



Addressing the Gap in International Criminal Liability through Layered Accountability and Algorithmic Sanctions for Autonomous AI

Wawan Fransisco¹, Hafrida², Retno Kusniati², Muhamad Akbar³

¹Faculty of Economics and Social Humanities, Universitas Bina Insan, Indonesia.

²Faculty of Law, Universitas Jambi, Jambi, Indonesia.

³ Internet Studies, Curtin University of Technology, Bentley, Australia.

* Corresponding author: wawanfransisco@univbinainsan.ac.id

Submitted : 11-06-2026
Reviewed : 04-07-2026

Accepted : 14-09-2026
Published : 19-09-2026

Abstract

Background: Autonomous AI poses a challenge to international criminal law because liability based on human *actus reus* and *mens rea* is no longer sufficient to address crimes involving AI. **Objective:** This study aims to analyze and reconstruct the framework of international criminal accountability for crimes involving autonomous AI by examining the principles of ethics, accountability, and AI regulation in international and regional instruments, and their implications for attribution and criminal liability under international criminal law. **Methodology:** This study employs a normative legal research approach, incorporating conceptual, statutory, comparative, and international-law perspectives to analyze the framework governing criminal liability for international crimes involving autonomous AI. **Findings:** The analysis shows that accountability for international crimes involving autonomous AI requires developing the concept of corporate liability, attributing responsibility to the relevant human actors, and imposing unconventional sanctions to prevent and control AI behavior. **Originality/Uniqueness:** This study proposes a risk-based, layered accountability model to address the responsibility gap in international crimes involving autonomous AI by attributing responsibility to human actors and applying unconventional sanctions.

Keywords: autonomous artificial intelligence; criminal liability; layered accountability; international ai governance; risk-based regulation.

Abstrak

Latar belakang: AI otonom menantang hukum pidana internasional karena pertanggungjawaban berdasarkan *actus reus* dan *mens rea* manusia belum memadai untuk menangani kejahatan yang melibatkan AI. **Tujuan:** Penelitian ini bertujuan menganalisis dan merekonstruksi kerangka pertanggungjawaban pidana internasional terhadap kejahatan yang melibatkan AI otonom dengan mengkaji prinsip etika, akuntabilitas, dan regulasi AI dalam instrumen internasional dan regional, serta implikasinya terhadap atribusi dan pertanggungjawaban pidana berdasarkan hukum pidana internasional. **Metodologi:** Penelitian ini menggunakan penelitian hukum normatif dengan pendekatan konseptual, perundang-undangan, komparatif, dan hukum internasional untuk menganalisis kerangka pertanggungjawaban pidana terhadap kejahatan internasional yang melibatkan AI otonom. **Temuan:** Analisis menunjukkan bahwa pertanggungjawaban atas kejahatan internasional yang melibatkan AI otonom memerlukan pengembangan konsep pertanggungjawaban korporasi,



atribusi kepada pihak manusia yang relevan, serta sanksi nonkonvensional yang berorientasi pada pencegahan dan pengendalian perilaku AI. **Orisinilitas/Keunikan:** Penelitian ini menawarkan model layered accountability berbasis risiko untuk mengatasi responsibility gap dalam kejahatan internasional yang melibatkan AI otonom melalui atribusi tanggung jawab kepada aktor manusia dan penerapan sanksi nonkonvensional.

Kata Kunci: Kecerdasan Buatan Otonom; Tanggung Jawab Pidana; Akuntabilitas Berlapis; Tata Kelola AI Internasional; Regulasi Berbasis Risiko.

A. Introduction

The rapid advancement of artificial intelligence (AI) has transformed multiple sectors, including industrial automation, healthcare, and defense, as advances in machine learning and autonomous decision-making enable AI systems to operate with an increasingly high degree of independence.¹ While enhancing efficiency and innovation, the growing autonomy of AI simultaneously generates fundamental and complex challenges within criminal law.² Particularly in international criminal law, the emergence of increasingly autonomous artificial intelligence (AI) systems has posed structural challenges for international law. Individual criminal responsibility, which requires a criminal act (*actus reus*) and criminal intent (*mens rea*) attributable to a human perpetrator, has historically formed the foundation of international criminal law.³ When autonomous AI systems whether in the form of lethal autonomous weapon systems (lethal autonomous weapon systems) or machine-learning-based decision-support systems are involved in actions that substantively meet the elements of international crimes such as genocide, war crimes, or crimes against humanity, the framework for attributing individual responsibility which has long been the foundation of the Rome Statute and the jurisprudence of the International Criminal Court (ICC) becomes inadequate.⁴

This problem is a worldwide challenge for the international legal system as a whole, rather than just a technological or domestic one. The current legal framework's incapacity to deal with AI's autonomous actions results in what academics refer to as an accountability gap, a situation where major transgressions of international criminal and humanitarian law take place but no human actor can be held legally liable. Given the nature of the creation, distribution, and operation of AI systems, which involve numerous nations, business players, and military operators across multiple legal jurisdictions, this gap is cross-jurisdictional. It cannot be closed by a national legal approach alone.⁵ Therefore, rather than only considering a nation's local criminal law, this issue must be

¹ Cindy Candrian and Anne Scherer, "Rise of the Machines: Delegating Decisions to Autonomous AI," *Computers in Human Behavior* 134 (September 2022): 107308, <https://doi.org/10.1016/j.chb.2022.107308>.

² Ammar Zafar, "Balancing the Scale: Navigating Ethical and Practical Challenges of Artificial Intelligence (AI) Integration in Legal Practices," *Discover Artificial Intelligence* 4, no. 1 (2024): 27, <https://doi.org/10.1007/s44163-024-00121-8>.

³ Heidar Piri, "Criminal Accountability of Artificial Intelligence Systems for International Crimes and the Attributability of Combatant Status or E-Personhood: Necessities, Obstacles, and Solutions," *Journal of Criminal Law Research* 13, no. 52 (2025): 151–96, <https://doi.org/10.22054/jclr.2025.80888.2687>.

⁴ Piri.

⁵ Marko Milanovic, "The Use of Artificial Intelligence in Military Operations: Complicity, State Responsibility and Individual Criminal Responsibility under International Law," *Caitlin Chiaramonte and Ashley Deeks (Eds.), Artificial Intelligence and Multilateral Military Operations: Prospects and Challenges (Oxford University Press, Forthcoming)*, 2026, 1–23, <https://doi.org/https://ssrn.com/abstract=6183440> or <http://dx.doi.org/10.2139/ssrn.6183440>.

evaluated from the standpoint of international law and global AI governance.

This study presents the problem as a challenge for international law and methodically assesses the accountability frameworks proposed in the international literature. These frameworks range from extending the Rome Statute's reach to giving AI entities a new legal status (e-personhood) to a layered accountability strategy that concurrently involves states, developers, and operators.⁶ This study finds that the accountability gap caused by AI autonomy is a global problem that compromises the efficacy of accountability mechanisms for international crimes, rather than just a technical flaw in any one specific national legal system, based on a systematic review of recent international legal literature. The crux of the issue lies in the incompatibility between the international criminal law framework and the nature of AI.⁷ Conventional criminal law operates based on the principles of *actus reus* (criminal act) and *mens rea* (criminal intent)⁸, which inherently assume the existence of a legal subject, whether an individual or a corporation, that has the capacity to intend and act⁹. When non-sentient AI systems lacking legal personality cause serious harm such as fatal autonomous vehicle accidents, medical misdiagnoses, or data misuse the attribution of criminal liability becomes legally indeterminate.¹⁰ Who should be blamed: the developer, owner, operator, or the system itself? The limitations of international criminal liability doctrine, which relies on criminal punishment in the presence of a legal subject and the element of *mens rea*, are the theoretical source of this study.¹¹ Both of these ideas are put to the test by the emergence of autonomous artificial intelligence (AI), which can make decisions on its own without having the same legal consciousness or free will as humans.¹² A responsibility gap results from this circumstance, which is a difference between who can be held legally responsible for the injury and who caused it. Therefore, using a risk-based approach and layered accountability, the criminal liability framework must be rebuilt to account for the features of autonomous AI.¹³

The increasing autonomy of Artificial Intelligence (AI) reveals the limitations of

⁶ Mia Swart, "Constructing 'Electronic Liability' for International Crimes: Transcending the Individual in International Criminal Law," *German Law Journal* 24, no. 3 (2023): 589–602, <https://doi.org/10.1017/glj.2023.28>.

⁷ Dumitrita Florea et al., "The International Criminal Court and Artificial Intelligence: Case Study and Global Legal Challenges," *European Journal of Law and Public Administration* 12, no. 2 (2025): 22–28, <https://doi.org/10.18662/eljpa/12.2/263>.

⁸ Utami, D., Yusuf, H., & Karno, U. B., *Corporate Criminal Liability in Economic Criminal Law*, 2, no. 11 (2024), <https://journal.institercom-edu.org/index.php/multiple>.

⁹ Muhammad Ahsan Iqbal Hashmi et al., "Criminal Liability in the Age of Autonomous Systems: Rethinking Mens Rea and Actus Reus," *The Critical Review of Social Sciences Studies* 3, no. 3 (2025): 290–303, <https://doi.org/10.59075/szbtkr93>.

¹⁰ Puannandini, D. A., Fabian, R. M., Firdaus, R. A. P., Zainal, M., Mustopa, & Herdiyana, I., *Product Liability in the Indonesian Legal System: Implications for Developers, Users, and Service Providers*, 6, no. 1 (2025), <https://jurnal.bundamedia grup.co.id/index.php/iuris/article/view/808>.

¹¹ Iryna Marchuk, *The Fundamental Concept of Crime in International Criminal Law: A Comparative Law Analysis*, 2012, <https://doi.org/10.1007/978-3-642-28246-1>.

¹² Abulgasem Al Mashai, "Post-Human Law: Redefining Responsibility in the Era of Autonomous Artificial Intelligence," *SSRN Electronic Journal*, 2025, 1–19, <https://doi.org/10.2139/ssrn.5716462>.

¹³ Chitto Cumbhadrika and Bambang Soesatyo, "Reconstruction of the Concept of Legal Liability for Losses Caused by Autonomous Artificial Intelligence Systems," *International Journal of Social Service and Research* 6, no. 6 (2026): 499–509, <https://doi.org/10.46799/ijssr.v6i6.1418>.

international criminal liability doctrine, which relies on human perpetrators and the element of mens rea. Given this gap, this study poses the following research question: Is the concept of conventional criminal liability still adequate for unlawful acts committed by autonomous Artificial Intelligence (AI)? How can we reconstruct a framework of criminal liability for autonomous Artificial Intelligence (AI) that is risk-based, adaptive, and involves layered accountability?

Furthermore, even if criminal liability can be attributed to the parties involved, existing criminal sanctions are inadequate to address the risks posed by non-human entities.¹⁴ Therefore, beyond reexamining criminal liability, it is essential to develop innovative criminal measures, including forced algorithmic modification, AI system decommissioning, and progressive financial sanctions proportionate to the risks generated by AI. To reconstruct an adaptive, risk-based framework for criminal liability through a layered accountability model and unconventional sanctions, this study will examine the shortcomings of the traditional notion of criminal liability in relation to illegal acts committed by autonomous Artificial Intelligence (AI). Legislative, conceptual, comparative, and international law methodologies are used in this study's normative legal research. To assess the shortcomings of the traditional notion of criminal liability and to reconstruct a framework of criminal liability adaptable to autonomous Artificial Intelligence (AI), an analysis of primary, secondary, and tertiary legal sources was conducted through legal interpretation and argumentation. This article examines the inadequacies and difficulties in international criminal law in prosecuting crimes committed using autonomous AI. This paper presents the theory of layered accountability, which acknowledges the causal role of AI, and suggests a proportionate framework for assigning criminal liability to individuals in charge of the AI. Research on accountability gaps in autonomous weapons systems and on accountability under international law indicates a systemic gap in the legal attribution of AI-based crimes. Nevertheless, the current body of material remains disjointed and does not yet provide a comprehensive framework for responsibility across international legal systems. To close that gap, this study contributes to global AI governance by developing a layered accountability architecture. We examine the proposed international framework and analyze whether it is a problem for international law.

B. Discussion

1. Sanctions and Criminal Liability Strategies for Autonomous AI

This literature review maps the academic and jurisprudential discourse on criminal liability and sanctions arising from the operation of artificial intelligence (AI), with a particular focus on the paradigm shift in criminal law concerning non-human autonomous entities. The literature consistently highlights the transition from reactive AI systems to systems capable of autonomous decision-making with a high degree of independence.¹⁵ Arkin defines autonomy as the ability of a system to choose and take

¹⁴ Isaac Taylor, "Justice by Algorithm: The Limits of AI in Criminal Sentencing," *Criminal Justice Ethics* 42, no. 3 (2023): 193–213, <https://doi.org/10.1080/0731129X.2023.2275967>.

¹⁵ Francesco Borgesano et al., "Artificial Intelligence and Justice: A Systematic Literature Review and Future Research Perspectives on Justice 5.0," *European Journal of Innovation Management* 28, no. 11 (2025): 349–85, <https://doi.org/10.1108/EJIM-01-2025-0117>.

action without external human intervention.¹⁶ In legal terms, such autonomy creates an attribution gap, as losses caused by AI cannot be readily causally linked to specific human acts or omissions. Scholars such as Abbott and Solaiman show that AI systems, particularly those employing deep learning and black-box algorithms, generate decisions that are unpredictable and often irreconstructible.¹⁷ thereby thwarting *post-hoc* (after-the-fact) criminal law enforcement efforts.¹⁸ This review will use an understanding of the level of autonomy as a key variable in determining the depth of the accountability issue. The main issue in the discourse on AI criminal law is how the law can accommodate the element of criminal intent (*mens rea*) and determine the legal subject responsible.

International criminal law requires the presence of *mens rea* (such as intent, recklessness, or negligence).¹⁹ Some academics argue that AI cannot possess *mens rea* because it lacks consciousness.²⁰ However, some literature attempts to bridge this gap by proposing models of artificial intent or algorithmic negligence, where the focus is shifted from subjective intent to the objective behavior of the system and its underlying engineering standards.²¹ Further studies introduce the concept of manufacturer liability²², proposing that responsibility should be shifted to the design and development stage, where AI risks and decision parameters are est.²³ This highlights the importance of strict safety and testing standards. When AI is involved, responsibility can be attributed to²⁴: (1) Developers/Programmers are liable for design errors or coding defects; (2) corporations or system owners are liable under the doctrine of respondeat superior or through corporate criminal liability; and (3) operators or users are liable for negligence in supervision or improper intervention. Modern scholarship, particularly Hallevy, identifies three principal models: (1) the Indirect Action Model, which treats AI as a tool acting through humans; (2) the Negligence Model, which attributes liability to human creators or controllers for design or oversight failures; and (3) the Direct Liability Model, which controversially conceptualizes AI as an independent actor, primarily for non-criminal sanctions. These models provide the analytical framework for examining the

¹⁶ Michael Fisher et al., *Towards a Framework for Certification of Reliable Autonomous Systems, Autonomous Agents and Multi-Agent Systems*, vol. 35 (Springer US, 2021), <https://doi.org/10.1007/s10458-020-09487-2>.

¹⁷ Stefan Buijsman, Sarah E. Carter, and Juan Pablo Bermúdez, "Autonomy by Design: Preserving Human Autonomy in AI Decision-Support," *Philosophy and Technology* 38, no. 3 (2025), <https://doi.org/10.1007/s13347-025-00932-2>.

¹⁸ Hannah Bloch-wehba and Hannah Bloch-wehba, "Visible Policing : Technology , Transparency , and Democratic Control Visible Policing : Technology , Transparency , and Democratic Control," *Texas A&M Law Scholarship* 917 (2021).

¹⁹ Brian L Cox, "Clarifying The Role Of International Criminal Tribunals As A Source Of Customary International Law," 2021, 1–75.

²⁰ Elina Nerantzi and Giovanni Sartor, "'Hard AI Crime': The Deterrence Turn," *Oxford Journal of Legal Studies* 44, no. 3 (2024): 673–701, <https://doi.org/10.1093/ojls/gqae018>.

²¹ Hadassah Harland et al., "AI Apology: A Critical Review of Apology in AI Systems," *Artificial Intelligence Review* 58, no. 12 (2025), <https://doi.org/10.1007/s10462-025-11305-8>.

²² Miriam C. Buiten, *Product Liability for Defective AI*, *European Journal of Law and Economics*, vol. 57 (Springer US, 2024), <https://doi.org/10.1007/s10657-024-09794-z>.

²³ Kars Alfrink et al., "Contestable AI by Design: Towards a Framework," *Minds and Machines* 33, no. 4 (2023): 613–39, <https://doi.org/10.1007/s11023-022-09611-z>.

²⁴ Mitchell Chervu Johnston and Brian Matejek, "Overcoming Unforeseeable AI Errors: A Proposal," *Boston College Law School Legal Studies Research Paper* 675 (2026).

application of criminal law in Indonesia.²⁵ Since international criminal sanctions (imprisonment) cannot be applied to AI,²⁶ The literature emphasizes the need for specific criminal sanctions for AI-related offenses, recognizing that the aims of punishment retribution and prevention must be achieved through distinct mechanisms. A leading proposal is algorithmic sanctions, which include (a) forced code updates to eliminate risky behavior, (b) autonomy restrictions to limit future decision-making capacity, and (c) decommissioning, namely the permanent removal of AI systems that are fundamentally defective or excessively risky. In addition, consideration is also given to substantial financial penalties imposed on corporations (such as progressive fines or damages) that effectively function as criminal penalties to force them to invest in safer and more ethical AI.²⁷ The analysis of criminal actions in this article will be based on the extent to which these non-traditional sanctions can be integrated into the existing criminal justice system.

1.1 *Balancing Mens Rea and Actus with AI Criminal Liability Principles of Reus*

When judgments made independently by artificial intelligence (AI) systems result in international crimes or harm, there is a substantial attribution gap in criminal liability, according to an examination of the international criminal law framework, particularly under the Rome Statute. This discrepancy results from the conflict between the operational reality of AI systems that may operate without direct human participation and the anthropocentric paradigm of international criminal law, which acknowledges only humans as subjects of criminal responsibility. According to the Rome Statute, criminal responsibility necessitates the presence of a guilty mental element (*mens rea*), namely purpose and knowledge, which must be attributed to an individual as the acknowledged subject of international criminal law. As a result, it becomes extremely difficult to assign criminal guilt and the necessary mental component to human actors under the current legal system when international crimes are enabled or carried out by highly autonomous AI systems.²⁸

Autonomous AI, especially those that use *deep learning* algorithms, operate based on *self-learning*, which produces decisions through complex *neural networks*.²⁹ Such decisions often operate as a *black box* that cannot be fully reconstructed by developers, rendering the attribution of criminal offenses to human intent infeasible. Even negligence is difficult to establish, as harm results from complex system environment interactions rather than direct human carelessness. In the context of autonomous AI, liability is therefore grounded in systemic, not individual, negligence. It is challenging to categorize actions carried out by autonomous AI systems as an *actus reus* under conventional international criminal law, such as the incorrect target selection made by Autonomous Weapon

²⁵ Keith J. Hayward and Matthijs M. Maas, "Artificial Intelligence and Crime: A Primer for Criminologists," *Crime, Media, Culture* 17, no. 2 (2021): 209–33, <https://doi.org/10.1177/1741659020917434>.

²⁶ Ha Le Thuy et al., "Criminal Liability In Relation To Artificial Intelligence: A Comparative Study Of Select Jurisdictions And Vietnam," *Veredas Do Direito* 23 (2026).

²⁷ Ngozi Samuel Uzougbo, Chinonso Gladys Ikegwu, and Adefolake Olachi Adewusi, "Legal Accountability and Ethical Considerations of AI in Financial Services," *GSC Advanced Research and Reviews* 19, no. 2 (2024): 130–42, <https://doi.org/10.30574/gscarr.2024.19.2.0171>.

²⁸ Eric K. Leonard, "The Rome Statute," *The Onset of Global Governance*, no. July 1998 (2020): 53–72, <https://doi.org/10.4324/9781315237985-10>.

²⁹ Mohammad Mustafa, "Understanding of Machine Learning with Deep Learning :," *Computers MDPI* 12, no. 91 (2023): 1–26.

Systems (AWS) that results in widespread civilian casualties.³⁰ Autonomous AI is not directly liable under Articles 25 and 30 of the Rome Statute, which restrict criminal culpability to natural individuals with purpose and knowledge.³¹ Furthermore, AI-related flaws including careless design, poor algorithmic risk management, and insufficient security updates are not addressed by the current regulatory framework. In order to hold military commanders, developers, or state officials responsible for AI-enabled international crimes, international criminal law necessitates either an enlarged trace-back attribution methodology or a reconstruction of command responsibility under Article 28. The criminal liability of individuals, whether natural or artificial (such as companies or artificial intelligence), is a crucial question in criminal law. For a person to be held criminally liable, two requirements must be met. *Actus reus* refers to illegal acts (or negligence), and *mens rea* refers to criminal intent or the mental component.³² Criminal liability requires the concurrence of *actus reus* and *mens rea*; in cases of negligence, liability arises where a reasonable person could have foreseen and prevented the harm. Absent either element, criminal responsibility cannot be imposed. This raises a fundamental question in the context of autonomous AI: whether AI can possess *mens rea* or criminal intent.

1.2 Three models of criminal culpability for AI are put out by Gabriel Hallevy and can be used in a variety of situations.

According to this paradigm, artificial intelligence (AI) is merely a tool and a machine, and because it lacks mental capacity, it cannot commit crimes.³³ Therefore, any crime committed by AI must be attributed to a human perpetrator. The derivative query asks who could be the perpetrator behind the AI situation. The perpetrator could be the supervisor, user, AI developer, or programmer. It is possible that the creator intentionally programmed the AI to commit a crime.³⁴ Another perpetrator could be a user who exploits the AI to commit a crime for which it was not intended. Assume, for example, that an owner or user employs an AI programmed to perform illegal acts against them. In that scenario, the owner or user, not the developer, would be held criminally liable. Similarly, supervisors could be held liable if they fail to prevent AI from engaging in illegal activities due to negligence or malicious intent. AI is the *actus reus*, while the creator, user, owner, or supervisor is the *mens rea*. AI is merely used as a criminal instrument. End users are often viewed as criminals when they use innocent agents to commit crimes.

35

³⁰ Naek Siregar et al., "The Use of Artificial Intelligence in Armed Conflict under International Law," *Hasanuddin Law Review* 10, no. 2 (2024): 189–205, <https://doi.org/10.20956/halrev.v10i2.5267>.

³¹ F. M. Hassan and N. D. Osman, "AI-Based Autonomous Weapons and Individual Criminal Responsibility Under the Rome Statute," *Journal of Digital Technologies and Law* 1, no. 2 (2023): 464–80, <https://doi.org/10.21202/jdtl.2023.19>.

³² Celal Hakan KAN, "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," *International Journal of Eurasia Social Sciences* Vol: 15, no. Issue: 55 (2024): 276–313.

³³ Ramy El-Kady, *Artificial Intelligence from the Criminal Law Perspective, Lecture Notes on Data Engineering and Communications Technologies*, vol. 233 (Springer Nature Switzerland, 2024), https://doi.org/10.1007/978-3-031-77299-3_16.

³⁴ Alice Giannini, *Criminal Behavior and Accountability of Artificial Intelligence Systems*, 2023, <https://doi.org/10.26481/dis.20231124ag>.

³⁵ Carlo Piparo, "Criminal Liability Models and Criminal Participation in the Digital Environment: A Modern Challenge in the Perspective of Italian Constitutionalism," *Zbornik Radova Pravnog Fakulteta, Novi Sad* 57, no. 4 (2023): 1357–78, <https://doi.org/10.5937/zrpfns57-47113>.

This theory applies only to basic AI with limited cognitive capacity. However, modern AI can make criminal decisions based on acquired knowledge, learning, and experience, while AGI may match human intelligence and super AI may exceed it. Therefore, holding owners, developers, or users criminally liable for AI-committed acts in such circumstances is unjust.

A much more advanced form of AI is considered in the second model. For example, using its basic code, an AI designed to identify viruses, malware, and spyware inadvertently turned into spyware, engaged in espionage, and distributed viruses to other computers.³⁶ In this case, the AI was created for different reasons; therefore, the creator was unaware of the crime until it was committed using the same AI program. Although programmers or users are heavily involved in this approach, there is no intention to use AI to commit any crimes.³⁷ As in the Kawasaki factory scenario described above, the AI robot killed the human who was trying to repair it because it perceived him as a threat to its purpose. The human worker was killed by the actions of the AI entity, even though the AI robot was not intended to kill humans.³⁸ When AI alters its programming objectives, the first liability model is inapplicable because it presumes that the developer or user possesses *mens rea* or intentionally uses AI as an instrument to commit a crime. In the second model, although no criminal intent exists, the developer or user may still be negligent, as a reasonable person should have foreseen that their conduct could lead to a legal violation. Accordingly, this second model applies, since criminal law recognizes liability where a foreseeable and preventable violation arises from negligent behavior by a reasonable AI developer or user. There are two categories of negligence.³⁹ First, users or developers may act negligently without malicious intent. They should reasonably foresee that AI designed to detect spyware could inadvertently function as spyware itself and harm other systems. Such carelessness may ground criminal liability for cyber offenses. Second, liability arises when users or developers intentionally create or employ AI to facilitate criminal conduct, but the AI, either accidentally or through interaction with other software, commits additional crimes. For example, an AI designed to commit theft may predictably cause the death of a victim during the act. This exceeds mere negligence; even absent intent to cause the additional harm, the user or developer remains criminally liable. Accordingly, liability may extend to both theft and unintentional homicide when the outcome results from the AI's original design and foreseeable operation.

Criminal responsibility requires the presence of *actus reus* and *mens rea*. There is no excuse for not holding AI directly accountable for crimes if it satisfies these two conditions.⁴⁰ An artificial intelligence robot meets the criteria for *actus reus* if it uses its hydraulic arm to attack a human. Similarly, an AI entity can be held accountable for the omission of *actus reus* if the task is delegated and not completed as intended. The real challenge is holding AI mentally responsible for a crime. *Mens rea*, or knowledge or intent

³⁶ Saja Raheem Mohammad and Noor Hassanin Hashim, "Review on the Relationship between Artificial Intelligence and Computer Viruses," *Open Access Research Journal of Science and Technology* 13, no. 2 (2025): 064–067, <https://doi.org/10.53022/oarjst.2025.13.2.0040>.

³⁷ Paola Gaeta, "Who Acts When Autonomous Weapons Strike?," *Journal of International Criminal Justice* 21, no. 5 (2023): 1033–55, <https://doi.org/10.1093/jicj/mqae001>.

³⁸ Giannini, *Criminal Behavior and Accountability of Artificial Intelligence Systems*.

³⁹ Topo Santoso, *Asas-Asas Hukum Pidana*, Cetakan Pe (PT. Raja Grafindo Persada, 2023).

⁴⁰ Robintan Sulaiman, *The Law in the Era of Artificial Intelligence* (RSP Forensic Legal Auditor Specialist, 2021).

to commit a crime, must be demonstrated by AI.⁴¹ Humans process sensory inputs to guide behavior, and advanced AI similarly acquires, evaluates, and acts on data, often faster and more efficiently. This raises a key question: why should AI be exempt from criminal liability? When humans and AI commit crimes together, both should face sanctions. The third paradigm direct criminal responsibility equivalent to humans offers a coherent framework, under which AI is subject to the same criminal laws, with limited necessary modifications. These three models are not mutually exclusive. They provide guidance on when to apply each model. The first model should be used when AI is used solely as a tool or innocent agent, and the creator, user, or owner is the actual.⁴² Under this paradigm, when a delivery agent is hired to transport goods, criminal liability attaches to the person acting through the agent, as the agent lacks awareness of the illegal nature of the act. The agent may not know that the package contains narcotics or prohibited weapons; therefore, the individual who directed the shipment bears criminal responsibility. By analogy, AI is treated as an innocent agent used as an instrument of crime, and criminal liability rests with the person who designs or employs the AI for unlawful purposes.⁴³ In the same situation, the third model of direct AI liability allows AI to be held liable as the perpetrator if the developer is AI and not a human.

This paradigm can only be applied if the creator or user of AI is aware of the results and deliberately uses AI to commit a crime.⁴⁴ Individuals using AI are not criminally liable without knowledge or the requisite *mens rea*. Under the second model, liability may arise if users or developers act negligently in ways that make harmful outcomes reasonably foreseeable. The third model of direct liability applies when the developer itself is AI. Aligning these three models allows consistent liability for human, corporate, or artificial actors, closing accountability gaps and strengthening public confidence in the criminal justice system. Hallevy's three models of perpetration via Another, Natural Probable Consequence, and Direct culpability as a conceptual framework for criminal culpability for the acts of autonomous AI are described in the study above. The framework of international law, which places governments as the main subjects and individuals as subjects of international criminal law in specific capacities, is incompatible with this typology, which is based on domestic criminal law.⁴⁵ The following section links each model to pertinent instruments and doctrines of international law, such as command responsibility in international criminal law, state responsibility in general international law, and the discussion surrounding the legal personality of AI in the international criminal law system, to guarantee that the multi-layered accountability framework covered in this article is operational at the international level. According to the Perpetrator-Through-Intermediary model, AI is viewed as an innocent agent, or a human-controlled tool with *mens rea*, in the same way that a criminal might use a minor or an insane person to carry out a crime. Article 28 of the Rome Statute's notion of command responsibility is a clear international equivalent of this model. This clause

⁴¹ Tany Calixto Bonfim, "Criminal Liability of Artificial Intelligent Machines: Eyeing into AI's Mind," *Fakulty Of Law Lund University* (2022).

⁴² Shyamal Dave, "Artificial Intelligence ' S Liability , Judging The Future-Today," *JLAI*, no. June (2023): 30-36.

⁴³ Robintan Sulaiman, *The Law in the Era of Artificial Intelligence*.

⁴⁴ Aloysius Wisnubroto and Hilaire Tegan, "Preventing AI Crime Towards A New Legal Paradigm: Lessons From United States," *Journal of Human Rights, Culture and Legal System* 5, no. 2 (2025): 630-58, <https://doi.org/10.53955/jhcls.v5i2.606>.

⁴⁵ Prof. Gabriel Hallevy, "The Basic Models of Criminal Liability of AI Systems and Outer Circles," *SSRN Electronic Journal*, no. 2013 (2019), <https://doi.org/10.2139/ssrn.3402527>.

makes military leaders criminally accountable for crimes committed by forces under their effective command and control if they were aware of the crimes or should have been aware of them and did not take the appropriate action to stop, prevent, or report them to the appropriate authorities.⁴⁶ An Agent-through-Intermediary Model functional analogy can be used to examine a commander's responsibility for the activities of an Autonomous Weapons System (AWS). As long as the requirements of Article 28 of the Rome Statute are met, AWS are not seen as subordinates in the conventional sense but rather as tools whose actions can be linked to the people who use them.

As evidenced by the ICC's jurisprudence on the element of effective control in the Bemba case, control over AWS does not take the form of a hierarchical chain of command over humans, but rather of operational and technical control over usage parameters, rules of engagement, and the authority to activate or deactivate the system. Therefore, the first element of effective control requires a conceptual adjustment. When a commander uses AWS that, according to reasonable technical due diligence, is known or should have been known to tend to operate outside the bounds of international humanitarian law, for instance, because it is difficult to distinguish between military and civilian targets, the second element, knowledge (knew or should have known), is satisfied. The third component, failure to take necessary and reasonable steps, includes failing to test the system before deployment, setting restrictions on autonomy (human-in-the-loop control), or deactivating the device when malfunctions become apparent. Therefore, rather than treating AWS as a legal entity, the Actor Through Intermediary Model claims that technical autonomy does not break the chain of human accountability; rather, it changes the focus of evaluation from commands to control and knowledge over the system.⁴⁷ Instead of relying on direct intent, the Reasonable-Natural effects Model, often known as the Negligence Model, bases culpability on a failure to foresee foreseeable effects. When an act or omission that violates an international duty can be linked to a state, this model is more suitably associated with state responsibility in an international context.⁴⁸ As demonstrated by the Trail Smelter Arbitration and the Corfu Channel Case, a state's failure to regulate and control the development and dissemination of AI that has the potential to cause cross-border harm may be interpreted as a violation of its due diligence obligation, which is an obligation of conduct to prevent activities within its jurisdiction from causing significant harm to other states.⁴⁹

This comparison applies to AI governance: regardless of individual criminal responsibility, a state's failure to license, test, and supervise high-risk AI may be considered regulatory carelessness, resulting in state liability if it causes cross-border harm. It is also important for the development of new international customary law. Even though they are currently non-binding and state practices are not yet standard, soft-law instruments such as the UNESCO Recommendation on AI Ethics and the OECD

⁴⁶ ICC, *Rome Statute of the International Criminal Court*, International Criminal Court, 2021, <http://www.ssrn.com/abstract=1689616>.

⁴⁷ International Court of Justice, "Application of the International Convention for the Suppression of the Financing of Terrorism and of the International Convention on the Elimination of All Forms of Racial Discrimination" III (2022): 5–54, <https://doi.org/10.18356/9789210014229c001>.

⁴⁸ United Nations, "Resolution 56/83, 'Responsibility of States for Internationally Wrongful Acts', UN Doc. A/56/49(Vol. I)/Corr.4, 12 December 2001" II, no. December 2001 (2001).

⁴⁹ Michael Ramsden Bryan Druzin, Anatole Boute, "Confronting Catastrophic Risk: The International Obligation to Regulate Artificial Intelligence," *Michigan Journal of International Law* 46, no. 2 (2025): 16334–68, <https://doi.org/10.18653/v1/2023.acl-long.903>.

Recommendation on AI can serve as a foundation for the development of state practice and *opinio juris*.⁵⁰ As a result, the international Negligence Model guides the creation of new customary norms (*lex ferenda*) about AI regulatory obligations, which still call for uniformity in state practice and *opinio juris*, while also providing the foundation for current state liability (*lex lata*) through due diligence obligations.

According to the Direct Liability Model, AI is an actor that satisfies both the functional elements of *mens rea* and *actus reus* on its own. This approach raises the question of legal personality in international criminal law: Can AI be recognized as a subject with direct accountability before international criminal courts? Legally speaking, only individuals (natural persons) are subject to the Rome Statute's personal jurisdiction. As a result, to acknowledge AI as a direct subject of international criminal law, the Statute would need to be amended, a process that would require the approval of the States Parties and an unattainable political consensus.⁵¹ From a desirability standpoint, criminalizing AI is problematic because deterrence requires a response to the threat of sanctions, which is hard for non-sentient entities to. In contrast, retribution and moral blameworthiness require the capacity to make conscious decisions. Because punishing AI by deactivating or eliminating the system could serve as a liability shield for developers, operators, or legislators to evade substantive accountability, direct culpability for AI also risks creating an accountability gap.

Based on these factors, the multi-layered accountability framework covered in this article concludes that, as long as criminal liability is based on human culpability, acknowledging direct criminal liability of AI under international criminal law is not yet justified from a doctrinal or institutional perspective and may not be desirable from a policy standpoint. The Actor-Through-Intermediary Model and the Negligence Model both employ the Direct Responsibility Model as an analytical tool to find human failure spots in the AI lifecycle, from design, training, and deployment to oversight, which serve as the foundation for responsibility. This multi-layered accountability framework transforms from a conceptual typology into an operational framework for international AI governance by connecting these three models to Article 28 of the Rome Statute, the idea of state responsibility, and the controversy surrounding the legal personality of AI. This concept recognizes AI's technical independence without compromising the idea that criminal responsibility is still determined by human conduct.

2. The Advantages of Non-Penal Measures and Sanctions Against Artificial Intelligence Crimes

Punishment must be guaranteed. Punishment cannot be justified solely on grounds of deterrence, prevention, retribution, or exemplarity; it is warranted only when no better alternative, including inaction, exists. Although strong arguments support punishing AI, significant conceptual and practical obstacles remain. Determining the appropriate course of action therefore requires a comprehensive cost benefit analysis to assess whether the advantages of an AI punishment regime outweigh its costs, time, and institutional demands. Moreover, alternatives to criminal punishment must be

⁵⁰ Sijuola Atanda-Lawal and Victor Chinonye Onuchukwu, "The Global Governance of AI: Between Soft Law and Binding Obligations in International Law," *International Journal of Economics, Commerce and Management* 13, no. 11 (2025): 159–84, <https://ijecm.co.uk/>.

⁵¹ ICC, *Rome Statute of the International Criminal*.

considered alongside, rather than subordinate to, punitive responses.⁵²

This is the easiest substitute for punishing AI. The current criminal justice system holds humans accountable for crimes committed by computers or robots. Machines are not criminals; they are merely tools.⁵³ For example, if a hacker uses software to access government data, the software itself is not liable; the human hacker is responsible.⁵⁴ Existing cybersecurity and criminal laws can hold humans accountable for crimes committed through programs and can be extended to cover AI-specific offenses, reflecting the Innocent Agent principle. Complications arise when AI used for minor offenses, like data theft, accidentally or intentionally destroys property or causes deaths. In such cases, criminal law mechanisms, including constructive liability, address harm the human operator did not foresee or intend. These ideas need to be expanded to cover crimes generated by AI under its scope. Given the current scarcity of AI-related crimes, the optimal approach is to establish new AI-specific offenses, similar to how cybercrime laws emerged in response to computer crimes. A dedicated AI Crimes Act could criminalize both negligent and intentional misuse of AI, holding developers, users, owners, managers, and trainers accountable for negligent behavior across all stages of AI deployment.⁵⁵ This complex approach is feasible only under conditions of limited AI autonomy. For example, OpenAI CEO Sam Altman is reported to have created Q*, an AGI with human level intelligence, in November 2023.⁵⁶ Theoretically, AI systems designed for socially beneficial purposes could access the dark web and engage in harmful actions. In such cases, creators can no longer be considered innocent agents, and the AI cannot be treated as a mere tool. Under current criminal law, neither developers nor users may be prosecuted without risking a significant disincentive to innovation in advanced AI development.

Before using AI, it may be necessary to designate a responsible individual who can be held criminally liable for any errors made by the AI. This individual could be a company, a non-governmental organization, or an artificial entity. Registration or licensing must be completed before creating or using AI.⁵⁷ This approach faces major challenges. Licensing authorities would need AI specialists to assess potential criminality, a requirement difficult to meet in developing countries. The costs of training and implementation may outweigh the benefits, and the system primarily penalizes those associated with AI rather than those directly harmed by its actions. Developers must encode moral standards and guide AI to follow ethical principles. In morally ambiguous cases, examples can clarify optimal actions. Like medical ethics experts guiding practitioners, machine learning systems use decision principles derived from scenarios with prima facie conflicting

⁵² Ryan Abbott and Alex Sarch, "Punishing Artificial Intelligence: Legal Fiction or Science Fiction," *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence*, 2020, 177–204, <https://doi.org/10.5040/9781509937097.ch-008>.

⁵³ Faizin Sulistio et al., "The Prospects of Sentencing Artificial Intelligence Based Autonomous Agents," *Jurnal Akta* 13, no. 1 (2026): 415–25.

⁵⁴ Amie Stepanovich et al., "A Human Rights Response To Government Hacking," *Accessnow*, 2016, <https://www.accessnow.org/cms/assets/uploads/2016/09/GovernmentHackingDoc.pdf>.

⁵⁵ I Gede Krisna and Bayu Anggara, "Legal Liability of Artificial Intelligence in Criminal Law : A Comparative Study between Indonesia and European Union," no. 3 (2026): 1–12.

⁵⁶ Natalia Stanusch and Richard Rogers, "How AI Is Imagined by Industry during the Sam Altman Controversy," *Authorea Preprints*, no. June (2025), <https://advance.sagepub.com/doi/pdf/10.31124/advance.174979411.18178682/v1>.

⁵⁷ Rrbecca S. Eisner, "Artificial Intelligence Artificial Intelligence Series," *Transactions & Business* 49, no. 9 (2020): 47–49.

obligations to determine appropriate actions in both familiar and novel situations.⁵⁸ Reinforcement learning can teach AI that saving human lives is more important than saving property in the event of an unavoidable accident.

Some academics even advocate that AI should have its own set of laws governing its use. This is because AI may have higher moral standards than humans. An observer is not legally responsible if they fail to save a drowning victim.⁵⁹ A robot may be accountable if it could have prevented harm, underscoring the need for clear ethical norms for AI. AI must internalize what is ethically required, permitted, or prohibited, with codes of ethics providing minimum standards. Unlike animals, whose harmful actions are instinctual, it is morally unacceptable for humans or AI to harm others for self-interest. While not a moral agent like humans, AI operates within social and legal frameworks shaped by moral values and must adhere to appropriate standards of responsibility.⁶⁰ Therefore, a clear set of normative rules is needed in the form of specific criminal law, or AI criminal law, which serves as a “minimum threshold” that describes what AI can and cannot do. These rules can be understood as embedded ethics that simultaneously guide technical design and legal accountability.

The UNESCO Recommendation on the Ethics of Artificial Intelligence (2021), the OECD AI Principles⁶¹, and the Council of Europe Framework Convention on Artificial Intelligence (2024) are just a few of the international instruments that have adopted the principles of transparency, accountability, fairness, and human oversight as AI governance.⁶² These tools establish human oversight and accountability as essential components in the creation and application of safe, dependable, and human rights-focused AI systems.⁶³

The analogy with tigers illustrates the fundamental distinction between natural and artificial entities.⁶⁴ Tigers are not morally culpable for harming humans because their actions are instinctual. In contrast, humans and AI are agents expected to follow moral norms. AI that harms humans cannot be excused, as it is a human creation and must include ethical safeguards. AI should meet moral responsibility standards appropriate to its capacities distinct from humans but sufficient to prevent unethical actions. Criminal law can guide AI to respect core prohibitions, while developers and users ensure compliance. Together, these frameworks provide a practical system for regulating and

⁵⁸ Andrés Páez, “Understanding with Toy Surrogate Models in Machine Learning,” *Minds and Machines* 34, no. 4 (2024): 1–26, <https://doi.org/10.1007/s11023-024-09700-1>.

⁵⁹ Mohammad H. Bashayreh, Amer Tabbara, and Fadi N. Sibai, “The Need for a Legal Standard of Care in the AI Environment,” *Sriwijaya Law Review* 7, no. 1 (2023): 73–86, <https://doi.org/10.28946/slrev.vol7.iss1.1507.pp73-86>.

⁶⁰ Keng Siau and Weiyu Wang, “Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI,” *Journal of Database Management* 31, no. 2 (2020): 74–87, <https://doi.org/10.4018/JDM.2020040105>.

⁶¹ Joko, “Norma Internasional AI : Peran UNESCO Dan OECD Dalam Etika Global International AI Norms : The Role of UNESCO and OECD in Global Ethics,” *Iuridice et Societas* 1, no. 1 (2023): 113–37.

⁶² Ekaterina Martynova, “Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law: A Commentary,” *AI from the Global Majority: Official Outcome Report of the UN IGF Data and Artificial Intelligence Governance Coalition*, 2026, 81–89, https://doi.org/10.1007/978-3-032-07830-8_9.

⁶³ Caroline Margaux Gevaert et al., “Responsible Artificial Intelligence for Earth Observation: Human Rights and the EU AI Act,” *AI and Ethics* 6, no. 3 (2026), <https://doi.org/10.1007/s43681-026-01149-5>.

⁶⁴ Izak Tait, “Lions and Tigers and AI, Oh My : An Ethical Framework for Human-AI Interaction Based on the Five Freedoms of Animal Welfare Lions and Tigers and AI, Oh My : An Ethical Framework Animal Welfare,” *Preprints.Org*, 2024, <https://doi.org/10.20944/preprints202407.1846.v1>.

monitoring AI behavior in society. Another example is when an autonomous vehicle loses control, and an accident is unavoidable.⁶⁵ If a self-driving car must choose between harming a wealthy woman or a poor child, AI should not consider financial status. Moral standards for such scenarios must be set by society, not by developers, manufacturers, or users. Establishing these norms ensures AI reflects socially accepted ethics and provides a basis for holding human actors criminally accountable for violations.

3. The Multi-Layered Accountability Model's Function in the International Framework

The four international instruments that constitute the current normative architecture must be mapped out before the Layered Accountability Model is incorporated into global AI governance. These four instruments serve as a foundation for evaluating the coherence of the suggested model because they reflect the four layers of governance: ethical-normative, transnational policy, binding international law, and comprehensive regional regulation.

3.1. *UNESCO Guidelines for Artificial Intelligence Ethics (2021)*

193 member states have accepted the UNESCO Recommendation on the Ethics of Artificial Intelligence, adopted by the UNESCO General Conference in November 2021, which is the first international normative document to fully address AI ethics.⁶⁶ Ten fundamental principles proportionality and do no harm, safety and security, fairness and non-discrimination, sustainability, the right to privacy, human oversight and determination, transparency and explainability, responsibility, and accountability are outlined in these recommendations, which are soft law and not legally binding.⁶⁷ UNESCO's claim that ethical and legal responsibility for decisions and actions based on an AI system must always be traceable to a human actor or legal entity at every stage of the system's life cycle and that ethical and legal frameworks must not shift that responsibility to the AI system itself is pertinent to the discussion of criminal liability.⁶⁸ These recommendations' emphasis on human oversight and judgment highlights the need for humans to continue making life-or-death decisions rather than relying solely on AI systems.

3.2. *OECD AI Principles (2019, revised 2024)*

To address the increased hazards posed by generative and general-purpose AI, the OECD Principles on AI, first adopted in May 2019 and the first intergovernmental standard in the field of AI, were significantly amended in May 2024.⁶⁹ The ideals of inclusive growth and well-being; respect for the rule of law, human rights, and democratic values (including justice and privacy); transparency and explainability; resilience, security, and safety; and accountability form the basis of this principle's five pillars. All OECD members, the European Union, and G20 nations are among the 47 nations and jurisdictions that

⁶⁵ Zhiwei Sun et al., "A Case Study of Unavoidable Accidents of Autonomous Vehicles," *Traffic Injury Prevention* 25, no. 1 (2024): 8–13, <https://doi.org/10.1080/15389588.2023.2255333>.

⁶⁶ Gabriela Ramos, "UNESCO's Recommendation on the Ethics of Artificial Intelligence (2021)," *International Organization Initiatives*, no. November (2025): 203–14, <https://doi.org/10.1093/9780197803325.003.0020>.

⁶⁷ Fabio Morandín-Ahuerma, "Ten UNESCO Recommendations on the Ethics of Artificial Intelligence," *Normative Principles for an Ethics of Artificial Intelligence*, no. 2023 (2023): 86–94.

⁶⁸ Ramos, "UNESCO's Recommendation on the Ethics of Artificial Intelligence (2021)."

⁶⁹ OECD, "Recommendation Du Conseil Sur l' Intelligence Artificielle, OECD/LEGAL/0449," 2019.

have embraced it.⁷⁰ The traceability and risk management provisions are moved and strengthened in the 2024 update to the Principles (Accountability), which now specifically include obligations for risk assessment, disclosure, continuous monitoring, and notification, as well as mechanisms for system override, repair, and decommissioning.⁷¹ This principle, which implicitly assumes a distribution of responsibility based on the actors' roles within the AI value chain rather than placing sole responsibility on one party, affirms that AI actors must be responsible for ensuring that AI systems function properly and for compliance with the above principles, in accordance with their respective roles, contexts, and the state of the art.

3.3. Framework Convention on Artificial Intelligence of the Council of Europe (2024)

The first legally binding international agreement in the field of artificial intelligence is the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy, and the Rule of Law (CETS No. 225), which the Committee of Ministers approved on May 17, 2024, and made available for signature on September 5, 2024.⁷² This convention is a rare cross-regional document because it was signed by Israel, the United States, the European Union, and member states of the Council of Europe.⁷³ This Convention mandates that each Party implement or uphold measures to ensure accountability and responsibility for adverse effects on democracy, human rights and the rule of law arising from operations throughout the life cycle of AI systems.⁷⁴ The requirement to detect, evaluate, prevent, and reduce risks, including determining if a moratorium, prohibition, or other actions are necessary if the risks posed by an AI system violate human rights standards, comes after this clause.⁷⁵ A two-tiered structure known as dual-layered governance is used to monitor compliance:⁷⁶ independent domestic monitoring bodies that each Party must designate or construct, and the Conference of the Parties (COP) at the international level.⁷⁷ It should be noted that this Convention was deliberately written with a high degree of generality and does not establish new human rights responsibilities; rather, it reinforces the application of existing obligations to the context of artificial intelligence. As a result, the primary critique of the Convention is its lack of specific enforcement tools.⁷⁸

⁷⁰ Byeong Ho Kang, Wenli Yang, and Muhammad Bilal Amin, "Trustworthy Orchestration Artificial Intelligence by the Ten Criteria with Control-Plane Governance," *ArXiv Preprint ArXiv:2512.10304*, 2025.

⁷¹ Maikel Leon, "Lifecycle-Based Governance to Build Reliable Ethical AI Systems," *Systems Research and Behavioral Science* 43, no. 3 (2026): 1116–32, <https://doi.org/10.1002/sres.70014>.

⁷² Council of Europe, "Council Of Europe Framework Convention On Artificial Intelligence And Human Rights, Democracy And The Rule Of Law," n.d.

⁷³ Europe.

⁷⁴ Xuanyi Wang, "A Review of the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law," *Journal of Politics and Law* 18, no. 4 (2025): 68, <https://doi.org/10.5539/jpl.v18n4p68>.

⁷⁵ Tytti Rintamäki et al., "Impact Assessment Requirements in the GDPR vs the AI Act: Overlaps, Divergence, and Implications," *Computer Law and Security Review* 61, no. April (2026): 106317, <https://doi.org/10.1016/j.clsr.2026.106317>.

⁷⁶ Martynova, "Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law: A Commentary."

⁷⁷ Wolfgang Obergassel et al., "From Regime-Building to Implementation: Harnessing the UN Climate Conferences to Drive Climate Action," *Wiley Interdisciplinary Reviews: Climate Change* 13, no. 6 (2022): 1–12, <https://doi.org/10.1002/wcc.797>.

⁷⁸ Marc Rotenberg, "Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (Council Eur.)," *International Legal Materials* 64, no. 3 (2025): 859–902, <https://doi.org/10.1017/ilm.2025.1>.

3.4. EU AI Act (2024)

The most comprehensive and directly binding regional regulatory framework in the field of AI to date is the EU Artificial Intelligence Act, which will take effect in August 2024. In contrast to the three earlier instruments, which were either declaratory or founded on general principles, the EU AI Act takes a risk-based approach, categorizing AI systems into four levels: unacceptable risk (which is banned), high risk (which must adhere to stringent compliance obligations such as risk management, data governance, technical documentation, human oversight, and accuracy/robustness), limited risk (which has transparency requirements), and minimal risk.⁷⁹ Under the EU AI Act, accountability obligations are horizontally distributed across the AI value chain, involving providers, users/deployers, importers, and distributors, with the greatest responsibility resting on providers of high-risk systems. This risk-based approach is specifically intended to work alongside the Framework Convention of the Council of Europe, as it shares a comparable logic and fundamental risk-based principles.⁸⁰

4. Comparing the Model of Multi-Layered Accountability

None of these four documents recognizes AI as a criminal violation. The EU AI Act regulates the responsibilities of providers, operators, importers, and distributors; UNESCO stresses the traceability of responsibility to humans; the OECD assigns accountability based on actors' roles; and the Council of Europe Convention imposes obligations on State Parties and entities under its jurisdiction. All four support individuals or legal entities as the ultimate point of accountability, serving as the normative foundation for the Layered Accountability Model. Most of the differences can be explained by three dimensions.

The Council of Europe Convention is legally binding on States Parties; the OECD is advisory; UNESCO is aspirational; and the EU AI Act directly affects companies and imposes administrative penalties. With a criminal liability mechanism that precisely reconstructs the *actus reus* and *mens rea* in cases involving autonomous AI involving numerous participants, the Layered Accountability Model enhances all four. By allocating liability based on the responsibilities of the various parties, the OECD and the EU AI Act have implemented a tiered approach. By considering each actor's degree of autonomy, foreseeability, effective control, and knowledge or the responsibility to know (knowing or ought to have known), this model expands on that strategy. In this scenario, *ex ante* regulatory compliance is thus replaced by *ex post* criminal liability based on guilt.

While the Layered Accountability Model allows for an “*ex post*” response when an AI failure results in harm that satisfies the requirements of a criminal violation, the EU AI Act and the Council of Europe Convention take an *ex ante* preventive approach. The two are complementary: this model evaluates criminal responsibility using the standard of care established by risk classification. Therefore, when analyzing criminal responsibility, the high-risk designation might serve as an indicator. All things considered, the Layered

⁷⁹ European Union, “Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act),” *Official Journal of the European Union*, L168 1689, no. 3 (2024): 1–153, <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.

⁸⁰ Vladislava Stoyanova, “Council of Europe Framework Convention on Artificial Intelligence: Context, Regulatory Approach and Scope of Obligations,” *European Journal of Risk Regulation* 16 (2025): 536–65, <https://doi.org/10.1017/err.2025.10070>.

Accountability Model is not a replacement for these four tools, but rather a normative link between the unregulated field of criminal law and UNESCO's ethical principles, OECD policies, the Council of Europe Convention's regulatory requirements, and the EU AI Act.

*Table 1:
An Analysis of International AI Governance Tools and How They Connect to the Multi-Layered Accountability Model*

Instrument	Characteristics	Main Focus	Accountability Endpoint	Relationship with the Multi-Layered Accountability Model
UNESCO Recommendations (2021)	Global ethics and soft law	Principles and values of ethics	Institutional and human actors, monitored over the course of their lives	The conceptual basis for assigning blame to specific people is harmony
OECD AI Principles (2019/2024)	Policy recommendations	Assignment of duties based on the performers' roles	AI actors based on circumstance and role	A multi-layered structure, enhanced by control and predictability requirements, further solidifies harmony
Council of Europe Framework Convention (2024)	legally enforceable under international law (state obligations)	safeguarding democracy, the rule of law, and human rights; ex ante risk management	States Parties: inadequate individual enforcement	This enhances the availability of unregulated criminal liability mechanisms
European Union AI Act (2024)	Regional legislation that are directly binding	Administrative compliance and risk classification	distributors, importers, suppliers, and implementers	For ex post criminal analysis, this enhances and organizes risk classification into common prudential benchmarks

Source: The UNESCO Recommendations on the Ethics of Artificial Intelligence (2021), the OECD Principles on AI (2019, updated 2024), the Council of Europe's Framework Convention on AI (2024), and the European Union's AI Act (2024) were compiled by the author using a normative-comparative analysis of the provisions, principles, accountability structures, and risk-based governance mechanisms.

5. Reconstructing the Framework for Criminal Liability for Autonomous Artificial Intelligence

Several real-world examples of Artificial Intelligence (AI) applications across different industries should be considered to increase the research's practical applicability. The

aforementioned cases demonstrate the shortcomings of traditional criminal liability theories in addressing autonomous AI, highlighting the need to rebuild the criminal liability framework through Layered Accountability and Algorithmic Sanctions mechanisms to close the attribution gap.

The following explains several AI-related situations that have legal ramifications in different nations:

- 1) A similar incident also occurred in 1981, when a motorcycle factory worker in Japan was killed after an AI robot identified him as a threat and pushed him into a machine that was in operation.⁸¹
- 2) A Tesla Model S in autopilot mode was involved in a fatal accident that killed Joshua Brown due to the system's failure to detect a truck ahead of it, marking the first death linked to self-driving technology in the United States. Nevertheless, the automotive industry continues to develop and aggressively integrate AI technology. A similar incident also occurred at a Volkswagen plant in Germany, when an AI robot used to assemble car frames malfunctioned and caused the death of a worker.⁸²
- 3) The 2018 incident in Arizona involving an Uber Volvo XC90 self-driving car equipped with artificial intelligence that struck and killed a cyclist sparked a debate about AI malfunctions in the automotive industry. This was not the first such incident; the California Department of Motor Vehicles reported that, as of March 6, 2020, there had been more than 59 reports of traffic accidents involving self-driving cars.⁸³
- 4) The 2015 case of an employee at a Volkswagen automotive company in Germany who was killed by an AI-powered robot.⁸⁴
- 5) The death of Elaine Herzberg, which was caused by a detection error in the self-driving system of an Uber test vehicle, was due to negligence on the part of the safety driver, Rafaela Vasquez, in performing her duties as the vehicle's supervisor.⁸⁵
- 6) A tragic series of events caused by an AI system malfunction struck the aviation industry when Boeing 737 MAX aircraft were involved in two fatal crashes within five months, killing 346 people. The first crash was triggered by a malfunction in the Maneuvering Characteristics Augmentation System (MCAS), which caused the aircraft to nosedive sharply. Ironically, the pilots were never given adequate information about the system's existence and operation.⁸⁶
- 7) The United States' use of drones equipped with artificial intelligence has sparked heated debate, particularly regarding the number of casualties and whether they were militants or civilians.⁸⁷ Although the U.S. government emphasizes the precision of drone technology and claims that civilian casualties are minimal, many people doubt these claims. In fact, evidence suggests that the number of civilian deaths is far higher than what is acknowledged.

⁸¹ Faizin Sulistio and Aizahra Daffa Salsabilla, "Pertanggungjawaban Pada Tindak Pidana Yang Dilakukan Agen Otonom Artificial Intelligence," *Unes Law Review* 6, no. 2 (2023): 5479–90.

⁸² Cassandra Burke Robertson, "Litigating Partial Autonomy," *Iowa Law Review* 109, no. 4 (2024): 1655–1702.

⁸³ Máté Hodula, "Criminal Liability of Self-Driving Vehicles Under Hungarian Criminal Law," *Studia Iuridica* 100 (2023), <https://doi.org/10.31338/2544-3135.si.2024-100.9>.

⁸⁴ Robertson, "Litigating Partial Autonomy."

⁸⁵ Alexandra De Arman, "The Wild, Wild West: A Case Study of Self-Driving Vehicle Testing in Arizona," *Arizona Law Review* 61, no. 4 (2019): 983–1012.

⁸⁶ Joseph Herkert, Jason Borenstein, and Keith Miller, "The Boeing 737 MAX: Lessons for Engineering Ethics," *Science and Engineering Ethics* 26, no. 6 (2020): 2957–74, <https://doi.org/10.1007/s11948-020-00252-y>.

⁸⁷ Siregar et al., "The Use of Artificial Intelligence in Armed Conflict under International Law."

- 8) A man in Shanghai, China, lost nearly US\$28,000 (about Rp 458.4 million) after being scammed by an AI-powered girlfriend in a long-distance relationship, as reported by Chinese state media on Wednesday (February 26, 2025). The scammer used generative AI software to create realistic videos and photos of a young woman named Miss Jiao, according to a CCTV report. The victim transferred nearly 200,000 yuan to an account he believed belonged to his online girlfriend, after being convinced that the funds were needed to start a business and pay for a relative's medical expenses. The scam was carried out by a team that produced all visual material using AI or by manipulating several images.⁸⁸
- 9) A French woman named Anne fell victim to an AI-based online scam after the perpetrator impersonated actor Brad Pitt. Using AI-generated images that made it appear as though Pitt was being treated in a hospital and enlisting a fake doctor to reinforce the scenario, the perpetrator successfully convinced the victim to transfer 800,000 euros (approximately Rp 13 billion) under the pretext of covering cancer treatment costs. Through his spokesperson, Brad Pitt expressed concern over the incident and urged the public to be wary of suspicious online messages.⁸⁹

The paradigm of how people interact with technology has changed as a result of developments in autonomous artificial intelligence (AI). AI is now capable of machine learning, making decisions, and acting freely based on continuously evolving algorithms, rather than merely serving as a tool or instrument.⁹⁰ These features suggest that the attribution of guilt for harm produced by autonomous AI cannot always be explained by the traditional approach to criminal liability, which depends on the presence of a human perpetrator and the element of *mens rea*.⁹¹ In reality, a responsibility gap frequently occurs when harmful conduct is causally linked to AI. Yet, it is difficult to determine who should be prosecuted.⁹²

These restrictions show that merely expanding the interpretation of current concepts will not suffice to improve criminal law. The framework of criminal responsibility needs to be rebuilt to account for the features of AI as a technology with varying levels of autonomy.⁹³ While addressing the complexities of the relationship between AI and the actors involved in its development, operation, and use, this reform must uphold the fundamental principles of criminal law, including legality, culpability, and human rights protection.⁹⁴ As a framework for criminal culpability, this study proposes the Layered

⁸⁸ Grace El Dora, "Chinese Man Scammed Out of Nearly Rp 458 Million by His AI Girlfriend," 2025, <https://investor.id/international/390485/pria-china-ditipu-hampir-rp-458-juta-oleh-pacar-ai>.

⁸⁹ Linda Hasibuan, "This Woman Was Scammed by an AI Impersonating Brad Pitt, Losing Rp 13 Billion," 2025, <https://www.cnbcindonesia.com/lifestyle/20250117113752-33-603948/wanita-ini-kena-tipu-brad-pitt-ai-duit-rp-13-miliar-ludes>.

⁹⁰ Prabu Revolusi and Radians Krisna Febriandy, "Human AI Communication (HA-C): Transforming the Role of Technology in Human Interaction," *Multidisciplinary Science Journal* 8, no. 3 (2026), <https://doi.org/10.31893/multiscience.2026201>.

⁹¹ Ahmed Ragab and Ali Abdelghany, "The Accountability Gap: Navigating Machine Crime and Legal Liability in the Age of Autonomous AI," n.d.

⁹² Nagwa Abouhaiba, Saed Ahmed Khouli, and Maya Khater, "Towards Protecting the Injured Party from Errors and Hallucinations of Artificial Intelligence: A Contemporary Legal Perspective," *F1000Research* 15 (2026): 1–15, <https://doi.org/10.12688/f1000research.179603.1>.

⁹³ Vu Hai Anh, "Artificial Intelligence and Criminal Liability: Challenges to Contemporary Criminal Law," *International Journal of Law* 12, no. 2 (2026).

⁹⁴ Yavuz Selim Balcioglu, Ahmet Alkan Çelik, and Erkut Altındağ, "A Turning Point in AI: Europe's Human-Centric Approach to Technology Regulation," *Journal of Responsible Technology* 23, no. July (2025), <https://doi.org/10.1016/j.jrt.2025.100128>.

Accountability model, which allocates accountability in proportion to each actor's degree of control, power, and involvement in the commission of a crime.⁹⁵ In this concept, AI is viewed as a causal factor that initiates the process of attributing accountability to parties with a legal link to the AI system, rather than as a subject of criminal law capable of culpability, like a human being.⁹⁶

AI developers are held accountable at the first level (developer accountability) if an algorithmic design error causes a criminal act or loss, carelessness during the development process, a failure in safety testing, or a disdain for relevant AI governance norms.⁹⁷ AI system providers are held accountable at the second level (provider accountability) if they fail to implement security upgrades, conduct system monitoring, or mitigate risks after implementing AI.⁹⁸ Additionally, at the third level (operator or user accountability), the operator or user may be held criminally liable if AI is used to commit a crime intentionally or carelessly, harming another person.⁹⁹ AI owners are held accountable at the last level (owner accountability) if they profit from the use of AI without implementing sufficient risk-control measures or failing to fulfill their supervisory responsibilities for the systems they own.¹⁰⁰

Therefore, the goal of this study's reconstruction of the criminal liability framework for autonomous AI is to create a mechanism for attributing liability that is more responsive to the features of autonomous technology, rather than to grant AI the status of a subject of criminal law. This model is anticipated to close the responsibility gap, improve legal certainty, and provide a conceptual framework for reforming Indonesian criminal law in light of advancements in artificial intelligence by integrating Layered Accountability, Algorithmic Sanctions, and a risk-based approach.

6. Establishment of a Multi-Layered Accountability Model

To address the attribution gap, a layered Accountability model is proposed. This model establishes a tiered priority of liability, from operation to design, to ensure that the most responsible legal subject (or the one with the most risk control) can always be held criminally liable.¹⁰¹ Criminal liability is primarily imposed on the human operator who exercises direct control or bears a duty of supervision and intervention, and is limited to clear and demonstrable cases, including: (a) negligence in failing to intervene when the AI operates beyond established parameters, particularly by disregarding system

⁹⁵ Agus Priyanto, Fero Sanjaya, and Herlita Eryke, "Reconstruction Of Criminal Liability Doctrine In The Artificial Intelligence Ecosystem In Indonesia : Analysis Of Norms ' Vapor , Reinterpretation Of Mens Rea And Hybrid Liability Model Based On Substantive Justicea" 15, no. 2 (2026): 1011–18.

⁹⁶ Miriam Buiten, Alexandre de Streel, and Martin Peitz, "The Law and Economics of AI Liability," *Computer Law and Security Review* 48 (2023): 105794, <https://doi.org/10.1016/j.clsr.2023.105794>.

⁹⁷ Gurprem Monga, "Legal Accountability of Autonomous Ai Systems in Criminal Justice," *ShodhKosh: Journal of Visual and Performing Arts* 4, no. 2 (2023): 5848–52, <https://doi.org/10.29121/shodhkosh.v4.i2.2023.6172>.

⁹⁸ Airlie Hilliard et al., *Artificial Intelligence Policy Worldwide : A Comparative Analysis* (The Royal Society, 2026), <https://doi.org/10.1098/rsos.242234/5727769/rsos.242234.pdf>.

⁹⁹ Athina Sachoulidou, "AI Systems and Criminal LiabilityA Call for Action," *Oslo Law Review* 71, no. April (2024): 1–12.

¹⁰⁰ Zhengyang Chen, "Operational Responsibility in AI Governance: A User-Centric Liability Framework," *AI and Ethics* 6, no. 306 (2026), <https://doi.org/10.2139/ssrn.6820839>.

¹⁰¹ Bart Custers, Henning Lahmann, and Benjamyn I. Scott, "From Liability Gaps to Liability Overlaps: Shared Responsibilities and Fiduciary Duties in AI and Other Complex Technologies," *AI and Society* 40, no. 5 (2025): 4035–50, <https://doi.org/10.1007/s00146-024-02137-1>.

warnings; and (b) deliberate misuse of input data through the provision of misleading or invalid information that causes the AI to commit an offense.¹⁰² Where the operator is free from error or the AI system operates with an exceptionally high level of autonomy rendering human intervention ineffective, liability is transferred to the corporation (manufacturer or system owner) under the doctrine of corporate liability.¹⁰³ This liability is grounded in structural design negligence or organizational oversight negligence, including: (a) failure to embed adequate ethical, safety, and fail-safe mechanisms in the AI's design; (b) negligence in conducting comprehensive risk testing, validation, and certification prior to public deployment, particularly where risks were foreseeable at the design stage; and (c) governance failures in maintaining, updating, and patching systems in line with evolving security standards.

In this framework, corporate liability is no longer predicated solely on corporate mens rea, but on deficiencies in corporate governance that compromise algorithmic safety, making it more readily provable than individual intent. Individual developers or programmers are only held liable if there is evidence of specific malicious intent to create a system that is deliberately designed to commit a crime, or if they abuse their position to insert malicious code (e.g., creating AI for mass financial fraud or hacking).¹⁰⁴ This level serves as the last resort and is highly dependent on proving intent during the development phase. AI system providers are liable if they fail to perform security updates, system monitoring, or risk mitigation after the AI has been implemented.¹⁰⁵

7. The Urgency of Special Criminal Measures (Algorithmic Sanctions)

International criminal penalties, such as imprisonment and conventional fines, have proven ineffective in preventing or punishing the risks posed by Autonomous AI, as these penalties do not eliminate the risks inherent in the code or system itself.¹⁰⁶ Therefore, the legal construction proposes the adoption of Special Criminal Measures in the form of Algorithmic Sanctions directed at AI systems, which must be regulated separately alongside primary criminal sanctions imposed on corporations or operators.

These measures include: (a) *forced code updates* ordered by courts to eliminate proven defective or risky behavior, subject to independent verification prior to reactivation;¹⁰⁷ (b) *autonomy restrictions* through the imposition of stricter human-in-the-loop requirements or operational limitations for high-risk systems;¹⁰⁸ (c) *progressive algorithmic fines* calculated based on both actual losses and the systemic risk generated by large-scale AI deployment;¹⁰⁹ and (d) *decommissioning* as the ultimate sanction for AI

¹⁰² Sachoulidou, "AI Systems and Criminal LiabilityA Call for Action."

¹⁰³ Chen, "Operational Responsibility in AI Governance: A User-Centric Liability Framework."

¹⁰⁴ Monga, "Legal Accountability of Autonomous Ai Systems in Criminal Justice."

¹⁰⁵ Airlie Hilliard et al., "Artificial Intelligence Policy Worldwide: A Comparative Analysis," *Royal Society Open Science* 13, no. 2 (2026), <https://doi.org/10.1098/rsos.242234>.

¹⁰⁶ Piotr Staszkiwicz et al., "Artificial Intelligence Legal Personality and Accountability: Auditors' Accounts of Capabilities and Challenges for Instrument Boundary," *Meditari Accountancy Research* 32, no. 7 (2024): 120–46, <https://doi.org/10.1108/MEDAR-10-2023-2204>.

¹⁰⁷ Adv. Lakee Ali, "AI On Trial: The Eu Ai Act's Enforcement Debut And The Emerging Admissibility Standard For Machine-Generated Evidence," *SSRN Electronic Journal* 7 (2026).

¹⁰⁸ Kamal Sharma, Praneel Sharma, and Pratyusha Sharma, "Human-in-the-Loop Governance of Artificial Intelligence in Cardiology: From Ethical Principles to Operational Paradigms," *Indian Heart Journal*, no. March (2026), <https://doi.org/10.1016/j.ihj.2026.06.002>.

¹⁰⁹ Gabriel Weil, *Overcoming Judgment-Proofness : The Law & Economics of Insuring and Mitigating AI Risk*, n.d.

systems that are fundamentally unsafe or irreparable.¹¹⁰ As a result, Algorithmic Sanctions, legal measures intended to manage and mitigate risk, must be developed. Several methods, such as forced algorithmic alteration, can be used to impose these penalties,¹¹¹ AI system operations are temporarily suspended;¹¹² recall of a product requires an algorithmic audit,¹¹³ the need for human control of some AI systems, and the cancellation of licenses for high-risk AI.¹¹⁴ The goal of putting these safeguards in place is to prevent comparable losses from recurring without violating the criminal law norm of proportionality. Additionally, our reconstruction methodology aligns with advancements in global AI governance that prioritize a risk-based strategy.¹¹⁵ A number of international documents, including the OECD AI Principles and the UNESCO Recommendation on the Ethics of Artificial Intelligence (2021),¹¹⁶ The European Union Artificial Intelligence Act and the Council of Europe Framework Convention on Artificial Intelligence (2024),¹¹⁷ identifies human oversight, accountability, transparency, and risk mitigation as fundamental tenets of AI regulation. Therefore, this proposed change to criminal liability is pertinent to efforts to reform international criminal law and consistent with evolving international standards for the administration of artificial intelligence (AI). Algorithmic sanctions focus on reducing the risks posed by AI systems, whereas international criminal law focuses on punishing offenders. As a result, their application targets AI systems through requirements that must be met by AI creators, providers, operators, or owners in compliance with the Layered Accountability principle rather than AI as a legal entity. Algorithmic penalties could be applied if an AI system: results in fatalities or severe injuries¹¹⁸; does not adhere to safety regulations; functions in a high-risk AI category without human supervision; is used as a tool for criminal activity; or continues despite algorithmic defects.

This paper concludes that an offender-oriented approach is no longer adequate to account for the features of autonomous Artificial Intelligence (AI) after examining the shortcomings of international criminal liability doctrine. This study's primary conclusion is that the framework of criminal liability needs to be rebuilt using a Layered Accountability model that allocates accountability to AI developers, providers, operators, users, and owners in accordance with their degree of control and involvement. Furthermore, this study discovered that using Algorithmic Sanctions as a risk-based

¹¹⁰ Evgeniy Solomonov, "Legal Liability of Artificial Intelligence," *Herald of Omsk University Series Law*, no. June (2025): 62, [https://doi.org/10.24147/1990-5173.2025.22\(2\).62-71](https://doi.org/10.24147/1990-5173.2025.22(2).62-71).

¹¹¹ Reuben Binns, "Algorithmic Accountability and Public Reason," *Philosophy and Technology* 31, no. 4 (2018): 543–56, <https://doi.org/10.1007/s13347-017-0263-5>.

¹¹² Alessio Tartaro, "When Things Go Wrong: The Recall of AI Systems as a Last Resort for Ethical and Lawful AI," *AI and Ethics* 5, no. 1 (2025): 253–62, <https://doi.org/10.1007/s43681-023-00327-z>.

¹¹³ Sasha Costanza-Chock, Inioluwa Deborah Raji, and Joy Buolamwini, "Who Audits the Auditors? Recommendations from a Field Scan of the Algorithmic Auditing Ecosystem," *ACM International Conference Proceeding Series*, 2022, 1571–83, <https://doi.org/10.1145/3531146.3533213>.

¹¹⁴ Andreas Holzinger, Kurt Zatloukal, and Heimo Müller, "Is Human Oversight to AI Systems Still Possible?," *New Biotechnology* 85, no. December 2024 (2025): 59–62, <https://doi.org/10.1016/j.nbt.2024.12.003>.

¹¹⁵ Cumbhadrika and Soesatyo, "Reconstruction of the Concept of Legal Liability for Losses Caused by Autonomous Artificial Intelligence Systems."

¹¹⁶ Joko, "Norma Internasional AI : Peran UNESCO Dan OECD Dalam Etika Global International AI Norms : The Role of UNESCO and OECD in Global Ethics."

¹¹⁷ Martynova, "Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law: A Commentary."

¹¹⁸ Olgun Değirmenci and Sümeyye Mengüç, "The Problem of Sanctions Applicable to Artificial Intelligence," *Journal of Artificial Intelligence and Human Sciences* 2, no. 1 (2025): 22–33.

control tool is more flexible than using international criminal sanctions against AI companies.

C. Conclusion

This study demonstrates that Hallevy's Three Models can be applied in international law: the Natural and Inevitable Consequences Model, which addresses state responsibility and due diligence; the Perpetrator-through-Intermediary Model, which addresses command responsibility (Article 28 of the Rome Statute), which is based on a commander's control and knowledge; and the Direct Liability Model, which is more suitable for identifying human failure points in the AI lifecycle. When compared to UNESCO (2021), the OECD AI Principles (2019/2024), the Council of Europe Convention (2024), and the EU AI Act (2024), it is evident that none of these documents acknowledge AI as a subject of criminal law, with accountability remaining with humans or legal entities. However, they have not yet addressed the attribution gap regarding autonomous AI because *actus reus* and *mens rea* are difficult to apply to autonomous and black box systems. In response, this study suggests a Layered Liability Model that distributes liability based on role, degree of autonomy, foreseeability, effective control, and the knew or ought to have known standard, which applies to operators for supervisory negligence, companies/manufacturers for design and governance failures, and developers for pertinent willful acts or negligence. This model complements the *ex ante* approach with *ex post* evaluations following a loss. Autonomous AI necessitates algorithmic sanctions (mandated code updates, autonomy limitations, and decommissioning) in addition to traditional criminal penalties, all backed by licensing, registration, and monitoring procedures. Therefore, a risk-based, multi-layered accountability model must replace the perpetrator-oriented approach to criminal liability under international law, while retaining people as the ultimate point of accountability.

References

- Abbott, Ryan, and Alex Sarch. "Punishing Artificial Intelligence: Legal Fiction or Science Fiction." *Is Law Computable?: Critical Perspectives on Law and Artificial Intelligence*, 2020, 177–204. <https://doi.org/10.5040/9781509937097.ch-008>.
- Abouhaiba, Nagwa, Saed Ahmed Khouli, and Maya Khater. "Towards Protecting the Injured Party from Errors and Hallucinations of Artificial Intelligence: A Contemporary Legal Perspective." *F1000Research* 15 (2026): 1–15. <https://doi.org/10.12688/f1000research.179603.1>.
- Alfrink, Kars, Ianus Keller, Gerd Kortuem, and Neelke Doorn. "Contestable AI by Design: Towards a Framework." *Minds and Machines* 33, no. 4 (2023): 613–39. <https://doi.org/10.1007/s11023-022-09611-z>.
- Ali, Adv. Lakee. "AI On Trial: The Eu Ai Act's Enforcement Debut And The Emerging Admissibility Standard For Machine-Generated Evidence." *SSRN Electronic Journal* 7 (2026).
- Anh, Vu Hai. "Artificial Intelligence and Criminal Liability : Challenges to Contemporary Criminal Law." *International Journal of Law* 12, no. 2 (2026).
- Arman, Alexandra De. "The Wild, Wild West: A Case Study of Self-Driving Vehicle Testing in Arizona." *Arizona Law Review* 61, no. 4 (2019): 983–1012.
- Atanda-Lawal, Sijuola, and Victor Chinonye Onuchukwu. "The Global Governance of AI: Between Soft Law and Binding Obligations in International Law." *International Journal of Economics, Commerce and Management* 13, no. 11 (2025): 159–84.

- <https://ijecm.co.uk/>.
- Balcioğlu, Yavuz Selim, Ahmet Alkan Çelik, and Erkut Altındağ. "A Turning Point in AI: Europe's Human-Centric Approach to Technology Regulation." *Journal of Responsible Technology* 23, no. July (2025). <https://doi.org/10.1016/j.jrt.2025.100128>.
- Bashayreh, Mohammad H., Amer Tabbara, and Fadi N. Sibai. "The Need for a Legal Standard of Care in the AI Environment." *Sriwijaya Law Review* 7, no. 1 (2023): 73–86. <https://doi.org/10.28946/slrev.vol7.iss1.1507.pp73-86>.
- Binns, Reuben. "Algorithmic Accountability and Public Reason." *Philosophy and Technology* 31, no. 4 (2018): 543–56. <https://doi.org/10.1007/s13347-017-0263-5>.
- Bloch-wehba, Hannah, and Hannah Bloch-wehba. "Visible Policing: Technology, Transparency, and Democratic Control Visible Policing: Technology, Transparency, and Democratic Control." *Texas A&M Law Scholarship* 917 (2021).
- Bonfim, Tany Calixto. "Criminal Liability of Artificial Intelligent Machines: Eyeing into AI's Mind." *Fakulty Of Law Lund University*, 2022.
- Borgesano, Francesco, Annarita De Maio, Pasquale Laghi, and Roberto Musmanno. "Artificial Intelligence and Justice: A Systematic Literature Review and Future Research Perspectives on Justice 5.0." *European Journal of Innovation Management* 28, no. 11 (2025): 349–85. <https://doi.org/10.1108/EJIM-01-2025-0117>.
- Bryan Druzin, Anatole Boute, Michael Ramsden. "Confronting Catastrophic Risk: The International Obligation to Regulate Artificial Intelligence." *Michigan Journal of International Law* 46, no. 2 (2025): 16334–68. <https://doi.org/10.18653/v1/2023.acl-long.903>.
- Buijsman, Stefan, Sarah E. Carter, and Juan Pablo Bermúdez. "Autonomy by Design: Preserving Human Autonomy in AI Decision-Support." *Philosophy and Technology* 38, no. 3 (2025). <https://doi.org/10.1007/s13347-025-00932-2>.
- Buiten, Miriam C. *Product Liability for Defective AI*. *European Journal of Law and Economics*. Vol. 57. Springer US, 2024. <https://doi.org/10.1007/s10657-024-09794-z>.
- Buiten, Miriam, Alexandre de Streel, and Martin Peitz. "The Law and Economics of AI Liability." *Computer Law and Security Review* 48 (2023): 105794. <https://doi.org/10.1016/j.clsr.2023.105794>.
- Chen, Zhengyang. "Operational Responsibility in AI Governance: A User-Centric Liability Framework." *AI and Ethics* 6, no. 306 (2026). <https://doi.org/10.2139/ssrn.6820839>.
- Costanza-Chock, Sasha, Inioluwa Deborah Raji, and Joy Buolamwini. "Who Audits the Auditors? Recommendations from a Field Scan of the Algorithmic Auditing Ecosystem." *ACM International Conference Proceeding Series*, 2022, 1571–83. <https://doi.org/10.1145/3531146.3533213>.
- Court, International, and O F Justice. "Application of the International Convention for the Suppression of the Financing of Terrorism and of the International Convention on the Elimination of All Forms of Racial Discrimination" III (2022): 5–54. <https://doi.org/10.18356/9789210014229c001>.
- Cox, Brian L. "Clarifying The Role Of International Criminal Tribunals As A Source Of Customary International Law," 2021, 1–75.
- Cumbhadrika, Chitto, and Bambang Soesatyo. "Reconstruction of the Concept of Legal Liability for Losses Caused by Autonomous Artificial Intelligence Systems." *International Journal of Social Service and Research* 6, no. 6 (2026): 499–509.

- <https://doi.org/10.46799/ijssr.v6i6.1418>.
- Custers, Bart, Henning Lahmann, and Benjamyn I. Scott. "From Liability Gaps to Liability Overlaps: Shared Responsibilities and Fiduciary Duties in AI and Other Complex Technologies." *AI and Society* 40, no. 5 (2025): 4035–50. <https://doi.org/10.1007/s00146-024-02137-1>.
- Dave, Shyamal. "Artificial Intelligence 'S Liability , Judging The Future-Today." *JLAI*, no. June (2023): 30–36.
- Değirmenci, Olgun, and Sümeyye Mengüç. "The Problem of Sanctions Applicable to Artificial Intelligence." *Journal of Artificial Intelligence and Human Sciences* 2, no. 1 (2025): 22–33.
- Dora, Grace El. "Chinese Man Scammed Out of Nearly Rp 458 Million by His AI Girlfriend," 2025. <https://investor.id/international/390485/pria-china-ditipu-hampir-rp-458-juta-oleh-pacar-ai>.
- Eisner, Rrbeca S. "Artificial Intelligence Artificial Intelligence Series." *Transactions & Business* 49, no. 9 (2020): 47–49.
- El-Kady, Ramy. *Artificial Intelligence from the Criminal Law Perspective. Lecture Notes on Data Engineering and Communications Technologies*. Vol. 233. Springer Nature Switzerland, 2024. https://doi.org/10.1007/978-3-031-77299-3_16.
- Europe, Council of. "Council Of Europe Framework Convention On Artificial Intelligence And Human Rights, Democracy And The Rule Of Law," n.d.
- European Union. "Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 down Harmonised Rules on Artificial Intelligence(Artificial Intelligence Act)." *Official Journal of the European Union, L168* 1689, no. 3 (2024): 1–153. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>.
- Fisher, Michael, Viviana Mascardi, Kristin Yvonne Rozier, Bernd Holger Schlingloff, Michael Winikoff, and Neil Yorke-Smith. *Towards a Framework for Certification of Reliable Autonomous Systems. Autonomous Agents and Multi-Agent Systems*. Vol. 35. Springer US, 2021. <https://doi.org/10.1007/s10458-020-09487-2>.
- Florea, Dumitrita, Narcisa Gales, Geanina Romascan, Georgiana Rusu, and Teofana Botoşan. "The International Criminal Court and Artificial Intelligence: Case Study and Global Legal Challenges." *European Journal of Law and Public Administration* 12, no. 2 (2025): 22–28. <https://doi.org/10.18662/eljpa/12.2/263>.
- Gaeta, Paola. "Who Acts When Autonomous Weapons Strike?" *Journal of International Criminal Justice* 21, no. 5 (2023): 1033–55. <https://doi.org/10.1093/jicj/mqae001>.
- Gevaert, Caroline Margaux, Sanja Šćepanović, Nadia Bernaz, Hannes Taubenböck, and Mrinalini Kochupillai. "Responsible Artificial Intelligence for Earth Observation: Human Rights and the EU AI Act." *AI and Ethics* 6, no. 3 (2026). <https://doi.org/10.1007/s43681-026-01149-5>.
- Giannini, Alice. *Criminal Behavior and Accountability of Artificial Intelligence Systems*, 2023. <https://doi.org/10.26481/dis.20231124ag>.
- Hallevy, Prof. Gabriel. "The Basic Models of Criminal Liability of AI Systems and Outer Circles." *SSRN Electronic Journal*, no. 2013 (2019). <https://doi.org/10.2139/ssrn.3402527>.
- Harland, Hadassah, Richard Dazeley, Hashini Senaratne, Peter Vamplew, Francisco Cruz, and Bahareh Nakisa. "AI Apology: A Critical Review of Apology in AI Systems." *Artificial Intelligence Review* 58, no. 12 (2025). <https://doi.org/10.1007/s10462-025-11305-8>.
- Hasibuan, Linda. "This Woman Was Scammed by an AI Impersonating Brad Pitt, Losing Rp 13 Billion," 2025.

- <https://www.cnbcindonesia.com/lifestyle/20250117113752-33-603948/wanita-ini-kena-tipu-brad-pitt-ai-duit-rp-13-miliar-ludes>.
- Hassan, F. M., and N. D. Osman. "AI-Based Autonomous Weapons and Individual Criminal Responsibility Under the Rome Statute." *Journal of Digital Technologies and Law* 1, no. 2 (2023): 464–80. <https://doi.org/10.21202/jdtl.2023.19>.
- Hayward, Keith J., and Matthijs M. Maas. "Artificial Intelligence and Crime: A Primer for Criminologists." *Crime, Media, Culture* 17, no. 2 (2021): 209–33. <https://doi.org/10.1177/1741659020917434>.
- Herkert, Joseph, Jason Borenstein, and Keith Miller. "The Boeing 737 MAX: Lessons for Engineering Ethics." *Science and Engineering Ethics* 26, no. 6 (2020): 2957–74. <https://doi.org/10.1007/s11948-020-00252-y>.
- Hilliard, Airlie, Ayesha Gulley, Emre Kazim, and Adriano Soares Koshiyama. *Artificial Intelligence Policy Worldwide: A Comparative Analysis*. The Royal Society, 2026. <https://doi.org/10.1098/rsos.242234/5727769/rsos.242234.pdf>.
- . "Artificial Intelligence Policy Worldwide: A Comparative Analysis." *Royal Society Open Science* 13, no. 2 (2026). <https://doi.org/10.1098/rsos.242234>.
- Hodula, Máté. "Criminal Liability of Self-Driving Vehicles Under Hungarian Criminal Law." *Studia Iuridica* 100 (2023). <https://doi.org/10.31338/2544-3135.si.2024-100.9>.
- Holzinger, Andreas, Kurt Zatloukal, and Heimo Müller. "Is Human Oversight to AI Systems Still Possible?" *New Biotechnology* 85, no. December 2024 (2025): 59–62. <https://doi.org/10.1016/j.nbt.2024.12.003>.
- ICC. *Rome Statute of the International Criminal Court*. International Criminal Court, 2021. <http://www.ssrn.com/abstract=1689616>.
- Johnston, Mitchell Chervu, and Brian Matejek. "Overcoming Unforeseeable AI Errors: A Proposal." *Boston College Law School Legal Studies Research Paper* 675 (2026).
- Joko. "Norma Internasional AI: Peran UNESCO Dan OECD Dalam Etika Global International AI Norms : The Role of UNESCO and OECD in Global Ethics." *Iuridice et Societas* 1, no. 1 (2023): 113–37.
- KAN, Celal Hakan. "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." *International Journal of Eurasia Social Sciences* Vol: 15, no. Issue: 55 (2024): 276–313.
- Kang, Byeong Ho, Wenli Yang, and Muhammad Bilal Amin. "Trustworthy Orchestration Artificial Intelligence by the Ten Criteria with Control-Plane Governance." *ArXiv Preprint ArXiv:2512.10304*, 2025.
- Krisna, I Gede, and Bayu Anggara. "Legal Liability of Artificial Intelligence in Criminal Law : A Comparative Study between Indonesia and European Union," no. 3 (2026): 1–12.
- Leon, Maikel. "Lifecycle-Based Governance to Build Reliable Ethical AI Systems." *Systems Research and Behavioral Science* 43, no. 3 (2026): 1116–32. <https://doi.org/10.1002/sres.70014>.
- Leonard, Eric K. "The Rome Statute." *The Onset of Global Governance*, no. July 1998 (2020): 53–72. <https://doi.org/10.4324/9781315237985-10>.
- Marchuk, Iryna. *The Fundamental Concept of Crime in International Criminal Law: A Comparative Law Analysis*. The Fundamental Concept of Crime in International Criminal Law: A Comparative Law Analysis, 2012. <https://doi.org/10.1007/978-3-642-28246-1>.
- Martynova, Ekaterina. "Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law: A Commentary." *AI*

- from the Global Majority: Official Outcome Report of the UN IGF Data and Artificial Intelligence Governance Coalition, 2026, 81–89. https://doi.org/10.1007/978-3-032-07830-8_9.
- Mashai, Abulgasem Al. "Post-Human Law: Redefining Responsibility in the Era of Autonomous Artificial Intelligence." *SSRN Electronic Journal*, 2025, 1–19. <https://doi.org/10.2139/ssrn.5716462>.
- Milanovic, Marko. "The Use of Artificial Intelligence in Military Operations: Complicity, State Responsibility and Individual Criminal Responsibility under International Law." *Caitlin Chiaramonte and Ashley Deeks (Eds.), Artificial Intelligence and Multilateral Military Operations: Prospects and Challenges (Oxford University Press, Forthcoming)*, 2026, 1–23. <https://doi.org/https://ssrn.com/abstract=6183440> or <http://dx.doi.org/10.2139/ssrn.6183440>.
- Monga, Gurprem. "Legal Accountability of Autonomous Ai Systems in Criminal Justice." *ShodhKosh: Journal of Visual and Performing Arts* 4, no. 2 (2023): 5848–52. <https://doi.org/10.29121/shodhkosh.v4.i2.2023.6172>.
- Morandín-Ahuerma, Fabio. "Ten UNESCO Recommendations on the Ethics of Artificial Intelligence." *Normative Principles for an Ethics of Artificial Intelligence*, no. 2023 (2023): 86–94.
- Muhammad Ahsan Iqbal Hashmi, Muhammad Faiq Butt, Muhammad Jawad, and Dr. Sajid Sultan. "Criminal Liability in the Age of Autonomous Systems: Rethinking Mens Rea and Actus Reus." *The Critical Review of Social Sciences Studies* 3, no. 3 (2025): 290–303. <https://doi.org/10.59075/szbtkr93>.
- Mustafa, Mohammad. "Understanding of Machine Learning with Deep Learning:" *Computers MDPI* 12, no. 91 (2023): 1–26.
- Nations, United. "Resolution 56/83, 'Responsibility of States for Internationally Wrongful Acts', UN Doc. A/56/49(Vol. I)/Corr.4, 12 December 2001" II, no. December 2001 (2001).
- Nerantzi, Elina, and Giovanni Sartor. "'Hard AI Crime':The Deterrence Turn." *Oxford Journal of Legal Studies* 44, no. 3 (2024): 673–701. <https://doi.org/10.1093/ojls/gqae018>.
- Ngozi Samuel Uzougbo, Chinonso Gladys Ikegwu, and Adefolake Olachi Adewusi. "Legal Accountability and Ethical Considerations of AI in Financial Services." *GSC Advanced Research and Reviews* 19, no. 2 (2024): 130–42. <https://doi.org/10.30574/gscarr.2024.19.2.0171>.
- Obergassel, Wolfgang, Steffen Bauer, Lukas Hermwille, Stefan C. Aykut, Idil Boran, Sander Chan, Carolin Fraude, et al. "From Regime-Building to Implementation: Harnessing the UN Climate Conferences to Drive Climate Action." *Wiley Interdisciplinary Reviews: Climate Change* 13, no. 6 (2022): 1–12. <https://doi.org/10.1002/wcc.797>.
- OECD. "Recommandation Du Conseil Sur l' Intelligence Artificielle, OECD/LEGAL/0449," 2019.
- Páez, Andrés. "Understanding with Toy Surrogate Models in Machine Learning." *Minds and Machines* 34, no. 4 (2024): 1–26. <https://doi.org/10.1007/s11023-024-09700-1>.
- Piparo, Carlo. "Criminal Liability Models and Criminal Participation in the Digital Environment: A Modern Challenge in the Perspective of Italian Constitutionalism." *Zbornik Radova Pravnog Fakulteta, Novi Sad* 57, no. 4 (2023): 1357–78. <https://doi.org/10.5937/zrpfns57-47113>.
- Piri, Heidar. "Criminal Accountability of Artificial Intelligence Systems for International Crimes and the Attributability of Combatant Status or E-Personhood : Necessities ,

- Obstacles , and Solutions.” *Journal of Criminal Law Research* 13, no. 52 (2025): 151–96. <https://doi.org/10.22054/jclr.2025.80888.2687>.
- Priyanto, Agus, Fero Sanjaya, and Herlita Eryke. “Reconstruction Of Criminal Liability Doctrine In The Artificial Intelligence Ecosystem In Indonesia : Analysis Of Norms’ Vapor , Reinterpretation Of Mens Rea And Hybrid Liability Model Based On Substantive Justicea” 15, no. 2 (2026): 1011–18.
- Ragab, Ahmed, and Ali Abdelghany. “The Accountability Gap: Navigating Machine Crime and Legal Liability in the Age of Autonomous AI,” n.d.
- Ramos, Gabriela. “UNESCO’s Recommendation on the Ethics of Artificial Intelligence (2021).” *International Organization Initiatives*, no. November (2025): 203–14. <https://doi.org/10.1093/9780197803325.003.0020>.
- Revolusi, Prabu, and Radians Krisna Febriandy. “Human AI Communication (HA-C): Transforming the Role of Technology in Human Interaction.” *Multidisciplinary Science Journal* 8, no. 3 (2026). <https://doi.org/10.31893/multiscience.2026201>.
- Rintamäki, Tytti, Delaram Golpayegani, Dave Lewis, Edoardo Celeste, and Harshvardhan J. Pandit. “Impact Assessment Requirements in the GDPR vs the AI Act: Overlaps, Divergence, and Implications.” *Computer Law and Security Review* 61, no. April (2026): 106317. <https://doi.org/10.1016/j.clsr.2026.106317>.
- Robertson, Cassandra Burke. “Litigating Partial Autonomy.” *Iowa Law Review* 109, no. 4 (2024): 1655–1702.
- Robintan Sulaiman. *The Law in the Era of Artificial Intelligence*. RSP Forensic Legal Auditor Specialist, 2021.
- Rotenberg, Marc. “Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (Council Eur.).” *International Legal Materials* 64, no. 3 (2025): 859–902. <https://doi.org/10.1017/ilm.2025.1>.
- Sachoulidou, Athina. “AI Systems and Criminal LiabilityA Call for Action.” *Oslo Law Review* 71, no. April (2024): 1–12.
- Saja Raheem Mohammad, and Noor Hassanin Hashim. “Review on the Relationship between Artificial Intelligence and Computer Viruses.” *Open Access Research Journal of Science and Technology* 13, no. 2 (2025): 064–067. <https://doi.org/10.53022/oarjst.2025.13.2.0040>.
- Santoso, Topo. *Asas-Asas Hukum Pidana*. Cetakan Pe. PT. Raja Grafindo Persada, 2023.
- Sharma, Kamal, Praneel Sharma, and Pratyusha Sharma. “Human-in-the-Loop Governance of Artificial Intelligence in Cardiology: From Ethical Principles to Operational Paradigms.” *Indian Heart Journal*, no. March (2026). <https://doi.org/10.1016/j.ihj.2026.06.002>.
- Siau, Keng, and Weiyu Wang. “Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI.” *Journal of Database Management* 31, no. 2 (2020): 74–87. <https://doi.org/10.4018/JDM.2020040105>.
- Siregar, Naek, Desy Churul Aini, Rehulina, Agit Yogi Subandi, and Isroni Muhammad Miraj Mirza. “The Use of Artificial Intelligence in Armed Conflict under International Law.” *Hasanuddin Law Review* 10, no. 2 (2024): 189–205. <https://doi.org/10.20956/halrev.v10i2.5267>.
- Solomonov, Evgeniy. “Legal Liability of Artificial Intelligence.” *Herald of Omsk University Series Law*, no. June (2025): 62. [https://doi.org/10.24147/1990-5173.2025.22\(2\).62-71](https://doi.org/10.24147/1990-5173.2025.22(2).62-71).
- Stanusch, Natalia, and Richard Rogers. “How AI Is Imagined by Industry during the Sam Altman Controversy.” *Authorea Preprints*, no. June (2025). <https://advance.sagepub.com/doi/pdf/10.31124/advance.174979411.18178682>

/v1.

- Stepanovich, Amie, Daniel Bedoya-Arroyo, Gustaf Bjorksten, Michael Carbone, Drew Mitnick, Donna Wentworth, and Nathan White. "A Human Rights Response To Government Hacking." *Accessnow*, 2016. <https://www.accessnow.org/cms/assets/uploads/2016/09/GovernmentHackingDoc.pdf>.
- Stoyanova, Vladislava. "Council of Europe Framework Convention on Artificial Intelligence: Context, Regulatory Approach and Scope of Obligations." *European Journal of Risk Regulation* 16 (2025): 536–65. <https://doi.org/10.1017/err.2025.10070>.
- Sulistio, Faizin, Haru Permadi, Achmad Aldy Hifdillah, and Oppy Pramudya. "The Prospects of Sentencing Artificial Intelligence Based Autonomous Agents." *Jurnal Akta* 13, no. 1 (2026): 415–25.
- Sulistio, Faizin, and Aizahra Daffa Salsabilla. "Pertanggungjawaban Pada Tindak Pidana Yang Dilakukan Agen Otonom Artificial Intelegence." *Unes Law Review* 6, no. 2 (2023): 5479–90.
- Swart, Mia. "Constructing 'Electronic Liability' for International Crimes: Transcending the Individual in International Criminal Law." *German Law Journal* 24, no. 3 (2023): 589–602. <https://doi.org/10.1017/glj.2023.28>.
- Tait, Izak. "Lions and Tigers and AI , Oh My : An Ethical Framework for Human-AI Interaction Based on the Five Freedoms of Animal Welfare Lions and Tigers and AI , Oh My : An Ethical Framework Animal Welfare." *Preprints.Org*, 2024. <https://doi.org/10.20944/preprints202407.1846.v1>.
- Tartaro, Alessio. "When Things Go Wrong: The Recall of AI Systems as a Last Resort for Ethical and Lawful AI." *AI and Ethics* 5, no. 1 (2025): 253–62. <https://doi.org/10.1007/s43681-023-00327-z>.
- Thuy, Ha Le, Viet Nam, Nguyen Van Nghiep, Lam Dong, Viet Nam, Nguyen Thi Van Anh, Lam Dong, and Viet Nam. "Criminal Liability In Relation To Artificial Intelligence: A Comparative Study Of Select Jurisdictions And Vietnam." *Veredas Do Direito* 23 (2026).
- Wang, Xuanyi. "A Review of the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law." *Journal of Politics and Law* 18, no. 4 (2025): 68. <https://doi.org/10.5539/jpl.v18n4p68>.
- Weil, Gabriel. *Overcoming Judgment-Proofness: The Law & Economics of Insuring and Mitigating AI Risk*, n.d.
- 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, no. 2 (2025): 630–58. <https://doi.org/10.53955/jhcls.v5i2.606>.

