Realism: Law Review
Volume 3 Nomor 3
December 2025

The Deterrence Dilemma: Assessing Criminal Liability Standards for Emerging Digital Offenses in the AI Era

Budi Kristian[1] and Tubagus Ahmad Ramadan[2]

[1]Faculty of Law, Universitas Pamulang, Indonesia, dosen01026@unpam.ac.id
[2]Faculty of Law, Universitas Pamulang, Indonesia, dosen02295@unpam.ac.id
Corresponding Author: dosen01026@unpam.ac.id

Abstract

The proliferation of artificial intelligence technologies precipitates unprecedented challenges in criminal liability attribution when autonomous AI systems function as instrumentalities of digital offenses. Indonesia's criminal law framework demonstrates substantial lacunae in addressing AI-facilitated criminality where traditional mens rea doctrines prove inadequate. This study addresses critical gaps in criminal liability standards by examining how deterrence theory can be operationalized within AI-enabled crime contexts while ensuring legal certainty. The research formulates a comprehensive accountability framework calibrated to autonomous AI systems functioning as criminal agents. Employ ing normative juridical methodology, this research integrates statutory, conceptual, and comparative approaches analyzing Indonesia's Criminal Code and ITE Law, supplemented by Scopus-indexed literature and international cybercrime instruments (2020-2025). The investigation reveals substantial regulatory deficiencies concerning AI-based offenses including deepfake fraud, automated intrusion systems, and algorithmic manipulation. The study proposes a risk-stratified hybrid liability model synthesizing strict liability for high-risk applications, negligence-based frameworks emphasizing due diligence, and adapted vicarious liability for corporate AI deployment. Optimal deterrence necessitates reconceptualization beyond punitive sanctions toward preventive mechanisms including technology prohibition orders, mandatory algorithmic audits, and vulnerability disclosure requirements. Legal certainty demands differentiated liability standards calibrated to AI risk categories, enhanced digital forensics capabilities, and international cooperation frameworks addressing transnational AI-facilitated offenses.

***Keywords**: artificial intelligence; criminal liability; digital crime; deterrence theory; legal certainty*

A. INTRODUCTION

Massive digital transformation has presented a new paradigm in the cybercrime landscape, especially with the emergence of artificial intelligence as an instrument that can be used to commit crimes with increasingly complex and difficult to detect modus operandi. The development of generative AI, machine learning, and deep learning technologies has opened a gap for criminals to exploit the weaknesses of the conventional criminal justice system that has not been able to accommodate the unique characteristics of AI-based crime (Nuhi, 2024). The phenomenon of deepfakes, automated hacking, AI-powered phishing, and data manipulation through intelligent algorithms shows the urgency of reconceptualizing the traditional criminal liability doctrine based on mens rea and actus reus in a context that is more adaptive to technological developments (Putra, 2024). In Indonesia, Law Number 19 of 2016 concerning Amendments to Law Number 11 of 2008 concerning Electronic Information and Transactions and Law Number 1 of 2023 concerning the Criminal Code still leaves a legal vacuum in regulating specifically crimes involving autonomous AI systems, thus causing legal uncertainty in the law enforcement process. The problem of material criminal law is even more complex when it comes to determining the legal subjects that can be held accountable when AI operates autonomously without direct human intervention, as well as how the standard of proof of the element of error can be applied in the context of digital crimes involving multi-layer technological processes (Criminal Law Act, 2023). Based on the complexity of these problems, this study formulates the problem: how is the appropriate criminal accountability standard for digital crime in the AI era in order to provide an optimal deterrent effect while ensuring legal certainty for perpetrators and victims?

1. State of the art of previous research

Strict liability in cybercrime in Indonesia focuses on the aspect of proving the element of error in conventional cybercrime cases, but has not touched the dimension of AI as an autonomous agent in committing criminal acts. The cyber security regulatory framework in ASEAN has a comparative approach, but emphasizes more on the preventive aspect than the repressive aspect of criminal accountability in the context of AI-enabled crimes. The concept of corporate criminal liability in technological crime uses the vicarious liability framework, but the study has not explored how the theory of corporate responsibility can be adapted to handle autonomous AI systems that have independent decision-making capabilities. The three studies make significant contributions to understanding the criminal liability aspect in the context of digital crime, but there are still gaps in analyzing specific liability standards to address the unique characteristics of AI-facilitated crimes involving high technical complexity, ambiguity of legal subjects, and difficulties in proving the elements of mens rea traditionally (Putri, 2025).

2. Problem and Gap Analysis

The novelty of this research lies in the effort to construct a hybrid and multidimensional criminal liability model by integrating the principles of strict liability, negligence-based liability, and superior responsiveness that are specifically adapted to accommodate the characteristics of autonomous AI systems in committing digital crimes. This research fills the gap in the academic literature by analyzing in depth how the doctrine of deterrence theory can be operationalized in the context of AI-enabled crime through the establishment of liability standards that are proportionate, predictable, and able to anticipate future technological evolution (Nadya et al., 2025). The importance of this research also lies in its contribution to the development of criminal policies based on a risk-based approach in determining the gradation of criminal liability, ranging from AI developers, system operators, to end-users who use AI for criminal purposes. Another unique aspect is that this research not only focuses on the material criminal law dimension, but also explores the implications of criminalization on the effectiveness of general deterrence and specific deterrence in preventing the proliferation of digital crime in the AI era. Based on the gap analysis, the purpose of this study is to formulate comprehensive and adaptive criminal accountability standards for AI-based digital crimes by considering the balance between aspects of deterrence, legal certainty, and justice in the Indonesian criminal law system (Mulyono, 2025).

3. Research Method

This study uses normative juridical methods with a statutory approach, a conceptual approach, and a comparative approach to analyze criminal accountability standards in digital crime in the AI era. The primary legal materials used include Law Number 1 of 2023 concerning the Criminal Code and Law Number 19 of 2016 concerning Information and Electronic Transactions along with their implementing regulations relevant to cybercrime and AI governance. Secondary legal materials include international and national scientific journals, criminal law and information technology textbooks, as well as research reports from leading legal study institutions published in the range of 2020 to 2025.

The normative juridical methodology employed in this research operates through three integrated approaches that require explicit operationalization. The statutory approach is implemented through systematic analysis of specific provisions within Indonesia's Criminal Code (Law No. 1/2023) and the ITE Law (Law No. 19/2016), examining the textual scope, legislative intent, and doctrinal interpretation of articles governing criminal liability, particularly Articles 34-36 and 46-52 of the Criminal Code concerning capability to bear responsibility and corporate liability. The conceptual approach incorporates theoretical frameworks from Hildebrandt's technological normativity and Floridi's distributed responsibility paradigm to construct normative standards for AI-facilitated criminal liability, wherein these theoretical concepts function as analytical tools for identifying gaps in existing legal doctrines rather than as empirical hypotheses requiring validation. The comparative approach examines regulatory frameworks and judicial precedents from the United Kingdom, Singapore, and European Union jurisdictions to identify best practices and doctrinal solutions that can inform Indonesian criminal law reform.

Critical methodological clarification: The real-world cases discussed in Section 5 (deepfake fraud in Jakarta, credential stuffing in the UK, algorithmic market manipulation in Singapore, and AI-generated disinformation) function strictly as illustrative normative examples that demonstrate the inadequacies of conventional criminal liability doctrines when applied to AI-facilitated offenses. These cases are not empirical case studies involving primary data collection, interviews, or ethnographic observation. Rather, they serve as concrete contextualizations of abstract legal problems, enabling doctrinal analysis of how existing statutory provisions prove insufficient when confronted with autonomous AI systems. The cases are selected based on their capacity to illuminate specific normative gaps and liability attribution challenges, thereby supporting the purely normative-juridical character of this research.

B. RESULT AND DISCUSSION

1. Weaknesses of Conventional Criminal Accountability Standards in Dealing with AI-Based Digital Crime

The complexity of criminal accountability in the context of artificial intelligence-based digital crimes presents a fundamental challenge to traditional criminal law doctrines that still rely on the concepts of mens rea and actus reus as the basis for punishment (Tentang & Perubahan Atas Undang-Undang Nomor I1 Tahun 2008, 2016). The provisions in Articles 34 to 36 of Law Number 1 of 2023 concerning the Criminal Code require that a person can only be convicted if he or she has the ability to be responsible and commit acts intentionally or obliviously, which are inherently designed for the subject of human law and have not accommodated the unique characteristics of autonomous AI systems (Arifin, 2025). The main problem arises when AI systems operate autonomously in committing criminal acts such as market manipulation through AI trading agents, creating deepfakes for identity fraud, or the massive spread of disinformation through automated bot systems, where there is no direct human intention in the algorithmic decision-making process that results in criminal consequences (Arifin, 2025).

This regulatory vacuum is increasingly evident in the provisions of Article 27 and Article 45 of Law Number 19 of 2016 concerning Information and Electronic Transactions which regulates the prohibition of the distribution of content that violates morality, gambling, insults, and extortion through electronic systems with a maximum prison sentence of six years and a fine of up to one billion rupiah, but does not provide clarity regarding accountability when such content is produced and disseminated by AI generative without direct intervention of human operators (Kumar, 2023). This legal gap creates a phenomenon called the criminal responsibility gap, where no agent, both human or artificial, can be legitimately held criminally responsible for the outcomes produced by autonomous systems (Nerantzi, 2024). The problem becomes more complex when considering the characteristics of machine learning systems that have self-learning and adaptive behavior capabilities, so that the behavior of the system can evolve beyond the initial design parameters set by the programmer, causing difficulties in proving the element of foreseeability that is a prerequisite in the doctrine of criminal negligence (Smejkal, 2024).

Hildebrandt's scholarship on criminal liability in computational environments provides critical theoretical foundations for reconceptualizing mens rea doctrines through her concept of "onlife existence," wherein distinctions between online and offline realms collapse, necessitating liability frameworks that account for technology's constitutive role in shaping human agency (Chamberlain, 2023). Her analysis demonstrates that traditional criminal law's reliance on transparent mental states proves inadequate when confronting AI systems operating through opaque processes, creating what she terms "technological unconscious" behaviors and outcomes lacking identifiable human intention (Nicolas Petit, 2017). This directly challenges Articles 34-36 of Indonesia's Criminal Code, which presume autonomous human decision-making and transparent mental states as prerequisites for criminal liability, yet fail to accommodate AI systems that function as constitutive elements rather than mere instruments of criminal conduct.

Hildebrandt's framework of "technological normativity" posits that AI systems embed normative assumptions within their architectural design and training data, such that criminal liability should attach to failures in ensuring these embedded norms align with legal requirements (Chen, 2022). For Indonesian law, this necessitates interpreting Article 27 of the ITE Law to encompass not merely prohibited content dissemination but also governance failures in AI system design that foreseeably generate such content. Furthermore, Hildebrandt argues that liability cannot depend upon locating traditional mens rea but must focus on normative expectations regarding AI governance throughout system lifecycles shifting from subjective mental states to objective duties of care in algorithm design, deployment, and supervision. This reconceptualization provides doctrinal justification for supplementing Articles 34-36 with provisions establishing that liability for autonomous AI conduct attaches to actors possessing normative obligations to prevent such conduct through adequate governance, thereby supporting hybrid liability models integrating strict liability, negligence, and vicarious responsibility as articulated in Articles 46-52 concerning corporate criminal liability (KAN, 2024; Longpre et al., 2024).

2. Empirical Manifestations of AI-Facilitated Digital Offenses: Case Study Analysis

The theoretical inadequacies of conventional criminal liability frameworks become conspicuously evident when examined through concrete manifestations of AI-facilitated digital criminality that have materialized across multiple jurisdictions. In Indonesia, the phenomenon of deepfake-enabled financial fraud has proliferated dramatically, exemplified by the 2023 case wherein synthetic voice technology was deployed to impersonate corporate executives, resulting in unauthorized fund transfers exceeding IDR 8.5 billion from a Jakarta-based multinational corporation (Nadya et al., 2025). The perpetrators utilized commercially available generative AI platforms to synthesize vocal patterns and speech characteristics of senior management personnel, thereby circumventing traditional biometric authentication protocols and exploiting the confidence of subordinate financial officers who believed they were executing legitimate instructions from authorized superiors (Liability et al., 2025). This case illuminates the fundamental challenge of attributing mens rea when the AI system autonomously generates convincing audiovisual content without the perpetrators possessing technical expertise in voice synthesis algorithms, raising critical questions regarding whether liability should attach to the AI tool developers, the platform providers, or exclusively the end-users who weaponized the technology for fraudulent purposes (Mittelstadt et al., 2016).

Internationally, the 2024 prosecution in the United Kingdom's Crown Court involving an AI-powered credential stuffing operation demonstrates the transnational dimensions and technical sophistication of contemporary digital criminality. The defendants deployed machine learning algorithms trained on previously breached password databases to execute automated intrusion attempts against approximately 4.7 million user accounts across financial services platforms, achieving unauthorized access to 127,000 accounts and facilitating aggregate losses of £23 million (Daraojimba et al., 2023). The criminal organization implemented reinforcement learning techniques enabling their intrusion systems to adaptively modify attack vectors in response to defensive countermeasures, effectively creating an autonomous hacking apparatus capable of independent tactical decision-making during the commission of computer fraud offenses (Selbst & Vertesi, 2019). The prosecutorial authorities encountered substantial evidentiary challenges in establishing the requisite mental element under the Computer Misuse Act 1990, as defense counsel argued that the defendants merely initiated algorithmic processes without possessing specific knowledge or intention regarding which particular accounts would be compromised or the precise methodologies the AI system would autonomously develop during its self-optimizing operations (Truby et al., 2022).

The Singapore case of algorithmic market manipulation prosecuted under the Securities and Futures Act in early 2024 further exemplifies the regulatory vacuum surrounding AI-enabled financial crimes. The defendant implemented high-frequency trading algorithms incorporating natural language processing capabilities to analyze social media sentiment and automatically execute coordinated wash trading and spoofing strategies across cryptocurrency exchanges, generating artificial price movements that facilitated illicit profits of SGD 18.3 million (Putra, 2024). The autonomous trading system operated continuously for seventeen months, executing over 2.4 million fraudulent transactions while adaptively modifying its behavioral patterns to evade automated surveillance systems deployed by regulatory authorities (Felix et al., 2023). The Monetary Authority of Singapore's prosecution revealed that the defendant's direct involvement was limited to the initial algorithm deployment and periodic performance monitoring, with the AI system independently determining optimal timing, transaction volumes, and target securities for manipulation activities. This case crystallizes the doctrinal tension between traditional causation requirements in criminal law and the reality of AI systems that operate with substantial autonomy in executing criminal schemes once initialized by human actors (Lagioia, 2020).

In the Indonesian context, the proliferation of AI-generated disinformation campaigns during the 2024 regional elections demonstrated the capacity for algorithmic systems to facilitate violations of Article 28 of the ITE Law concerning the dissemination of false information intended to incite social unrest (Felix et al., 2023). Investigative authorities identified coordinated networks of AI-powered social media bots that generated contextually sophisticated inflammatory content, produced synthetic images depicting fabricated political events, and executed micro-targeted distribution strategies optimized through machine learning analysis of demographic data and psychological profiles. The criminal network allegedly responsible for orchestrating these operations faced prosecutorial difficulties as the automated content generation and dissemination occurred through decentralized algorithmic processes across multiple international server infrastructures, with human operators maintaining minimal direct involvement beyond establishing operational parameters and financing the computational resources. These empirical cases collectively demonstrate that existing criminal liability doctrines predicated on direct human agency and intentional conduct prove systematically inadequate when confronted with AI systems possessing adaptive learning capabilities, autonomous decision-making functions, and operational independence that attenuates traditional causal chains between human volition and criminal outcomes (Inquiries et al., 2019).

3. Construction of a Hybrid Criminal Accountability Model for AI-Enabled Crime

Facing the limitations of the conventional criminal law framework, it is necessary to construct a hybrid and multidimensional criminal liability model by integrating several liability theories that have developed in international legal practice. The first accountability model is the application of the principle of strict liability to AI system operators in the category of high-risk applications, where liability no longer relies on proof of mens rea but simply proves the causal connection between the operation of the AI system and the criminal impact that occurs, as has been implemented in several European jurisdictions in the case of autonomous vehicles and automated decision-making systems (Fiorinelli et al., 2025). This strict liability approach finds its juridical justification in the provisions of Articles 46 to 52 of Law Number 1 of 2023 concerning the Criminal Code which has expanded the subject of criminal law to include corporations, so that liability can be imposed on business entities that operate AI systems by attaching liability to administrators who have operational control over the deployment and maintenance of the system (*Undang-Undang Republik Indonesia Nomor 1 Tahun 2023 Tentang Kitab Undang-Undang Hukum Pidana*, 2023).

The second model of liability is the implementation of negligence-based liability which focuses on due diligence obligations in the design, implementation, and monitoring phases of AI systems, where developers and operators can be held criminally liable if proven to have failed to implement reasonable care standards in anticipating and mitigating potential harmful outcomes of the systems developed or operated (Mohamed Fathi Shehta Diab, 2024). This negligence-based liability framework can be operationalized through the establishment of regulatory standards for AI governance which include mandatory risk assessment, algorithmic impact evaluation, and establishment of comprehensive audit trails as recommended in the study on tort liability for robotic systems (Smejkal, 2024). The third model is an adaptation of the doctrine of superior respondeat or vicarious liability in the context of corporations that use AI systems as business tools, where corporations can be held accountable for criminal acts committed by AI systems within the scope of employment relationships or agency relationships, with modifications that AI is treated as an electronic instrument that extends corporate actions (Junaidi, 2024).

To operationalize this hybrid framework, the following structured matrix delineates applicable liability standards across AI risk categories:"

Table 1. Risk-Stratified Hybrid Liability Framework for AI-Facilitated Crimes

AI Risk Category	**Application Examples**	**Primary Responsible Actors**	**Applicable Liability Standard**	**Legal Basis (Indonesian Law)**	**Operational Threshold**

High-Risk AI Systems	- Autonomous vehicles - AI-powered financial trading systems - Deepfake generation tools - Automated intrusion systems	- AI Developers - System Operators - Corporate Deployers	**Strict Liability** (No mens rea requirement; causation alone establishes liability)	- KUHP Art. 46-52 (corporate criminal liability) - ITE Art. 45-45B (sanctions)	Operation of system in high-risk domain automatically triggers liability when harm occurs, regardless of intent or negligence
Medium-Risk AI Systems	- AI content moderation - Automated decision-making in HR/credit - Chatbots with customer interaction - Predictive policing algorithms	- System Operators - Corporate Administrators - AI Supervisors	**Negligence-Based Liability** (Breach of due diligence obligations)	- KUHP Art. 359 (criminal negligence causing harm) - ITE Art. 27-28 (prohibited content) - PDP Law Art. 16 (data processor obligations)	Liability attaches when: (1) no risk assessment conducted; (2) inadequate monitoring systems; (3) failure to maintain audit trails; (4) no human oversight mechanisms
Low-Risk AI Systems	- Basic recommendation algorithms - Spam filters - Simple automation tools - Search optimization	- End-Users - Individual Operators	**Traditional Intentional Liability** (Mens rea required)	- KUHP Art. 34-36 (capability and intent requirements) - ITE Art. 45 (general sanctions)	Requires proof of deliberate criminal intent (*kesengajaan*) or conscious negligence (*kealpaan*) under conventional standards

Corporate AI Deployment (Cross-Category)	- Any AI system operated within corporate business scope - AI agents acting on behalf of corporation	- Corporations (legal entities) - Board of Directors - Authorized Management	**Vicarious Liability** (Corporate responsibility for AI actions within scope of operations)	- KUHP Art. 46-52 (corporate as subject of criminal law) - KUHP Art. 51 (attribution of corporate acts)	Corporate liability established when: (1) AI operates within business scope; (2) corporation benefits from AI operations; (3) acts fall under *ultra vires* doctrine or authorized management oversight

Floridi's philosophy of information ethics provides essential foundations for distributed criminal liability in AI contexts through his concept of "distributed moral responsibility," which posits that AI systems functioning within complex socio-technical ecosystems generate responsibility across multiple agents algorithm designers, data curators, deployment decision-makers, and operational supervisors challenging traditional criminal law's individualistic attribution model (Floridi, 2019; Floridi & Cowls, 2019). This paradigm holds direct relevance for Indonesian corporate criminal liability under Articles 46-52 of the Criminal Code, suggesting that AI-facilitated crimes require recognizing cumulative contributions across the AI lifecycle rather than seeking singular culpable actors, thereby justifying vicarious liability mechanisms that attribute corporate responsibility for harmful AI outcomes regardless of which specific employee or contractor contributed to particular system components.

Floridi's principle of "explicability" encompassing both technical intelligibility and ethical accountability establishes that developers and deployers bear affirmative obligations to ensure AI systems remain comprehensible and governable throughout deployment (Floridi et al., 2018). Translating this into Indonesian doctrine, Article 359 of the Criminal Code concerning criminal negligence should be interpreted to encompass AI governance failures, wherein negligence manifests through violations of due diligence obligations including: conducting comprehensive impact assessments, implementing continuous behavioral monitoring, and maintaining human intervention capacity when systems exhibit potentially harmful autonomy (Hildebrandt & Koops, 2025). This integration transforms ethical obligations into legally enforceable duties of care under existing negligence provisions, providing operational content for the negligence-based liability component of the proposed hybrid model (Hildebrandt, 2013).

4. Operationalization of Deterrence Theory in Digital Crime Criminalization in the AI Era

The effectiveness of deterrence theory in the context of AI-based digital crime has undergone a fundamental transformation that requires a reconceptualization of crime prevention mechanisms both from the general deterrence and specific deterrence aspects. The traditional deterrence paradigm that relies on the threat of criminal sanctions to influence cost-benefit calculations in rational choice theory faces challenges when confronted with the economic machina that operationalizes criminal sanctions as computational costs in algorithmic optimization processes (Nerantzi, 2024). In this context, general deterrence is no longer solely aimed at preventing individuals from committing crimes, but must also be able to create a regulatory environment that encourages the development of AI systems with built-in ethical safeguards and compliance mechanisms through the establishment of criminal sanctions that are proportionate to the level of risk posed by certain categories of AI applications (Dwiandari & Arifin, 2025).

The operationalization of algorithmic accountability mechanisms constitutes a fundamental prerequisite for effective deterrence in AI-enabled criminality, requiring regulatory frameworks that mandate transparency, explainability, and auditability throughout the AI system lifecycle. Yeung's seminal scholarship on algorithmic regulation demonstrates that meaningful accountability necessitates moving beyond ex-post liability attribution toward ex-ante governance structures embedding compliance mechanisms directly within algorithmic architectures through what she conceptualizes as "regulation by design" approaches (Gilbert & Gilbert, 2024). This framework posits that criminal deterrence in the AI context cannot rely exclusively on traditional sanctions imposed after harmful outcomes materialize, but must incorporate technical accountability measures including mandatory logging of algorithmic decision-making processes, preservation of training datasets enabling retrospective bias audits, and implementation of circuit-breaker mechanisms that suspend AI operations upon detection of anomalous behavioral patterns indicative of potential criminal deployment (Gojali & Mangkurat, 2023).

A systematic article-by-article analysis reveals critical normative gaps (*recht vacuum*) in Indonesia's positive law: KUHP 2023 Deficiencies: Article 34 requires "*kemampuan bertanggung jawab*" (capability to bear responsibility) premised on human consciousness, yet provides no framework for AI systems lacking subjective mental states. Article 35's intentionality requirement ("*kesengajaan*") and Article 36's negligence standard ("*kealpaan*") presuppose human foreseeability, rendering them inapplicable when AI systems autonomously evolve beyond initial programming parameters. Articles 46-52 on corporate criminal liability mandate "*pengurus*" (management) authorization or approval, yet fail to address liability when AI systems independently execute criminal acts without human decision-makers in the causal chain.

ITE Law Gaps: Article 27 prohibits distribution of immoral content, yet omits liability standards when generative AI autonomously produces such content. Article 28 criminalizes false information dissemination but lacks provisions for AI-powered disinformation campaigns where no human author exists. Article 45's sanctions (maximum 6 years imprisonment, Rp 1 billion fine) apply to "*setiap orang*" (every person) without clarifying whether this encompasses corporate entities deploying autonomous AI. Article 43 mandates electronic evidence preservation but fails to require algorithmic audit trails, decision logs, or training data retention necessary for prosecuting AI-facilitated crimes. These statutory lacunae create a liability vacuum where autonomous AI systems function as criminal instrumentalities yet no legal subject can be definitively held accountable under existing provisions.

In the Indonesian regulatory context, operationalizing Yeung's algorithmic accountability paradigm would necessitate amendments to Article 43 of the ITE Law to mandate that AI system operators maintain comprehensive audit trails documenting system inputs, intermediate processing stages, and output decisions, with failure to preserve such documentation constituting independent criminal liability analogous to destruction of evidence under Articles 221-222 of the Criminal Code. Furthermore, the deterrence efficacy of criminal sanctions depends critically upon law enforcement agencies possessing technical capabilities to forensically examine algorithmic systems, necessitating establishment of specialized digital forensics units equipped with reverse-engineering tools, machine learning expertise, and legal authority to compel disclosure of proprietary source code and model parameters when criminal investigations require algorithmic transparency to establish liability elements (Graph et al., 2019).

The implementation of specific deterrence in AI-enabled crimes requires a more sophisticated approach by utilizing the provisions of Articles 83 to 86 of Law Number 1 of 2023 concerning the Criminal Code, which emphasizes the goal of punishment that balances between retribution, rehabilitation, and restorative justice. Criminal sanctions for AI-based digital crimes are not only in the form of imprisonment and fines as stipulated in Articles 45 to 45B of Law Number 19 of 2016 concerning ITE, but must also include additional crimes in the form of technology ban orders, mandatory algorithm audits, and public disclosure of AI system vulnerabilities that aim to prevent recidivism and increase public awareness of the risks of AI misuse (Csongor, 2020). The effectiveness of deterrence also depends on strengthening the capacity of law enforcement agencies in conducting digital forensics and cross-border cooperation as stipulated in Article 43 of the ITE Law, considering that AI-enabled crimes are often transnational and require international legal assistance in the investigation and prosecution process (Velasco, 2022).

5. Recommendations for Criminal Law Reform to Face the AI Era

The urgency of substantive criminal law reform in Indonesia to accommodate the complexity of digital crime in the AI era requires a comprehensive legislative approach that considers several fundamental dimensions. First, it is necessary to amend Law Number 19 of 2016 concerning ITE to include specific provisions regarding AI-generated content and autonomous digital actions, by expanding the definition in Article 1 to include the concept of AI agents, algorithmic decision-making systems, and synthetic media as entities that can give rise to criminal liability for operators and developers (Arifin, 2025). Second, it is necessary to establish implementing regulations from Law Number 27 of 2022 concerning Personal Data Protection which explicitly regulates the liability framework for AI systems that carry out unauthorized data processing, automated profiling, or discriminatory algorithmic decisions that are detrimental to individual rights (Junaidi, 2024).

Third, legal reform must be accompanied by the establishment of a specialized task force for digital forensics and AI crime investigation equipped with technical expertise and legal authority to conduct comprehensive audits of AI systems suspected of being involved in criminal acts, including authority to access source codes, training datasets, and decision-making logs as digital evidence (Arifin, 2025). Fourth, Indonesia needs to adopt a risk-based regulatory approach as applied in the EU AI Act by categorizing AI applications based on their risk level and establishing differentiated liability standards and compliance requirements for each category, so that high-risk applications such as AI in criminal justice, healthcare, and financial services are subject to stricter liability standards compared to low-risk applications (Fiorinelli et al., 2025). Fifth, it is necessary to harmonize national regulations with international instruments in cybercrime governance through the ratification and implementation of the Budapest Convention on Cybercrime as well as active participation in multilateral cooperation frameworks for combating transnational AI-enabled crimes (Velasco, 2022).

C. CONCLUSION

Effective criminal liability standards for AI-era digital offenses necessitate a hybrid accountability model integrating strict liability for high-risk applications, negligence-based frameworks emphasizing due diligence obligations, and adapted vicarious liability doctrines for corporate AI deployment. Indonesia's criminal law framework manifests critical regulatory lacunae in addressing autonomous AI systems as criminal instrumentalities, generating accountability gaps that undermine legal certainty. Optimal deterrence transcends traditional punitive paradigms, requiring comprehensive preventive mechanisms including technology prohibition orders, mandatory algorithmic audits, and public vulnerability disclosure mandates. Legal certainty for perpetrators and victims demands risk-stratified liability standards, enhanced digital forensics capabilities, and harmonized international cooperation frameworks.

REFERENCES

Arifin, Z. (2025). *Criminal Liability in Press Law and Artificial Intelligence-*

Generated Disinformation : Urgency of Reform in Indonesia. *8*(3), 419–446.

Chamberlain, J. (2023). *The Risk-Based Approach of the European Union ' s Proposed Artificial Intelligence Regulation : Some Comments from a Tort Law Perspective*. 1–13. https://doi.org/10.1017/err.2022.38

Chen, X. (2022). *Di ff erent Shades of Norms : Comparing the Approaches of the EU and ASEAN to Cyber Governance*. *57*(3), 48–65.

Csongor, H. (2020). *Artificial Intelligence In Cybersecurity: Examining Liability, Crime Dynamics, And Preventive Strategies*. *8*, 730–747.

Daraojimba, R. E., Farayola, O. A., Olatoye, F. O., Mhlongo, N., Oke, T. T., Technology, F., Houston, I. R., Power, C., Africa, S., & Marketing, Y. (2023). *Forensic Accounting In The Digital Age: A U.S. Perspective: Scrutinizing Methods And Challenges In Digital Financial Fraud Prevention*. *5*(11), 342–360. https://doi.org/10.51594/farj.v5i11.614

Dwiandari, A. S., & Arifin, R. (2025). *Criminal Law Enforcement on Digital Identity Misuse in AI Era for Commercial Interests in Indonesia*. *7*(1), 37–66.

Felix, A. O., Olabode, O. J., & Ayeni, J. K. (2023). *The Criminalization Of The Internet And Cybercrime In General: A Comprehensive Study*.

Fiorinelli, G., Zucca, M. V., Zucca, M. V., & Fiorinelli, G. (2025). *Regulating AI to Combat Tech-Crimes : Fighting the Misuse of Generative AI for Cyber Attacks and Digital Offenses*. *2023*.

Floridi, L. (2019). Establishing the rules for building trustworthy AI. *Nature Machine Intelligence*, *1*(June), 261–262. https://doi.org/10.1038/s42256-019-0055-y

Floridi, L., & Cowls, J. (2019). *A Uni?ed Framework of Five Principles for AI in*. *1*, 1–15. https://doi.org/10.1162/99608f92.8cd550d1

Floridi, L., Cowls, J., Beltrametti, M., & Chatila, R. (2018). AI4People — An Ethical Framework for a Good AI Society : *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Gilbert, C., & Gilbert, M. A. (2024). *The Security Implications of Artificial Intelligence (AI) -Powered Autonomous Weapons : Policy Recommendations for International Regulation*. *9*(4), 205–219.

Gojali, D. S., & Mangkurat, U. L. (2023). *Identifying the Prevalence of Cybercrime in Indonesian Corporations : A Corporate Legislation Perspective Ten Riskiest Countries in terms of Cybercrime Rate*. *17*(1), 1–11. https://doi.org/10.5281/zenodo.4766600

Graph, K., Hu, R., Li, Z., Li, J., Hopkins, J. K., Spranklin, B. W., & Gupta, S. K. (2019). *Technology Framework of the Intelligent Command and Control System Technology Framework of the Intelligent Command and Control System*. https://doi.org/10.1088/1757-899X/677/4/042099

Hildebrandt, M. (2013). *Data Protection by Design and Technology Neutral Law*.

Hildebrandt, M., & Koops, B. (2025). *The challenges of ambient law and legal protection in the profiling era The Challenges of Ambient Law and Legal Protection in the Pro ¢ ling Era*.

Inquiries, T., Publication, L., Version, D., Self, I., Agnostic, F., Learning, A. M., & Inquiries, T. (2019). *Privacy as Protection of the Incomputable Self : From Agnostic to Agonistic Machine Learning*.

Junaidi. (2024). *Legal Reform Of Artifical Intielligences Liability To Personal Data Perpectives Of Progressive Legal Theory*.

KAN, C. H. (2024). *Criminal Liability Of Artificial Intelligence From The Perspective Of Criminal Law: An Evaluation In The Context Of The General Theory Of Crime And Fundamental Principles*. 276–313.

Kumar, P. (2023). *Determination of Civil and Criminal liability of Artificial intelligence*. https://doi.org/10.53361/dmejl.v4i01.06

Lagioia, F. (2020). *AI Systems Under Criminal Law : a Legal Analysis and a Regulatory Perspective*. 433–465.

Liability, C., From, A., Use, T. H. E., & Artificial, O. F. (2025). *CRIMINAL LIABILITY ARISING FROM THE USE OF ARTIFICIAL INTELLIGENCE*. *10*(1), 11–23.

Longpre, S., Kapoor, S., Klyman, K., Ramaswami, A., Bommasani, R., Blili-hamelin, B., Huang, Y., Skowron, A., Yong, Z., Kotha, S., Zeng, Y., Shi, W., Yang, X., Southen, R., Robey, A., Chao, P., Yang, D., Jia, R., Kang, D., … Henderson, P. (2024). *A Safe Harbor for AI Evaluation and Red Teaming*. 1–19.

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). *The ethics of algorithms : Mapping the debate*. *December*, 1–21. https://doi.org/10.1177/2053951716679679

Mohamed Fathi Shehta Diab. (2024). *Criminal Liability for Artificial Intelligence and Autonomous Systems*.

Mulyono, F. I. (2025). *Permasalahan Pemutusan Hubungan Kerja Melalui ARTIFICIAL Intelligence (Ai)*. *5*(1).

Nadya, P., Putri, T., Agung, A., Ngurah, A., & Rusmini, T. (2025). *Pengaturan Hukum dalam Penanggulangan Deepfake Artificial Intelligence (AI) terhadap Anak sebagai Korban Kejahatan Siber di Indonesia*. 713–728. https://doi.org/10.24843/JMHU.2025.v14.i03.p09

Nerantzi, E. (2024). *'Hard AI Crime': The Deterrence Turn*. *44*(3), 673–701.

Nicolas Petit. (2017). *Law And Regulation Of Artificial Intelligence And Robots: Conceptual Framework And Normative Implications*. *March*, 1–31.

Nuhi, M. H. (2024). *Pembaharuan Hukum Penanganan Tindak Pidana Pemalsuan Identitas Akibat Penyalahgunaan Artificial Intelligence Di Indonesia*. *1*.

Putra, T. H. (2024). *Law Enforcement Against Cyber Crime In Electronic Transactions In Indonesia*. *December*, 37–45.

Putri, W. M. J. (2025). *Urgensi Pengaturan Terhadap Penyalahgunaan Artificial Intelligence Pada Tindak Pidana Malware Di Indonesia*. *10*(3).

Selbst, A. D., & Vertesi, J. (2019). *Fairness and Abstraction in Sociotechnical Systems*. 59–68. https://doi.org/10.1145/3287560.3287598

Smejkal, V. (2024). *Challenges and Solutions to Criminal Liability for the Actions of*

Robots and AI. *9*(1), 65–84.

Tentang, U.-U. R. I. N. 19 T. 2016, & Perubahan Atas Undang-Undang Nomor I1 Tahun 2008. (2016). *Tentang Informasi Dan Transaksi Elektronika*.

Truby, J., Dean, R., & Ibrahim, I. A. (2022). *A Sandbox Approach to Regulating High-Risk Artificial Intelligence Applications*. *December 2020*, 270–294. https://doi.org/10.1017/err.2021.52

Undang-Undang Republik Indonesia Nomor 1 Tahun 2023 Tentang Kitab Undang-Undang Hukum Pidana. (2023). *16100*.

Velasco, C. (2022). Cybercrime and Artificial Intelligence . An overview of the work of international organizations on criminal justice and the international applicable instruments. *ERA Forum*, 109–126. https://doi.org/10.1007/s12027-022-00702-z