.<sup>53</sup>

Despite the historical influence of colonial law, Indonesian legal reforms demonstrate a deliberate effort to develop national law grounded in the principles, philosophy, and norms of Indonesian society. This reformist approach has drawn inspiration from European legal systems, particularly French and Dutch 19th-century jurisprudence, but incorporates the principle of balance to establish a distinctly national legal foundation. Such legal modernization is crucial in the context of globalization and information technology, particularly for regulating digital communication, business, and international trade, while ensuring fairness and consistency in law enforcement.<sup>54</sup>

A key example of Indonesia's commitment to national legal development is the enactment of Law No. 8 of 1981 concerning Criminal Procedure (KUHAP). The considerations outlined in the law explicitly emphasize the reform of colonial legal structures into national legal frameworks, aiming to unify and codify laws consistent with the Indonesian archipelagic perspective. The law also underscores the

---

<sup>52</sup> Ian D Marder, 'Mapping Restorative Justice and Restorative Practices in Criminal Justice in the Republic of Ireland', *International Journal of Law, Crime and Justice*, 70 (2022), 100544 <https://doi.org/https://doi.org/10.1016/j.ijlcrj.2022.100544>

<sup>53</sup> Atul Kumar Srivastava and others, 'Leveraging Deep Learning for Comprehensive Multilingual Hate Speech Detection', *Procedia Computer Science*, 252 (2025), 832–40 <https://doi.org/https://doi.org/10.1016/j.procs.2025.01.044>

<sup>54</sup> Ishmael Mugari and Caleb Muzinda, 'Social Media: The Double-Edged Sword for Perpetuating and Preventing Criminal Activities in Zimbabwe', *Safer Communities*, 20.4 (2021), 225–36 <https://doi.org/https://doi.org/10.1108/SC-02-2021-0004>

internalization of societal obligations, the proper development of law enforcement attitudes, and the promotion of justice, legal certainty, and protection of human dignity. By codifying procedural rules that align with Pancasila and the 1945 Constitution, KUHAP ensures that all citizens enjoy equality before the law while obliging compliance with legal norms and governmental authority.<sup>55</sup>

Indonesia's approach to defamation, particularly through the ITE Law and national criminal procedural law, reflects a systematic effort to balance historical legal continuity with the demands of modern governance. These measures aim to regulate electronic communications, protect reputations, and strengthen legal certainty, while fostering the development of national legal identity and adherence to the principles of justice, equality, and human rights enshrined in the 1945 Constitution. The ongoing adaptation and reform of legal frameworks demonstrate Indonesia's commitment to establishing a law-based state capable of addressing both traditional and contemporary challenges in a rapidly evolving digital and social context.<sup>56</sup>

The theory of justice values emphasizes dignity, as it views the Indonesian positive legal system as a tolerant framework that incorporates the five major legal traditions developed by humanity over time. Ilham Basri asserts that the Indonesian legal system is a highly complex and comprehensive system, in which these five global legal traditions function as interrelated and mutually influential elements. According to Basri, no branch or subsystem of Indonesian law can be examined in isolation, as all elements operate in synergy, similar to the organs of a human body whose functionality depends on their interconnectedness.<sup>57</sup>

Moreover, the theory of justice values posits that a positive legal system is inherently goal-oriented. Within such a system, the whole is greater than the sum of its parts, and each subsystem operates in relation to the broader environment. The interaction of system components generates value, while harmonization and unifying mechanisms maintain coherence. Furthermore, internal mechanisms of control, correction, supervision, and feedback function to sustain the system's continuity. The Indonesian positive legal system, from this perspective, is a living, functioning system that regulates societal order on a daily basis. It is a cohesive entity whose components interact to achieve the overarching objectives of law and governance, ensuring that each legal issue or problem finds resolution within the system itself.<sup>58</sup>

---

<sup>55</sup> Louis B Schlesinger and Tirza Z Ben-Ari, 'Personality Disorders and Criminal Behavior', in *Encyclopedia of Forensic Sciences, Third Edition (Third Edition)*, ed. by Max M Houck, Third Edition (Oxford: Elsevier, 2023), pp. 133–43 <https://doi.org/10.1016/B978-0-12-823677-2.00057-X>

<sup>56</sup> Roberto Labadie-Tamayo and others, 'Distilling Knowledge from Large Language Models: A Concept Bottleneck Model for Hate and Counter Speech Recognition', *Information Processing & Management*, 63.2, Part A (2026), 104309 <https://doi.org/10.1016/j.ipm.2025.104309>

<sup>57</sup> Aaron Sell and Daniel Sznycer, 'Societal Institutions Echo Evolved Human Nature: An Analysis of the Western Criminal Justice System and Its Relation to Anger', *Evolution and Human Behavior*, 44.3 (2023), 210–21 <https://doi.org/10.1016/j.evolhumbehav.2023.01.007>

<sup>58</sup> Mahima Choudhary, Basant Agarwal and Vishnu Goyal, 'Hate Speech Detection: Leveraging LLM-GPT2 with Fine-Tuning and Multi-Shot Techniques', *Procedia Computer Science*, 258 (2025), 2817–25 <https://doi.org/10.1016/j.procs.2025.04.542>

The Indonesian legal system, based on Pancasila as a legal philosophy, also embodies this intrinsic unity. Positive law in Indonesia functions as an integrated whole, within which every legal problem must have a solution grounded in the system. This principle renders the legal system dignified and reflective of societal values. Within this framework, legal reconstruction emphasizes not only regulatory compliance but also the integration of justice values, ensuring that the system maintains both effectiveness and legitimacy in practice.<sup>59</sup>

The reconstruction of normative values in the context of this study suggests that Indonesia's regulation of criminal defamation conducted via the internet, which previously lacked a justice-oriented foundation, must now align with the principles of justice. By grounding legal provisions in justice values, the legal framework seeks to balance the protection of reputations, individual dignity, and freedom of expression, while ensuring that legal enforcement is fair, consistent, and aligned with the philosophical foundations of the Indonesian legal system. This approach reflects a broader effort to modernize Indonesian law, incorporating both national values and the evolving demands of digital society.<sup>60v</sup>

The Restorative Justice (RJ) approach functions not merely as a procedural mechanism to amicably terminate cases but fundamentally emphasizes the fulfillment of justice for all parties involved. Within the context of defamation offenses in the electronic information and transactions sector, the implementation of RJ requires adherence to several essential principles. First, the mediator must effectively communicate the significance of mediation, manage the emotional responses of the parties, and clarify that the incident constitutes a misstep that necessitates rectification. Through this mediation process, any harm incurred can be addressed and restored in a manner that is equitable to all parties. Second, the mediator is responsible for identifying specific steps to compensate for damages or restore the victim's circumstances, including reinstating the victim's dignity and reputation through clarification or monetary and non-monetary compensation, as mutually agreed upon by the perpetrator and the victim. Third, the restorative process must actively involve all stakeholders to reconstruct and harmonize the social relationship between the victim and the perpetrator, ensuring that previously disrupted interactions return to a constructive equilibrium. Fourth, the process quality must be prioritized, as mediation does not serve to establish winners or losers but rather to cultivate mutual respect, awareness, and the achievement of a mutually beneficial resolution.<sup>61</sup>

Law No. 1 of 2024 explicitly governs defamation through Article 45, Paragraph (4), stipulating that any individual who intentionally assaults the honor or reputation of another by making allegations intended for public dissemination via electronic

---

<sup>59</sup> Zhilong Guo, Jie Hao and Lewis Kennedy, 'Protection Path of Personal Data and Privacy in China: Moving from Monism to Dualism in Civil Law and Then in Criminal Law', *Computer Law & Security Review*, 52 (2024), 105928 <https://doi.org/https://doi.org/10.1016/j.clsr.2023.105928>

<sup>60</sup> Krisda Chaemsathong and Kewalin Simuang, 'Lawyers' Metapragmatic Expressions in the Opening Speech of Anglo-American Criminal Trials', *Lingua*, 289 (2023), 103536 <https://doi.org/https://doi.org/10.1016/j.lingua.2023.103536>

<sup>61</sup> Randi Solhjell, 'How Acts Become Hate Crime: The Police's Documenting of Criminal Cases', *International Journal of Law, Crime and Justice*, 72 (2023), 100574 <https://doi.org/https://doi.org/10.1016/j.ijlcj.2022.100574>



information and/or electronic documents through an electronic system, as defined in Article 27A, may face imprisonment of up to two years and/or a fine of up to IDR 400,000,000.00 (four hundred million rupiah). Paragraph (5) further specifies that this offense is complaint-based, allowing prosecution only upon the victim's complaint, thereby emphasizing victim autonomy in the legal process and reducing the potential for institutional overreach in prosecuting defamation in the digital sphere. This framework ensures that the victim retains agency while providing legal clarity and accountability in electronic defamation cases.<sup>62</sup>

In addition, Law No. 1 of 2023, known as the New Criminal Code (*KUHP Baru*), reinforces the prohibition against acts that attack personal honor or reputation for public dissemination under Article 433. This regulation is set to take effect in 2026 and establishes a modernized, coherent legal foundation for safeguarding individual reputations, particularly in digital environments. By integrating restorative justice mechanisms into both the Electronic Information and Transactions Law and the New Criminal Code, the legislation provides a robust legal basis for initiating mediation and restitution processes, addressing challenges previously encountered under Police Regulation No. 8 of 2021.<sup>63</sup>

Under Perpol No. 8 of 2021 on Handling Criminal Acts Based on Restorative Justice, Article 15 mandates that restorative justice proceedings be initiated through a written application submitted by the perpetrator, the victim, or other relevant parties, accompanied by a peace declaration and evidence of restitution. Following this submission, investigators conduct a special case review to determine the appropriateness of terminating the case for legal reasons, as outlined in Article 16. However, this regulatory structure positions investigators in a predominantly passive role, precluding proactive initiation of restorative justice. Such limitations introduce operational challenges, as any proactive effort by investigators may be perceived as biased and subject to internal oversight complaints, thereby constraining law enforcement officers' ability to apply restorative justice effectively in digital defamation cases, where timely resolution and reputation rehabilitation are essential.<sup>64</sup>

Consequently, the incorporation of restorative justice principles into defamation regulations, both within the UU ITE and the New Criminal Code, underscores the necessity of a humanistic, participatory, and victim-centered legal approach. Restorative justice prioritizes not only the imposition of sanctions on perpetrators but also the restoration of victims' reputations, the conduct of high-quality mediation, and the reconciliation of social relationships. This approach establishes a legal framework that is equitable, transparent, and effective, ensuring comprehensive

---

<sup>62</sup> Evan N Shenkin and Michele Abee, 'International Spaces for Feminist Cross-Border Resistance', *Political Geography*, 112 (2024), 103111 <https://doi.org/https://doi.org/10.1016/j.polgeo.2024.103111>

<sup>63</sup> Tay Keong Tan, 'Chapter 17 - Curse of the Cyborg Mammoths: The Use of Artificial Intelligence in Manipulating and Mobilizing Human Emotions', in *Ethics in Online AI-Based Systems*, ed. by Santi Caballé, Joan Casas-Roma, and Jordi Conesa, Intelligent Data-Centric Systems (Academic Press, 2024), pp. 347–64 <https://doi.org/https://doi.org/10.1016/B978-0-443-18851-0.00018-4>

<sup>64</sup> Craig Wright, 'Chapter 21 - Information Systems Legislation', in *The IT Regulatory and Standards Compliance Handbook*, ed. by Craig Wright (Burlington: Syngress, 2008), pp. 609–71 <https://doi.org/https://doi.org/10.1016/B978-1-59749-266-9.00021-7>

protection for society in the digital era. Effective implementation of these principles requires explicit regulatory support, professional mediator training, and heightened legal awareness among all stakeholders, enabling restorative justice to operate not merely as an alternative dispute resolution mechanism but as a primary strategy for upholding justice in cyberspace.<sup>65</sup>

Law No. 1 of 2024, which constitutes the second amendment to Law No. 11 of 2008 on Electronic Information and Transactions, exhibits several substantive normative weaknesses. Article 27A employs the terms “suatu hal” (“a matter”) and “orang lain” (“another person”), which are overly subjective and prone to multiple interpretations in application. This ambiguity undermines legal certainty, particularly in the context of law enforcement officers determining the boundaries of actions considered attacks on individual honor or reputation. This concern aligns with Gustav Radbruch’s theory, which identifies legal certainty (*Rechtssicherheit*) as a fundamental element of law alongside justice and utility; ambiguous terminology risks generating injustice through divergent interpretations.<sup>66</sup>

To address this issue, Article 27A should be reformulated with more precise and operational language. Specifically, “orang lain” should be replaced with “individual,” and “suatu hal” should be rephrased as “an act that undermines the honor or reputation of an individual.” Consequently, the revised provision would read: “Every person who intentionally accuses or undermines the honor or reputation of an individual, with the intent that such act be publicly disseminated through Electronic Information and/or Electronic Documents conducted via an Electronic System.” This formulation enhances legal certainty and supports human rights protection, as guaranteed in Article 28G paragraph (1) of the 1945 Constitution, which ensures personal security and protection from threats to human dignity and honor.<sup>67</sup>

Further weaknesses exist in Article 45, paragraph (5), which governs complaint-based criminal acts. Although it restricts prosecution to the victim or affected party, the provision does not explicitly incorporate a restorative justice framework. Contemporary legal paradigms emphasize the resolution of disputes through restorative mechanisms to achieve equitable justice for victims, perpetrators, and society. As articulated by Howard Zehr, restorative justice prioritizes the restoration of social relationships over punitive measures. Therefore, the normative reconstruction should explicitly integrate restorative justice principles, consistent with Satjipto Rahardjo’s concept of progressive law, which treats law as a living instrument designed to realize substantive justice in society.<sup>68</sup>

---

<sup>65</sup> David Kloos, ‘A Crazy State’, *Bijdragen Tot de Taal-, Land- En Volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia*, 170.1 (2014), 25–65 <https://doi.org/https://doi.org/10.1163/22134379-17001003>

<sup>66</sup> Zora A Sukabdi, ‘Bridging the Gap: Contributions of Academics and National Security Practitioners to Counterterrorism in Indonesia’, *International Journal of Law, Crime and Justice*, 65 (2021), 100467 <https://doi.org/https://doi.org/10.1016/j.ijlcj.2021.100467>

<sup>67</sup> Endang Wahyu Pamungkas, Valerio Basile and Viviana Patti, ‘A Joint Learning Approach with Knowledge Injection for Zero-Shot Cross-Lingual Hate Speech Detection’, *Information Processing & Management*, 58.4 (2021), 102544 <https://doi.org/https://doi.org/10.1016/j.ipm.2021.102544>

<sup>68</sup> Ahmad Ghandour, Viktor Shestak and Konstantin Sokolovskiy, ‘Security and Privacy of Adolescents in Social Applications and Networks: Legislative Aspects and Legal Practice of Countering Cyberbullying

Integrating restorative justice into the legal framework offers multiple benefits. First, it provides a structured mechanism for reconciling victims and perpetrators, facilitating the restoration of the victim's dignity and reputation. Second, it emphasizes mediation and compensation as primary tools, reducing reliance on adversarial litigation and mitigating protracted judicial processes. Third, it encourages community participation and accountability, ensuring that justice is achieved collectively rather than solely through punitive enforcement. By incorporating these principles into Articles 27A and 45, the law evolves into a human-centered instrument capable of addressing challenges associated with digital defamation effectively.<sup>69</sup>

The proposed normative reconstruction strengthens the coherence and functionality of Law No. 1 of 2024 within Indonesia's legal system. Clarifying terminology, embedding restorative justice, and aligning the law with human rights principles create a legal framework that is fair, predictable, and enforceable. This approach enhances legal certainty while enabling the law to function as a mechanism for substantive justice, responsive to societal needs in the digital era. Consequently, the reformed provisions embody the principle that law must operate as a dynamic, living system, capable of safeguarding dignity, rights, and social harmony in an increasingly digital society.<sup>70</sup>

## CONCLUSION

Based on the discussion, it can be concluded that, first, the regulation of internet-based defamation in Indonesia under Law Number 1 of 2024, which amends Law Number 11 of 2008 concerning Electronic Information and Transactions, remains substantively and structurally inadequate, resulting in legal uncertainty and limiting the public's ability to exercise freedom of expression. The use of ambiguous terminology in Article 27A, including phrases such as "another person" and "a matter," allows for multiple interpretations, creating challenges for law enforcement, prosecutors, and the judiciary in determining the scope of criminal liability. This vagueness undermines legal certainty, a principle emphasized by Gustav Radbruch, and may lead to overcriminalization and arbitrary application, which is inconsistent with the guarantees provided under Articles 28E paragraph three and 28F of the 1945 Constitution, as well as related statutes including the Law on Public Information Transparency, the Press Law, the Consumer Protection Law, and Articles 310 and 311 of the Criminal Code. Second, the current legal framework insufficiently incorporates restorative justice mechanisms, which are essential for achieving equitable outcomes and restoring social harmony. While Article 45 paragraph five limits prosecution to complaint-based cases, it does not explicitly prioritize restorative approaches, reducing opportunities for mediation, victim compensation, and the repair of social relationships between perpetrators and victims. Restorative justice emphasizes

---

on Example of Developed and Developing Countries', *Journal of Information, Communication and Ethics in Society*, 19.4 (2021), 433–45 <https://doi.org/https://doi.org/10.1108/JICES-09-2020-0101>

<sup>69</sup> Jemma Purdey, 'Describing Kekerasan: Some Observations on Writing about Violence in Indonesia after the New Order', *Bijdragen Tot de Taal-, Land- En Volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia*, 160.2 (2004), 189–225 <https://doi.org/https://doi.org/10.1163/22134379-90003726>

<sup>70</sup> Samsul Maarif, 'Re-Establishing Human-Nature Relations: Responses of Indigenous People of Indonesia to Covid-19', *Interdisciplinary Journal for Religion and Transformation in Contemporary Society*, 7.2 (2021), 447–72 <https://doi.org/https://doi.org/10.30965/23642807-bja10023>

repairing social harm, fostering mutual respect, and encouraging participatory resolution rather than focusing solely on punitive measures, which is particularly important in online defamation cases where reputational damage can be immediate and widespread. Third, normative reconstruction is necessary to enhance clarity, justice, and human rights protections. Replacing vague terms in Article 27A with precise language by using “individual” instead of “another person” and “an act that undermines the honor or reputation of an individual” instead of “a matter,” and amending Article 45 paragraph five to prioritize restorative justice, would establish a balanced, transparent, and participatory legal framework. Such reconstruction would not only safeguard individual reputations but also protect freedom of expression, facilitate equitable dispute resolution, and ensure that law enforcement aligns with contemporary digital realities, thereby advancing a modern, just, and humanistic legal system in Indonesia.

## References

- Agarwal, Shivang and C Ravindranath Chowdary, ‘Combating Hate Speech Using an Adaptive Ensemble Learning Model with a Case Study on COVID-19’, *Expert Systems with Applications*, 185 (2021), 115632 <https://doi.org/https://doi.org/10.1016/j.eswa.2021.115632>
- Anis Mashdurohatun, Yusfandi Usman, Toni Ariadi Efendi, Purwatik Purwatik and Istiniyati Istiniyati, ‘Rethinking Palm Oil Plastic Regulations for Sustainable and Ecological Justice’, *Journal of Human Rights, Culture and Legal System*, 5 (2025), 500–530 <https://doi.org/10.53955/jhcls.v5i2.681>
- Barber, Kate, ‘Legal Cynicism in Men’s Rights Discourses: Using Corpus Linguistics to Investigate How Distrust in the Legal System Excuses and Perpetuates Sexual Violence against Women’, *Applied Corpus Linguistics*, 5 (2025), 100148 <https://doi.org/https://doi.org/10.1016/j.acorp.2025.100148>
- Borda, Aldo Zammit, ‘International Criminal Courts and Tribunals’, in *Encyclopedia of Violence, Peace, & Conflict (Third Edition)*, ed. by Lester R Kurtz, Third Edition (Oxford: Academic Press, 2022), pp. 627–33 <https://doi.org/https://doi.org/10.1016/B978-0-12-820195-4.00155-2>
- Bray, Kerry, Nils Braakmann and John Wildman, ‘Austerity, Welfare Cuts and Hate Crime: Evidence from the UK’s Age of Austerity’, *Journal of Urban Economics*, 141 (2024), 103439 <https://doi.org/https://doi.org/10.1016/j.jue.2022.103439>
- Chaemsaitong, Krisda and Kewalin Simuang, ‘Lawyers’ Metapragmatic Expressions in the Opening Speech of Anglo-American Criminal Trials’, *Lingua*, 289 (2023), 103536 <https://doi.org/https://doi.org/10.1016/j.lingua.2023.103536>
- Chao, August FY, Chen-Shu Wang, Bo-Yi Li and Hong-Yan Chen, ‘From Hate to Harmony: Leveraging Large Language Models for Safer Speech in Times of COVID-19 Crisis’, *Heliyon*, 10 (2024), e35468 <https://doi.org/https://doi.org/10.1016/j.heliyon.2024.e35468>
- Chetty, Naganna and Sreejith Alathur, ‘Hate Speech Review in the Context of Online Social Networks’, *Aggression and Violent Behavior*, 40 (2018), 108–18 <https://doi.org/https://doi.org/10.1016/j.avb.2018.05.003>
- Choudhary, Mahima, Basant Agarwal and Vishnu Goyal, ‘Hate Speech Detection: Leveraging



- LLM-GPT2 with Fine-Tuning and Multi-Shot Techniques', *Procedia Computer Science*, 258 (2025), 2817–25 <https://doi.org/https://doi.org/10.1016/j.procs.2025.04.542>
- Crowder, Stan, Shawn Mikulay and Brent E Turvey, 'Chapter 2 - Ethical Issues for Students of Criminal Justice', in *Ethical Justice*, ed. by Brent E Turvey and Stan Crowder (San Diego: Academic Press, 2013), pp. 21–62 <https://doi.org/https://doi.org/10.1016/B978-0-12-404597-2.00002-4>
- DeLisi, Matt and Michael G Vaughn, 'Foundation for a Temperament-Based Theory of Antisocial Behavior and Criminal Justice System Involvement', *Journal of Criminal Justice*, 42 (2014), 10–25 <https://doi.org/https://doi.org/10.1016/j.jcrimjus.2013.11.001>
- Demenko, Anna, 'Jurisdiction in Online Speech Crime Cases in the European Union', *International Journal of Law, Crime and Justice*, 80 (2025), 100718 <https://doi.org/https://doi.org/10.1016/j.ijlcj.2024.100718>
- Dillon, Claire, Vikramaditya Joshi and Amra Sabic-El-Rayess, 'Addressing Extremist Abuses of Medieval Past: A Connection-First Approach to Narratives of Hate', *International Journal of Educational Development*, 111 (2024), 103160 <https://doi.org/https://doi.org/10.1016/j.ijedudev.2024.103160>
- Donley, Amy M, James D Wright and Jerian Benwell-Lybarger, 'Violence, Criminalization, and the Homeless', in *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*, ed. by James D Wright, Second Edition (Oxford: Elsevier, 2015), pp. 111–15 <https://doi.org/https://doi.org/10.1016/B978-0-08-097086-8.64101-6>
- Firmino, Anderson Almeida, Cláudio de Souza Baptista and Anselmo Cardoso de Paiva, 'Improving Hate Speech Detection Using Cross-Lingual Learning', *Expert Systems with Applications*, 235 (2024), 121115 <https://doi.org/https://doi.org/10.1016/j.eswa.2023.121115>
- Ghandour, Ahmad, Viktor Shestak and Konstantin Sokolovskiy, 'Security and Privacy of Adolescents in Social Applications and Networks: Legislative Aspects and Legal Practice of Countering Cyberbullying on Example of Developed and Developing Countries', *Journal of Information, Communication and Ethics in Society*, 19 (2021), 433–45 <https://doi.org/https://doi.org/10.1108/JICES-09-2020-0101>
- Chenai, Amira, Zeinab Noorian, Hadiseh Moradisani, Parya Abadeh, Caroline Erentzen and Fattane Zarrinkalam, 'Exploring Hate Speech Dynamics: The Emotional, Linguistic, and Thematic Impact on Social Media Users', *Information Processing & Management*, 62 (2025), 104079 <https://doi.org/https://doi.org/10.1016/j.ipm.2025.104079>
- Giorgini Pignatiello, Giacomo, 'Countering Anti-Lgbti+ Bias in the European Union. A Comparative Analysis of Criminal Policies and Constitutional Issues in Italian, Spanish and French Legislation', *Women's Studies International Forum*, 86 (2021), 102466 <https://doi.org/https://doi.org/10.1016/j.wsif.2021.102466>
- Gorwa, Robert, 'Elections, Institutions, and the Regulatory Politics of Platform Governance: The Case of the German NetzDG', *Telecommunications Policy*, 45 (2021), 102145 <https://doi.org/https://doi.org/10.1016/j.telpol.2021.102145>
- Griffin, Rachel, 'New School Speech Regulation as a Regulatory Strategy against Hate Speech on Social Media: The Case of Germany's NetzDG', *Telecommunications Policy*, 46 (2022), 102411 <https://doi.org/https://doi.org/10.1016/j.telpol.2022.102411>

- Guo, Zhilong, Jie Hao and Lewis Kennedy, 'Protection Path of Personal Data and Privacy in China: Moving from Monism to Dualism in Civil Law and Then in Criminal Law', *Computer Law & Security Review*, 52 (2024), 105928 <https://doi.org/https://doi.org/10.1016/j.clsr.2023.105928>
- Hashmi, Ehtesham, Sule Yildirim Yayilgan, Muhammad Mudassar Yamin, Mohamed Abomhara and Mohib Ullah, 'Self-Supervised Hate Speech Detection in Norwegian Texts with Lexical and Semantic Augmentations', *Expert Systems with Applications*, 264 (2025), 125843 <https://doi.org/https://doi.org/10.1016/j.eswa.2024.125843>
- Helfgott, Jacqueline B, 'Criminal Behavior and the Copycat Effect: Literature Review and Theoretical Framework for Empirical Investigation', *Aggression and Violent Behavior*, 22 (2015), 46–64 <https://doi.org/https://doi.org/10.1016/j.avb.2015.02.002>
- Ibrahim, Yasmine M, Reem Essameldin and Saad M Darwish, 'An Adaptive Hate Speech Detection Approach Using Neutrosophic Neural Networks for Social Media Forensics', *Computers, Materials and Continua*, 79 (2024), 243–62 <https://doi.org/https://doi.org/10.32604/cmc.2024.047840>
- Ibrohim, Muhammad Okky and Indra Budi, 'Hate Speech and Abusive Language Detection in Indonesian Social Media: Progress and Challenges', *Heliyon*, 9 (2023), e18647 <https://doi.org/https://doi.org/10.1016/j.heliyon.2023.e18647>
- Jorgensen, Julia, 'The Functions of Sarcastic Irony in Speech', *Journal of Pragmatics*, 26 (1996), 613–34 [https://doi.org/https://doi.org/10.1016/0378-2166\(95\)00067-4](https://doi.org/https://doi.org/10.1016/0378-2166(95)00067-4)
- Kaminska, Olha, Chris Cornelis and Veronique Hoste, 'Fuzzy Rough Nearest Neighbour Methods for Detecting Emotions, Hate Speech and Irony', *Information Sciences*, 625 (2023), 521–35 <https://doi.org/https://doi.org/10.1016/j.ins.2023.01.054>
- Kapil, Prashant and Asif Ekbal, 'A Deep Neural Network Based Multi-Task Learning Approach to Hate Speech Detection', *Knowledge-Based Systems*, 210 (2020), 106458 <https://doi.org/https://doi.org/10.1016/j.knosys.2020.106458>
- Kar, Purbani and Swapan Debbarma, 'Sentimental Analysis & Hate Speech Detection on English and German Text Collected from Social Media Platforms Using Optimal Feature Extraction and Hybrid Diagonal Gated Recurrent Neural Network', *Engineering Applications of Artificial Intelligence*, 126 (2023), 107143 <https://doi.org/https://doi.org/10.1016/j.engappai.2023.107143>
- Kloos, David, 'A Crazy State', *Bijdragen Tot de Taal-, Land- En Volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia*, 170 (2014), 25–65 <https://doi.org/https://doi.org/10.1163/22134379-17001003>
- Labadie-Tamayo, Roberto, Djordje Slijepčević, Xihui Chen, Adrian Jaques Böck, Andreas Babic, Liz Freimann, and others, 'Distilling Knowledge from Large Language Models: A Concept Bottleneck Model for Hate and Counter Speech Recognition', *Information Processing & Management*, 63 (2026), 104309 <https://doi.org/https://doi.org/10.1016/j.ipm.2025.104309>
- Lanza, Giulia, 'The Fundamental Role of International (Criminal) Law in the War in Ukraine', *Orbis*, 66 (2022), 424–35 <https://doi.org/https://doi.org/10.1016/j.orbis.2022.05.010>
- Levin, Jack and Jack McDevitt, 'Hate Crimes', in *Encyclopedia of Violence, Peace, & Conflict (Third Edition)*, ed. by Lester R Kurtz, Third Edition (Oxford: Academic Press, 2022), pp.

28–34 <https://doi.org/https://doi.org/10.1016/B978-0-12-820195-4.00040-6>

Lievens, Eva, ‘Bullying and Sexting in Social Networks: Protecting Minors from Criminal Acts or Empowering Minors to Cope with Risky Behaviour?’, *International Journal of Law, Crime and Justice*, 42 (2014), 251–70 <https://doi.org/https://doi.org/10.1016/j.ijlcj.2014.02.001>

Maarif, Samsul, ‘Re-Establishing Human-Nature Relations: Responses of Indigenous People of Indonesia to Covid-19’, *Interdisciplinary Journal for Religion and Transformation in Contemporary Society*, 7 (2021), 447–72 <https://doi.org/https://doi.org/10.30965/23642807-bja10023>

Maaß, Sabrina, Jil Wortelker and Armin Rott, ‘Evaluating the Regulation of Social Media: An Empirical Study of the German NetzDG and Facebook’, *Telecommunications Policy*, 48 (2024), 102719 <https://doi.org/https://doi.org/10.1016/j.telpol.2024.102719>

Marder, Ian D, ‘Mapping Restorative Justice and Restorative Practices in Criminal Justice in the Republic of Ireland’, *International Journal of Law, Crime and Justice*, 70 (2022), 100544 <https://doi.org/https://doi.org/10.1016/j.ijlcj.2022.100544>

McKearney, Aidan, ‘Changing Contexts: From Criminal to Citizen’, *Journal of Organizational Change Management*, 33 (2019), 515–26 <https://doi.org/https://doi.org/10.1108/JOCM-11-2018-0331>

Miller, Ronald and Brent E Turvey, ‘Chapter 9 - Ethical Issues for Criminal Prosecutors’, in *Ethical Justice*, ed. by Brent E Turvey and Stan Crowder (San Diego: Academic Press, 2013), pp. 257–316 <https://doi.org/https://doi.org/10.1016/B978-0-12-404597-2.00009-7>

Modha, Sandip, Prasenjit Majumder, Thomas Mandl and Chintak Mandalia, ‘Detecting and Visualizing Hate Speech in Social Media: A Cyber Watchdog for Surveillance’, *Expert Systems with Applications*, 161 (2020), 113725 <https://doi.org/https://doi.org/10.1016/j.eswa.2020.113725>

Mossie, Zewdie and Jenq-Haur Wang, ‘Vulnerable Community Identification Using Hate Speech Detection on Social Media’, *Information Processing & Management*, 57 (2020), 102087 <https://doi.org/https://doi.org/10.1016/j.ipm.2019.102087>

Mugari, Ishmael and Caleb Muzinda, ‘Social Media: The Double-Edged Sword for Perpetuating and Preventing Criminal Activities in Zimbabwe’, *Safer Communities*, 20 (2021), 225–36 <https://doi.org/https://doi.org/10.1108/SC-02-2021-0004>

Muhammad, Rustamaji, Shalih Mangara Sitompul, Tojiboyev Sarvar Zafarovich and Rahimah Embong, ‘The Reduction of Criminal Justice Policy in Indonesia: Justice versus Virality’, *Journal of Human Rights, Culture and Legal System*, 5 (2025), 442–72 <https://doi.org/10.53955/jhcls.v5i2.637>

Nave, Eva and Lottie Lane, ‘Countering Online Hate Speech: How Does Human Rights Due Diligence Impact Terms of Service?’, *Computer Law & Security Review*, 51 (2023), 105884 <https://doi.org/https://doi.org/10.1016/j.clsr.2023.105884>

Pamungkas, Endang Wahyu, Valerio Basile and Viviana Patti, ‘A Joint Learning Approach with Knowledge Injection for Zero-Shot Cross-Lingual Hate Speech Detection’, *Information Processing & Management*, 58 (2021), 102544 <https://doi.org/https://doi.org/10.1016/j.ipm.2021.102544>

- Parvaresh, Vahid, 'Covertly Communicated Hate Speech: A Corpus-Assisted Pragmatic Study', *Journal of Pragmatics*, 205 (2023), 63–77  
<https://doi.org/https://doi.org/10.1016/j.pragma.2022.12.009>
- Purdey, Jemma, 'Describing Kekerasan; Some Observations on Writing about Violence in Indonesia after the New Order', *Bijdragen Tot de Taal-, Land- En Volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia*, 160 (2004), 189–225  
<https://doi.org/https://doi.org/10.1163/22134379-90003726>
- Sadeghi, Atiyeh, Sebastian Pape and David Harborth, 'The Impact of Individuals' Social Environments on Contact Tracing App Use: Survey Study', *JMIR Human Factors*, 10 (2023)  
<https://doi.org/https://doi.org/10.2196/45825>
- Schlesinger, Louis B and Tirza Z Ben-Ari, 'Personality Disorders and Criminal Behavior', in *Encyclopedia of Forensic Sciences, Third Edition (Third Edition)*, ed. by Max M Houck, Third Edition (Oxford: Elsevier, 2023), pp. 133–43  
<https://doi.org/https://doi.org/10.1016/B978-0-12-823677-2.00057-X>
- Sell, Aaron and Daniel Sznycer, 'Societal Institutions Echo Evolved Human Nature: An Analysis of the Western Criminal Justice System and Its Relation to Anger', *Evolution and Human Behavior*, 44 (2023), 210–21  
<https://doi.org/https://doi.org/10.1016/j.evolhumbehav.2023.01.007>
- Shenkin, Evan N and Michele Abee, 'International Spaces for Feminist Cross-Border Resistance', *Political Geography*, 112 (2024), 103111  
<https://doi.org/https://doi.org/10.1016/j.polgeo.2024.103111>
- Solhjell, Randi, 'How Acts Become Hate Crime: The Police's Documenting of Criminal Cases', *International Journal of Law, Crime and Justice*, 72 (2023), 100574  
<https://doi.org/https://doi.org/10.1016/j.ijlcrj.2022.100574>
- Spaid, Sue, 'Chapter 18 - On Deterring Hate Speech, While Maximizing Security and Privacy', in *Ethics in Online AI-Based Systems*, ed. by Santi Caballé, Joan Casas-Roma, and Jordi Conesa, Intelligent Data-Centric Systems (Academic Press, 2024), pp. 365–81  
<https://doi.org/https://doi.org/10.1016/B978-0-443-18851-0.00021-4>
- Srivastava, Atul Kumar, Mitali Srivastava, Sanchali Das, Vikas Jain and Tej Bahadur Chandra, 'Leveraging Deep Learning for Comprehensive Multilingual Hate Speech Detection', *Procedia Computer Science*, 252 (2025), 832–40  
<https://doi.org/https://doi.org/10.1016/j.procs.2025.01.044>
- Sukabdi, Zora A, 'Bridging the Gap: Contributions of Academics and National Security Practitioners to Counterterrorism in Indonesia', *International Journal of Law, Crime and Justice*, 65 (2021), 100467  
<https://doi.org/https://doi.org/10.1016/j.ijlcrj.2021.100467>
- Sultan, Daniyar, Aigerim Toktarova, Ainur Zhumadillayeva, Sapargali Aldeshov, Shynar Mussiraliyeva, Gulbakhram Beissenova, and others, 'Cyberbullying-Related Hate Speech Detection Using Shallow-to-Deep Learning', *Computers, Materials and Continua*, 74 (2022), 2115–31  
<https://doi.org/https://doi.org/10.32604/cmc.2023.032993>
- Sundari, Elisabeth, Helidorus Chandera Halim and Ousu Mendy, 'Should Indonesia Adopt Legal Representation in Civil Cases?', *Journal of Human Rights, Culture and Legal System*, 5 (2025), 554–80  
<https://doi.org/10.53955/jhcls.v5i2.604>
- Tan, Tay Keong, 'Chapter 17 - Curse of the Cyborg Mammoths: The Use of Artificial



- Intelligence in Manipulating and Mobilizing Human Emotions’, in *Ethics in Online AI-Based Systems*, ed. by Santi Caballé, Joan Casas-Roma, and Jordi Conesa, Intelligent Data-Centric Systems (Academic Press, 2024), pp. 347–64 <https://doi.org/https://doi.org/10.1016/B978-0-443-18851-0.00018-4>
- Turvey, Brent E and Aurelio Coronado, ‘Chapter 2 - Victimhood: Entering the Criminal Justice System’, in *Forensic Victimology (Third Edition)*, ed. by Brent E Turvey, Third Edition (San Diego: Academic Press, 2023), pp. 33–76 <https://doi.org/https://doi.org/10.1016/B978-0-12-821768-9.00002-1>
- Udanor, Collins and Chinatu C Anyanwu, ‘Combating the Challenges of Social Media Hate Speech in a Polarized Society’, *Data Technologies and Applications*, 53 (2019), 501–27 <https://doi.org/https://doi.org/10.1108/DTA-01-2019-0007>
- del Valle-Cano, Gloria, Lara Quijano-Sánchez, Federico Liberatore and Jesús Gómez, ‘SocialHaterBERT: A Dichotomous Approach for Automatically Detecting Hate Speech on Twitter through Textual Analysis and User Profiles’, *Expert Systems with Applications*, 216 (2023), 119446 <https://doi.org/https://doi.org/10.1016/j.eswa.2022.119446>
- Wagner, Ben, Matthias C Kettemann, Anna Sophia Tiedeke, Felicitas Rachinger and Marie-Therese Sekwenz, ‘Mapping Interpretations of the Law in Online Content Moderation in Germany’, *Computer Law & Security Review*, 55 (2024), 106054 <https://doi.org/https://doi.org/10.1016/j.clsr.2024.106054>
- Wisnubroto, Aloysius and Hilaire Tegan, ‘Preventing AI Crime Towards A New Legal Paradigm: Lessons From United States’, *Journal of Human Rights, Culture and Legal System*, 5 (2025), 630–58 <https://doi.org/10.53955/jhcls.v5i2.606>
- Wright, Craig, ‘Chapter 21 - Information Systems Legislation’, in *The IT Regulatory and Standards Compliance Handbook*, ed. by Craig Wright (Burlington: Syngress, 2008), pp. 609–71 <https://doi.org/https://doi.org/10.1016/B978-1-59749-266-9.00021-7>
- Yasmin, Musarat, Tahira Jabeen and Sania Noor, ‘Examining Digital Hate: Gendered Discursive Strategies in Online Harassment of Pakistani Influencers’, *Computers in Human Behavior Reports*, 19 (2025), 100759 <https://doi.org/https://doi.org/10.1016/j.chbr.2025.100759>
- Zahrah, Fatima, Jason RC Nurse and Michael Goldsmith, ‘Unmasking Hate in the Pandemic: A Cross-Platform Study of the COVID-19 Infodemic’, *Big Data Research*, 37 (2024), 100481 <https://doi.org/https://doi.org/10.1016/j.bdr.2024.100481>
- Zapata Roza, Andrés, Alejandra Campo-Archbold, Daniel Díaz-López, Ian Gray, Javier Pastor-Galindo, Pantaleone Nespole, and others, ‘Cyber Democracy in the Digital Age: Characterizing Hate Networks in the 2022 US Midterm Elections’, *Information Fusion*, 110 (2024), 102459 <https://doi.org/https://doi.org/10.1016/j.inffus.2024.102459>