Journal of International Islamic Law, Human Right and Public Policy

Journal of International Islamic Law, Human Rights and Public Policy
(JISHUP Journal) | ISSN(e): 3031-2280
Website https://jishup.org/index.php/ojs/index | Volume 4 No. 4 (2026)

CRIMINAL LIABILITY FOR MISUSE OF ARTIFICIAL INTELLIGENCE TECHNOLOGY (ARTIFICIAL INTELLIGENCE/AI) IN THE PERSPECTIVE OF THE NATIONAL CRIMINAL CODE

Agus Siagian*, Erniyanti, Irpan Husein, Markus Gunawan
Universitas Batam, Batam, Indonesia
E-mail: siagian.agus76@gmail.com* , erniyanti@univbatam.ac.id , irpan@univbatam.ac.id , markusgunawan@hotmail.com
Corresponding Author: erniyanti@univbatam.ac.id

Received: 03/07/2026 | **Revised:** 10/07/2026 | **Accepted:** 29/07/2026 | **Published:** 10/08/2026

Abstract

The development of Artificial Intelligence (AI) technology has provided significant benefits across various sectors; however, it has also led to various forms of misuse, including deepfakes, digital fraud, data manipulation, and cybercrime. The regulation of criminal liability for AI misuse in Indonesia is currently based on the Indonesian Criminal Code (KUHP), Law Number 1 of 2024 concerning the Second Amendment to Law Number 11 of 2008 on Electronic Information and Transactions, as well as other relevant laws and regulations. Its implementation still faces several challenges, including the absence of specific legal provisions, difficulties in digital evidence, and the determination of the responsible legal subject. Therefore, strengthening legal regulations and developing adaptive criminal law policies are essential. This study aims to analyze the legal framework, implementation, challenges, and solutions concerning criminal liability for the misuse of Artificial Intelligence from the perspective of the Indonesian Criminal Code. This study employs the theories of Criminal Liability, Legal Certainty, and Law Enforcement as the analytical framework. The conceptual framework encompasses criminal liability, the misuse of Artificial Intelligence, and the Indonesian Criminal Code. This research applies a normative legal research method using statutory, conceptual, and case approaches. The study relies on secondary data consisting of primary, secondary, and tertiary legal materials, which are analyzed qualitatively. The findings indicate that the current legal framework governing criminal liability for AI misuse has not yet provided adequate legal certainty, resulting in challenges related to evidence and the determination of responsible parties. The novelty of this research lies in the development of a criminal liability concept that integrates the principles of fault, accountability of AI developers, providers, and users as part of the reform of Indonesia's national criminal law. This study concludes that the regulation of criminal liability for the misuse of Artificial Intelligence should be strengthened to accommodate technological advancements while ensuring legal certainty. It recommends the enactment of specific AI regulations, harmonization with the Indonesian Criminal Code, and the enhancement of law enforcement capacity in addressing AI-related crimes.

Keywords: Criminal Liability; Misuse of Artificial Intelligence (AI); Indonesian Criminal Code; Law Enforcement; Legal Certainty.

INTRODUCTION

Advances in science and technology have transformed patterns of social interaction, economics, governance, and law enforcement. One of the most prominent innovations in digital transformation is Artificial Intelligence (AI), a technology that enables computer systems to perform learning, pattern recognition, language processing, data analysis, and decision-making that mimic human intellectual capabilities. The use of AI increases efficiency and productivity in the education, health, trade, industry, transportation, public services, and law enforcement sectors. However, this same capacity can also be used to commit crimes on a scale, speed, and reach that surpasses conventional crimes. Forms of abuse of concern include deepfakes for identity fraud, facial and voice manipulation, digital fraud, the spread of false information, theft of personal data, automated cyberattacks, and the use of malicious software. These actions can cause material losses, reputational damage, privacy violations, and disruptions to public order and security. Because AI output can be generated through complex algorithmic processes, law enforcement

must not only prove the existence of prohibited content or effects but also determine the relationship between system design, user commands, operational controls, and the resulting criminal consequences. The changing nature of AI-based crimes raises fundamental issues for criminal law. In conventional crimes, the relationship between the perpetrator's actions and the consequences is generally easier to trace. In AI systems, output can be influenced by training data, model design, service provider configuration, application developer integration, user instructions, and the system's learning mechanisms. This complexity raises questions about who should be held accountable: the user who issues the instructions, the developer who created the system, the platform provider, the corporate owner or manager, or any other party who benefits from the use of AI.

Indonesian criminal law still places fault as the basis for punishment. The principle of "geen straf zonder schuld" requires that punishment be imposed only on parties capable of responsibility, who committed the act intentionally or negligently, and who have no justification or excuse (Chazawi, 2022; Hamzah, 2022; Hiariej, 2023). AI has not yet been recognized as a subject of criminal law with the will and capacity to take responsibility. Therefore, this technology is essentially positioned as an instrument, while responsibility is directed at humans or corporations that have a relationship of fault, control, contribution, or benefit from the crime.

Normatively, Indonesia does not yet have a specific law that comprehensively regulates criminal liability for AI misuse. The legal basis is still scattered across Law Number 1 of 2023 concerning the Criminal Code, Law Number 1 of 2024 concerning the Second Amendment to the Electronic Information and Transactions Law, Law Number 27 of 2022 concerning Personal Data Protection, and other sectoral provisions. These regulations can be used to address specific consequences or actions, but they do not specifically define AI, the division of liability between actors, the standard of care, and the method of attribution of fault within the multi-layered technology ecosystem.

The absence of specific regulations has the potential to lead to differing interpretations, inconsistent application of the law, and difficulties in establishing evidence. Additional challenges arise because AI-based crimes often utilize anonymous identities, overseas servers, cross-border platforms, and volatile electronic data. Institutionally, law enforcement requires digital forensic skills, an understanding of algorithms, laboratory support, and coordination with electronic system administrators. At the societal level, low digital literacy increases the risk of victimization and weakens the reporting and preservation of electronic evidence.

Based on these issues, this study analyzes three issues. First, how the legal regulations governing criminal liability for AI misuse are constructed from the perspective of the National Criminal Code. Second, how the principle of criminal liability is implemented when AI is used as a means of criminal activity. Third, the main obstacles to law enforcement and the necessary legal reform solutions are identified. The study aims to formulate a model of accountability that remains grounded in the principles of legality and culpability, yet maps the roles of developers, providers, users, managers, and corporations in the chain of AI use.

LITERATURE REVIEW

Criminal Liability and the Principle of Fault

Criminal liability is the transfer of objective blame for a criminal act to a perpetrator who subjectively qualifies for punishment. An act that meets the elements of a crime does not automatically result in a criminal offense, as it must also be proven that the perpetrator was responsible, whether intentionally or negligently, and that there was no reason to excuse the crime. Therefore, criminal law distinguishes between criminal acts and liability for those acts. This distinction is crucial in AI-based cases because the system can produce unlawful output, while the fault must still be traced to a human actor or corporation (Hamzah, 2022; Sofyan & Azisa, 2023).

The principle of actus non facit reum nisi mens sit rea asserts that an act does not render a person guilty without a blameworthy mental attitude. In the use of AI, actus reus can take the form of creating, distributing, processing, accessing, manipulating, or using AI output for illicit purposes. Mens rea can be reflected in instructions deliberately designed to generate false content, knowledge that the system will be used for fraud, omission of a real risk, or failure to fulfill a duty of care. Because technical processes involve multiple parties, proving fault needs to be directed at relevant decisions and controls, not at the assumption that every party in the automated technology chain is responsible.

AI does not yet have the status of a subject of criminal law. The system lacks legal consciousness, moral capacity, or the ability to accept and undergo punishment in the human sense. Granting criminal personality to AI also risks breaking the chain of human accountability, as perpetrators could shift blame to the machine. Therefore, a more consistent approach with the National Criminal Code is to maintain human and corporate legal status, while treating AI's technical autonomy as a factual circumstance that impacts the proof of causal relationships, knowledge, control, and negligence.

Corporate liability becomes relevant when AI development or provision occurs within an organizational structure. Corporate culpability can be assessed through policies, management decisions, compliance culture, omissions, oversight failures, or the benefits derived from system use. This concept allows the law to address violations not solely stemming from end users, but also from business design or governance that knowingly ignores risks. However, attribution to the corporation must be based on a demonstrable link between the corporate activity and the crime, to avoid absolute liability, which contradicts the principle of fault (Sjahdeini, 2022).

Previous studies have highlighted accountability for the use of AI, AI-based crimes within the framework of the new Criminal Code, and developer responsibility for deepfakes and electronic document forgery (Haris & Tantimin, 2022; Wijanarko et al., 2025; Sativa et al., 2025; Budiman et al., 2026). This literature demonstrates the need to expand the analysis from the direct perpetrator to the actor ecosystem. This study continues this discussion by integrating the principles of culpability, technical roles, control, benefits, duty of care, and corporate liability into a single, layered attribution model.

Legal Certainty and the Legal System

Legal certainty demands clear, consistent, understandable, and enforceable rules. In the context of AI, legal certainty relates not only to the formulation of prohibitions but also to determining which parties are obligated to take preventative measures, store data, maintain system security, and provide access to information for law enforcement purposes. When regulations only address electronic acts in general, authorities can adopt a technology-neutral approach, but the scope for interpretation becomes wider in cases involving machine learning, probabilistic outputs, and decisions that are not fully explainable.

Legal certainty must be balanced with justice and utility. Regulations that are too general may fail to protect victims, while regulations that are too rigid risk stifling innovation and quickly becoming left behind by technological developments. Therefore, criminal law reform needs to focus on risks and functions, not just the name of the technology. Acts that harm legal interests remain explicitly defined, while actors' obligations are adjusted to the level of control and foreseeable risks. This approach maintains the principle of legality while allowing the law to respond to changes in the form and capabilities of AI systems.

Lawrence M. Friedman's legal system theory explains the effectiveness of law through three components: legal substance, legal structure, and legal culture. Legal substance includes the National Criminal Code, the Electronic Information and Transactions Law, the Personal Data Protection Law, and sectoral regulations that form the basis for enforcement. Legal structure includes the police, prosecutors, courts, supervisory agencies, ministries, electronic system providers, and forensic institutions. Legal culture encompasses the level of public awareness, industry compliance, developer ethics, and the ability of officials to understand technological risks.

These three components are interdependent. Good criminal law is ineffective without the technical capability to identify data sources, models, accounts, devices, and transaction traces. Conversely, strong forensic capabilities cannot replace a clear legal basis. The culture of AI use also determines the effectiveness of prevention, as some violations occur due to the habit of spreading unverified content, sharing personal data, using anonymous services, or ignoring account security. Therefore, the solution cannot simply consist of new criminalizations but must encompass procedures, institutional capacity, technical standards, and public education.

Conceptual Boundaries

In this study, AI abuse is understood as the use of algorithmic and data-based systems to generate, modify, automate, or amplify unlawful acts. Its scope includes digital fraud, deepfakes, electronic information manipulation, identity theft, personal data breaches, fake pornographic content, and cyberattacks. The National Criminal Code (KUHP) is positioned as a general framework for criminal liability, while sectoral laws define specific acts and protect specific legal interests. This definition positions AI as a means or environment for criminal acts, rather than a stand-alone perpetrator.

Research Gap and Analytical Framework

The literature on law and AI in Indonesia is developing in several directions. Some studies discuss whether AI can be held independently accountable; others assess the responsibility of users or developers in cases of deepfakes, forgery, and autonomous systems. These studies show two tendencies. First, positive law continues to position humans or corporations as subjects. Second, the division of responsibility between actors lacks uniform parameters, particularly when adverse outcomes arise from the interaction between system design, data, configuration, and user instructions.

The research gap lies in the lack of integration of criminal norm analysis with the structure of the AI ecosystem. Simply labeling users as perpetrators is not always sufficient if providers know that certain features are being systematically used for crime and still allow it. Conversely, placing responsibility on developers simply for creating technology is also unfair when the system is used deviantly beyond its intended purpose. Parameters are needed to distinguish the creation of versatile technology from actual contributions to crime.

This study uses five attribution parameters. First, role, namely position as developer, provider, deployer, operator, user, administrator, or corporation. Second, control, namely the ability to regulate models, data, access, configuration, and use. Third, knowledge or foreseeability, namely whether the actor knew or should have known about the risk of misuse. Fourth, causal contribution, namely the relationship between an action or omission and the occurrence of a consequence. Fifth, benefit, namely the benefit obtained and its relevance to the obligation of supervision. These parameters do not replace the element of crime, but rather help apply the element of culpability to complex technological structures.

This framework also distinguishes three situations. In direct misuse, a user intentionally exploits AI to produce or distribute criminal objects. In facilitated misuse, a system provider or administrator provides access, support, or condone the commission of a crime. In governance failure, the consequences arise from a failure to meet security standards or adequate oversight. Each situation requires different levels of evidence, particularly regarding intent, negligence, participation, and corporate responsibility.

The novelty of this research is its multi-layered accountability formulation, which neither recognizes AI as a criminal actor nor limits the analysis to end users. The model links the principle of culpability to the accountability of each actor based on control, knowledge, contribution, benefit, and duty of care. With this framework, the law can avoid two extremes: impunity, where actions are attributed to machines, and overcriminalization, where all parties in the technological chain are held responsible without proof of fault.

METHOD

This research is a normative legal study with a descriptive-analytical nature. The study focuses on the norms governing criminal liability for AI misuse from the perspective of the National Criminal Code. The research not only describes the applicable provisions but also assesses their adequacy in addressing the characteristics of AI systems, identifies implementation issues, and formulates recommendations for criminal law reform.

The first approach is a legislative approach. The materials reviewed include the 1945 Constitution of the Republic of Indonesia, Law Number 1 of 2023 concerning the Criminal Code, Law Number 1 of 2024 concerning the Second Amendment to the Electronic Information and Transactions Law, Law Number 27 of 2022 concerning Personal Data Protection, provisions of criminal procedure law, the Telecommunications Law, the Pornography Law, and other regulations related to electronic systems and cybercrime.

The second approach is a conceptual approach to examine the doctrines of criminal liability, fault, capacity to be responsible, intent, negligence, participation, corporate liability, legal certainty, and the legal system. The third approach is a case-based approach, analyzing case patterns related to deepfakes, digital fraud, data manipulation, the dissemination of false information, personal data breaches, and the use of autonomous AI systems. Because the case materials in the research sources are used to identify normative patterns, the research does not conclude empirical facts or the prevalence of crime.

Research sources consist of primary, secondary, and tertiary legal materials. Primary materials consist of binding laws and regulations. Secondary materials include books on criminal law and cyber law, journal articles, theses, dissertations, institutional reports, and expert opinions. Tertiary materials include legal dictionaries, encyclopedias, and technology terminology references. The collection of materials was conducted through a literature review, which involved inventorying, identifying, classifying, and selecting materials based on their relevance to the problem formulation.

The analysis was conducted qualitatively through several stages. First, identifying the legal interests and applicable elements of criminal acts. Second, mapping the roles of actors in the AI ecosystem. Third, examining the relationship between actions, errors, control, and consequences. Fourth, assessing the adequacy of the legal substance, structure, and culture. Fifth, developing prescriptions in the form of accountability models and regulatory reform steps. Conclusions were drawn deductively, moving from general principles of criminal law to their application to forms of AI abuse.

The validity of the argument is maintained through consistency between norms, doctrines, and conclusions. Each proposed responsibility is tested against the principles of legality and the principle of no crime without fault. Thus, the research does not position technological risk as a stand-alone basis for criminal action, but rather as a

context for assessing the obligations, knowledge, negligence, and contributions of actors. This method is appropriate for formulating legal policy because the primary research issue is the gap between norms and the construction of responsibility, not the measurement of population behavior.

RESULTS AND DISCUSSION

1. Regulatory Framework for Criminal Liability in AI Misuse

The research findings indicate that criminal liability provisions for AI misuse remain fragmentary. The National Criminal Code provides general principles regarding criminal offenses, misconduct, participation, attempted violations, and corporate liability. The Electronic Information and Transactions (ITE) Law regulates various unlawful acts in the electronic space, while the Personal Data Protection Law protects data from unauthorized collection, use, processing, and dissemination. Telecommunications and pornography provisions can be invoked when AI is used over a network or to create illicit content. These instruments collectively form the basis for enforcement, but do not yet constitute a unified AI regime.

The positive legal approach is technology-neutral. This means the primary focus is not on specific technologies, but rather on prohibited acts and consequences. Deepfakes can be assessed through elements of fraud, forgery, defamation, blackmail, pornography, or data breach, depending on the facts of the case. This approach has the advantage of remaining applicable rules as technology changes. The disadvantage is the lack of clarity regarding technical obligations and the division of responsibility within AI systems, particularly when harm arises from features designed, integrated, and operated by different parties.

Table 1. Legal Instruments Relevant to AI Misuse

Legal Instrument	Normative Focus	Relevance to AI Misuse
Law No. 1 of 2023 (National Criminal Code)	General principles of offenses, fault, participation, punishment, and corporate liability.	Provides the principal framework for attributing criminal responsibility to persons and corporations.
Law No. 1 of 2024 (Electronic Information and Transactions)	Unlawful electronic information, access, manipulation, distribution, and electronic systems.	Applies when AI is used to create, alter, access, or disseminate prohibited electronic content.
Law No. 27 of 2022 (Personal Data Protection)	Lawful processing, security, confidentiality, and protection of personal data.	Relevant to AI-based scraping, profiling, identity misuse, and unauthorized processing or disclosure.
Telecommunications and Pornography Laws	Misuse of networks and prohibited pornographic content.	Supports prosecution where AI-enabled conduct uses telecommunication infrastructure or creates unlawful synthetic content.

The table shows that existing law governs legal interests and consequences, not AI architecture. Therefore, criminal enforcement must begin with identifying the relevant offense, followed by attribution of conduct and culpability. The use of the term AI cannot replace proving the elements of a crime. Conversely, the technical complexity of AI should not be used to eliminate liability if human or corporate involvement can be proven.

Regulatory Attribution to Human and Corporate Actors

The National Criminal Code still places humans as the primary legal subject. If a user intentionally instructs a user to create a fake voice and use it in a fraudulent manner, then AI functions as a tool. Faults can be traced to the intent, command pattern, victim selection, deployment actions, and expected benefits. The situation becomes more complex when the system produces outputs that are not fully predicted. In such circumstances, the law needs to distinguish truly unpredictable risks from those that are known, foreseeable, or intentionally ignored.

Developers cannot be held liable simply for creating a general-purpose technology. Liability becomes relevant when developers intentionally embed malicious functionality, collaborate with perpetrators, remove safeguards to enable prohibited uses, or ignore specific risks they are obligated to undertake. Service providers can be assessed based on access controls, usage policies, abuse monitoring, log retention, and response to reports. Deployers, or organizations implementing systems, are responsible for configuration, data selection, intended use, operator oversight, and output control within their operational environment.

Users are the parties closest to the direct action when AI is used to create, distribute, or use the object of a crime. However, this proximity does not eliminate the possibility of involvement of other parties. Parties providing facilities, instructions, technical assistance, account access, victim data, or distribution support can be analyzed using the concept of participation based on their contribution. If the act is committed for and within the scope of corporate activities, accountability can also be directed to the corporation and its management based on the decision-making structure, benefits, oversight, and compliance culture.

This layered approach aligns more with the principle of culpability than the single-responsibility model. Each actor is tested individually or organizationally: what actions were taken; what obligations were incurred; what risks were known; what controls were in place; whether the action or omission contributed to the outcome; and whether there was justification or excusal. The results of the tests can vary from actor to actor, so not all parties should be punished on the same basis or with the same degree of culpability.

From a legal certainty perspective, the main weakness of the current regulation is the lack of an operational definition of AI, risk classification, documentation requirements, and minimum standards for developers and providers. This gap does not necessarily preclude the application of existing offenses, but it does complicate the proof of negligence because the due care requirement is not clearly defined. Specific regulations are needed, particularly to establish obligations for prevention, auditing, record-keeping, incident reporting, and cooperation with law enforcement, which can then form the basis for assessing culpability in criminal cases.

2. Implementation of Criminal Liability

Implementation of criminal liability begins with determining the prohibited act, not the AI label. Authorities must identify whether the act involves fraud, forgery, extortion, defamation, illegal access, manipulation of electronic information, data breaches, or prohibited content. The evidence then focuses on who committed the act, how the AI was used, the purpose of the use, and the causal relationship to the harm. This step prevents the use of AI from being treated as a new offense without legal basis while ensuring its technical nature is taken into account.

Electronic evidence plays a central role. Relevant data can include command history, account logs, metadata, network addresses, payment records, device data, model versions, configurations, editing history, and content distribution paths. With deepfakes, it's not enough to simply prove the content is fake; investigators also need to link the creation and distribution process to the perpetrator. The integrity of the evidence must be maintained through documentation of acquisition, storage, examination, and analysis to ensure its validity in court.

Proving difficulties arise because AI services can be cloud-based, use fake accounts, store data outside of jurisdictions, or fail to retain logs long-term. Generative systems can also produce different outputs for the same instructions, making technical reconstructions less straightforward. In such cases, the assessment must combine technical evidence with contextual evidence: the perpetrator's communications, motive, preparation, target selection, account control, transaction patterns, and subsequent actions. Technical evidence does not stand alone but is placed within a framework that explains intent and contribution.

Intentionality can be proven when an actor knowingly uses AI to achieve a prohibited effect. Negligence becomes relevant when the actor had a duty of care, knew or should have known of the risk, had the ability to prevent it, but failed to take reasonable action. However, the application of negligence requires a clear standard. Without specific technical standards or regulations, assessments can be overly subjective. Therefore, governance guidelines and security obligations play a crucial role not only in prevention but also in providing a normative measure for criminal prosecution.

In systems that appear autonomous, autonomy does not automatically sever the causal link to humans. Investigations must examine who sets the goals, selects the data, activates the system, changes parameters, authorizes access, and decides on the use of outputs. If consequences arise despite all reasonable obligations being met and are unforeseeable, criminal penalties should not be imposed. Conversely, if autonomy is used as an excuse to cover up risky design or deliberate omissions, the law must be able to penetrate claims that "the machine did it."

Institutional Implementation and Actor Mapping

Based on legal system theory, implementation faces challenges in substance, structure, and culture. Substantively, regulations do not yet define the division of responsibilities within the AI chain. Structurally, case management requires coordination between investigators, prosecutors, digital forensic experts, AI experts, courts, ministries, and service providers. Culturally, the public often disseminates content without verification, while some industry players have not yet prioritized documentation, transparency, and incident reporting as key elements of governance.

Actor mapping is necessary from the investigation stage. This mapping helps determine what data should be secured and who should be requested for information. Investigators need to distinguish between the base model developer, interface provider, application developer, system implementation organization, operator, end user, account holder, beneficiary, and the party who disseminates the results. Without mapping, investigations risk focusing solely on the final content and missing evidence regarding the decisions that enabled the crime.

Table 2. Role-Based Attribution in the AI Ecosystem

Actor	Relevant Conduct or Duty	Possible Basis of Attribution
Developer	Designs model functions, safeguards, and technical limitations.	Intentional facilitation, participation, or culpable disregard of a specific foreseeable risk.
Provider / Platform	Controls access, policies, logging, moderation, and incident response.	Knowledge, control, assistance, omission despite a legal duty, or corporate policy.
Deployer / Organization	Selects purpose, data, configuration, operators, and operational controls.	Fault in deployment, supervision, data use, security, or organizational governance.
User / Operator	Gives instructions, creates or uses outputs, selects victims, and disseminates content.	Direct intent, negligence, attempt, participation, or use of AI as an instrument of crime.
Corporation / Management	Receives benefit, establishes policies, supervises personnel, and controls compliance.	Corporate attribution, management decisions, failure of supervision, or a culture of non-compliance.

This table does not formulate new offenses, but rather provides a framework for investigation and analysis. The basis for criminal liability must still be established within applicable legal provisions. Mapping helps assess whether an actor is a direct perpetrator, participant, accomplice, responsible administrator, or has no sufficient culpable connection. This way, law enforcement can avoid generalizing that all developers are responsible or that only end users are subject to criminal penalties.

Implementation also requires standard procedures for data preservation and information requests from electronic system providers. Time is crucial because logs can be deleted and content can be moved. Cooperation with providers must maintain data protection, legal authority, and restrictions on information use. In court, experts are needed to explain how the system works in a way that is understandable without displacing judges in assessing the elements of a crime and culpability.

3. Obstacles to Effective Enforcement

The first obstacle is the lack of specific regulations. Disparate laws allow for enforcement against various consequences, but they fail to define AI, classify systems, require documentation, and limit the liability of each actor. Consequently, law enforcement must interpret general norms. Interpretation remains possible, but differences in understanding can lead to disparities, particularly in assessing the negligence of developers or providers who did not directly engage in the act.

The second obstacle is the complexity of proving the truth. Synthetic content can closely resemble the original object, while its creation process involves different models, data, devices, and services. Perpetrators can use anonymous identities, private networks, temporary accounts, or difficult-to-trace payments. In addition to content authenticity, authorities must prove account ownership or control, the source of instructions, the perpetrator's intent, causal relationships, and the integrity of the evidence. Limited access to logs or models can exacerbate the difficulty.

The third obstacle is institutional capacity. Handling cases requires knowledge of digital forensics, cybersecurity, AI architecture, data processing, and the characteristics of electronic evidence. Expertise and resources are not always evenly available. Coordination can also be hampered because cases involve multiple institutions and electronic system administrators. Without a clear division of roles, data requests, device inspections, and technical analysis risk being delayed, resulting in evidence being lost or not being used optimally. The fourth barrier is the transnational nature of AI. AI services can be developed in one country, operated from another, store data in multiple locations, and be used to target victims in Indonesia. Differences in jurisdiction, data protection standards, log

retention policies, and legal aid procedures present challenges. Obtaining information from foreign providers can also take time, while speed is crucial in digital investigations.

The fifth barrier is digital culture and literacy. People who are unable to recognize deepfakes or fraudulent patterns are more likely to fall victim. Victims may also delete conversations, transfer funds without documentation, or report them late. On the other hand, the use of AI without ethical understanding encourages the normalization of identity manipulation and the spread of fake content. Low awareness of data security increases the data sources available for training or generating synthetic content.

These five obstacles reinforce each other. A lack of norms weakens the need for caution; limited capacity slows down the process of establishing evidence; the transnational nature of data hinders access; and a weak digital culture increases the number and impact of abuse. Therefore, solutions must be systemic. Increasing criminal penalties without improving technical standards, evidentiary procedures, cooperation, and literacy will not result in effective law enforcement.

4. Legal and Institutional Solutions

The first solution is the establishment of a specific regulatory framework for AI that aligns with the National Criminal Code. Regulations should include functional definitions, risk classifications, principles of use, actor rights and obligations, security standards, documentation, audits, incident reporting, and oversight mechanisms. Criminal provisions should be directed at serious and clearly defined offenses, while administrative and technical obligations should be used to establish deterrence and precautionary measures. This approach avoids overcriminalization while closing accountability gaps.

The second solution is harmonization with the ITE Law, the Personal Data Protection Law, the National Criminal Code, and sectoral regulations. Harmonization is necessary to prevent overlapping or unclear jurisdiction within a single act. Specific AI norms should clarify the relationship with existing crimes, including when AI is used for forgery, fraud, extortion, defamation, pornography, illegal access, or data misuse. Prosecution guidelines can help determine the most appropriate article based on the dominant legal interest and consequences.

The third solution is strengthening the capacity of law enforcement officials. Education and training programs should include generative AI, deepfake analysis, log preservation, device and cloud forensics, transaction tracing, data protection, and expert examination. Capacity building should be supported by laboratories, analytical tools, evidence management protocols, and expert networks. Developing joint technical guidelines between the police, prosecutors, and courts will reduce disparities in the treatment of electronic evidence.

The fourth solution is to strengthen national and international cooperation. At the national level, governments, law enforcement, academia, industry, and civil society need to establish information-sharing mechanisms on methods, technical indicators, and incidents. At the international level, mutual legal assistance, law enforcement cooperation, and operational relationships with global providers need to be accelerated. Agreements on data preservation and emergency contact points are crucial because digital evidence can be lost before formal procedures are completed.

The fifth solution is improving digital literacy and industry governance. The public needs to understand how to verify content sources, protect personal data, preserve evidence, and report incidents. The industry must implement security by design, including risk testing, usage restrictions, monitoring, complaint channels, and appropriate transparency. These preventative measures reduce the opportunity for criminal activity and provide a trail that assists law enforcement in the event of violations.

The overall solution must maintain a balance between protecting society and fostering innovation. Regulation should not assume all AI is dangerous, but should differentiate between the level of risk and the capacity of actors. High-risk systems require stricter controls, while low-risk uses can be regulated proportionately. The key principle remains the same: the greater the control, benefits, and ability to prevent risks, the greater the duty of care that can be imposed.

5. An Integrated Role-Based Liability Model

Based on the analysis, this study proposes a multi-layered liability model. The model begins with the identification of the offense and the legal interest violated. The next step is to map the actors, actions, control, knowledge, causal contribution, benefit, and duty of care. Each actor is then tested based on the relevant form of fault: intent, negligence, participation, assistance, or corporate misconduct. The model does not establish automatic liability but rather provides a sequence of analysis to ensure that technological complexity does not obscure legal elements.

The first layer is the responsibility of the user or direct operator. This liability is strong if the user designs the instructions, selects the target, generates the content, and uses it for illicit purposes. The second layer is the responsibility of the facilitator. The developer or provider may be liable if they knowingly provide features, access, or support for a specific crime or continue to assist after knowing about the illicit use. The third layer is liability for governance failures, which can only be applied when there is a clear legal obligation, foreseeable risk, preventability, and negligence contributing to the outcome.

The fourth layer is corporate accountability. The analysis focuses on whether the crime was committed within the scope of corporate activities, provided benefits, implemented policies, or occurred due to failures in oversight and compliance. Corporations cannot hide behind a division of labor that makes no one appear to control the entire process. Conversely, corporate criminal charges should not be imposed simply because their products are used for crime. The link between policies, control, benefits, and organizational misconduct must still be proven.

This model rejects the imposition of criminal liability on AI. Making AI an agent raises issues regarding culpability, capacity to accept blame, the purpose of punishment, and the implementation of sanctions. More importantly, it could diminish human accountability. Technical autonomy should encourage more detailed scrutiny of the decisions that shape the system, not shift responsibility to technological objects. AI can be a source of evidence, a tool, an object, or a condition influencing actions, but the subjects of the crime remain human and corporate.

The layered model also provides a foundation for prevention. Actors with high levels of control must maintain logs, manage access, respond to reports, and implement appropriate safeguards. Actors using the system must comply with prohibitions and not exploit output to harm others. Corporations must ensure policies, training, and oversight are in place. These obligations will clarify the extent of negligence and improve the quality of evidence, ensuring that governance regulations and criminal law enforcement are mutually reinforcing.

Normatively, this model is consistent with the principle of legality because punishment is imposed only based on the applicable offense. It is also consistent with the principle of culpability because any attribution requires evidence of blameworthy mental attitude or negligence. From a legal certainty perspective, the parameters of role, control, knowledge, contribution, benefit, and obligation provide a structure for assessment. From a utility perspective, the model encourages prevention by those best able to mitigate risks without hindering the legitimate use of AI.

Policy Implications

The first implication is the need to distinguish between criminal and governance norms. Not every violation of AI standards requires criminal prosecution. Administrative violations can be addressed through remedial orders, administrative fines, service restrictions, or audit requirements. Criminal penalties are directed at acts with a high degree of culpability and harm, such as deliberate use for fraud, blackmail, identity manipulation, or serious data breaches. This differentiation maintains proportionality and prevents criminal law from becoming the sole instrument for regulating technology.

The second implication is the need for documentation standards. Without logs, change logs, and configuration information, proof will always rely on conjecture. Documentation requirements must be proportionate to the risk and consider data protection. Such standards could include administrator identities, system versions, relevant data sources, access settings, mitigation actions, and incident reports. Information does not have to be fully disclosed, but it should be available to authorized authorities through legitimate procedures.

The third implication is the creation of attribution guidelines for law enforcement. These guidelines could address, at a minimum, the following questions: what offense is alleged; who the actors are in the system chain; who has control; what evidence demonstrates knowledge or intent; what obligations were breached; how the actions contributed to the consequences; and whether the corporation benefited. Guidelines do not replace case assessment, but they do increase consistency and reduce the risk of criminal prosecution based solely on technical proximity to the system.

CONCLUSION

Criminal liability regulations for AI misuse in Indonesia have not yet been established under a specific regime. The National Criminal Code and sectoral laws can be used to prosecute specific acts and consequences, but they do not provide a comprehensive division of liability between developers, providers, deployers, users, administrators, and corporations. AI has not and should not be classified as a subject of criminal law. Accountability must remain directed at humans or corporations based on the elements of offense, fault, control, causal contribution, benefit, and duty of care.

Implementation faces challenges such as difficulties in authenticating and attributing evidence, cross-border systems, limited logs, digital forensic capacity, institutional coordination, and public literacy. The study proposes a layered model that begins with crime identification, actor mapping, and individual culpability testing. This model avoids impunity by emphasizing machine autonomy while simultaneously preventing absolute accountability for all parties in the AI ecosystem.

Legal reforms need to be implemented through risk-based AI regulations, harmonization with the National Criminal Code and sectoral regulations, strengthening the capacity of authorities, evidence preservation procedures, international cooperation, and digital literacy. Regulations must differentiate administrative and criminal sanctions proportionally. With this approach, national criminal law can provide certainty, justice, and effective protection without hindering responsible technological innovation.

REFERENCES

Ali, M. (2022). Dasar-dasar hukum pidana. Sinar Grafika.

Arief, M. I. (2024). Unsur-unsur tindak pidana dan teknik penerapan pasal KUHP (Undang-Undang Nomor 1 Tahun 2023). MCL Publisher.

Budiman, T. N., & Hernawati, R. A. S. (2026). Pertanggungjawaban pidana pengembang artificial intelligence atas penyalahgunaan teknologi deepfake di Indonesia. Cessie: Jurnal Ilmiah Hukum.

Chazawi, A. (2022). Pelajaran hukum pidana. Rajawali Pers.

Djoen, T. M., & John, T. (2025). Pertanggungjawaban pidana atas manipulasi foto berbasis kecerdasan buatan bermuatan pornografi dalam hukum pidana Indonesia. Jurnal Ilmu Hukum, Humaniora dan Politik.

Fuady, M. (2022). Teori hukum pidana modern. Citra Aditya Bakti.

Hamzah, A. (2022). Asas-asas hukum pidana Indonesia. Rineka Cipta.

Haris, M. T. A. R., & Tantimin. (2022). Analisis pertanggungjawaban hukum pidana terhadap pemanfaatan artificial intelligence di Indonesia. Jurnal Komunikasi Hukum, 8(1), 307-316.

Hartanto, & Ismail, H. M. (2022). Kitab Undang-Undang Hukum Pidana 2022. LPMI.

Hiariej, E. O. S. (2023). Prinsip-prinsip hukum pidana (Edisi revisi). Cahaya Atma Pustaka.

Ibrahim, A. (2023). Hukum siber dan kejahatan digital di Indonesia. Kencana.

Kadir, A. (2024). Hukum tindak pidana khusus dalam KUHP Nasional. Minhaj Pustaka.

Ramadhani, M., Hidayat, R., & Hosnah, A. U. (2025). Pertanggungjawaban pidana atas tindak pidana yang dihasilkan oleh sistem AI otonom (self-learning algorithm): Kajian terhadap asas kesalahan dalam hukum pidana Indonesia. Jurnal Review Pendidikan dan Pengajaran.

Rodliyah, & Salim, H. S. (2024). Pengantar hukum pidana mengacu pada KUHP baru (UU Nomor 1 Tahun 2023). Sinar Grafika.

Sativa, T. I., et al. (2025). Pertanggungjawaban pidana terhadap kejahatan berbasis artificial intelligence dalam kerangka KUHP baru. Causa: Jurnal Hukum dan Kewarganegaraan.

Sjahdeini, S. R. (2022). Pertanggungjawaban pidana korporasi (Edisi pembaruan). Grafiti.

Sofyan, A. M., & Azisa, N. (2023). Hukum pidana Indonesia. Prenadamedia Group.

Wijanarko, S., Ekaningsih, L., & Farida, A. (2025). Konstruksi pertanggungjawaban pidana korporasi pengembang kecerdasan buatan terhadap tindak pidana pemalsuan dokumen elektronik. Lex Journal: Kajian Hukum dan Keadilan, 10(2).

Republik Indonesia. (1999). Undang-Undang Nomor 36 Tahun 1999 tentang Telekomunikasi.

Republik Indonesia. (2008). Undang-Undang Nomor 44 Tahun 2008 tentang Pornografi.

Republik Indonesia. (2022). Undang-Undang Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi.

Republik Indonesia. (2023). Undang-Undang Nomor 1 Tahun 2023 tentang Kitab Undang-Undang Hukum Pidana.

Republik Indonesia. (2024). Undang-Undang Nomor 1 Tahun 2024 tentang Perubahan Kedua atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik.