Adaptive Criminal Justice Governance For Artificial Intelligence-Driven Crime In Indonesia

Ani Purwati[1]*, Liza Marina[2], Saiful Anam[3], Taryono[4]

[1,2,3,4]Lecturer, Postgraduate Program, Faculty of Law, Universitas Sahid, Indonesia
*Corresponding Author:
Email: anipurwati@usahid.ac.id

Abstract.
Artificial intelligence (AI) changes the conditions under which criminal conduct is planned, executed, concealed, and investigated. Its significance lies not only in enabling new forms of offending but also in distributing relevant actions and information across users, developers, deployers, platforms, and transnational infrastructures. This article examines how Indonesian criminal justice can respond to AI-mediated crime without weakening legality, culpability, due process, and human rights. It employs normative legal research through statutory, conceptual, and functional-comparative approaches, analysing Indonesian criminal law, electronic information regulation, personal data protection, international AI and cybercrime instruments, and selected governance approaches. The analysis identifies five interconnected deficits: normative fragmentation, distributed attribution, synthetic-evidence vulnerability, jurisdictional dispersion, and institutional adaptation lag. In response, the article develops Adaptive Criminal Justice Governance (ACJG), an ecosystem-centred framework organised around three governance layers normative, institutional, and technological-network—and six operational pillars. Its original analytical contributions are distributed attribution, chain of provenance, constrained adaptability, and continuous institutional learning. Distributed attribution uses control, knowledge, foreseeability, and contribution as investigative and analytical dimensions, not as substitute elements of criminal offences. Chain of provenance supplements chain of custody by tracing the creation and transformation of digital objects before and after acquisition. The article argues that criminal justice should preserve a stable normative core while adapting institutional capabilities under enforceable rights-based constraints. For Indonesia, the priority is therefore integrated institutional capacity and evidentiary reliability, rather than automatic expansion of technology-specific criminal offences.
Keywords: *artificial intelligence; criminal justice; distributed attribution; digital evidence and adaptive governance..*

I. INTRODUCTION

Artificial intelligence (AI), particularly generative and agentic systems, is changing the operational conditions of contemporary criminality. Systems capable of generating text, images, audio, video, code, and synthetic identities can be repurposed for deception, impersonation, social engineering, automated fraud, malicious code development, and the manipulation of information. The relevant transformation is not that every AI-enabled harmful act constitutes a new offence. Rather, AI may increase the scale, speed, credibility, automation, and concealment of conduct already prohibited by criminal law. King et al. (2020) identify AI crime as an interdisciplinary problem requiring attention to both foreseeable threats and institutional responses. Criminal law nevertheless remains organised around legally defined conduct, culpability, and attributable responsibility. AI complicates the factual pathway between an actor and a harmful result. A single incident may involve a user who prompts a model, a provider that develops or hosts it, a deployer that integrates it into a service, a platform that distributes the output, and infrastructure providers that process or store relevant data. Their technical connection to the event does not establish equivalent knowledge, intent, control, or legal responsibility. A defensible response must therefore avoid two symmetrical errors: treating technological complexity as a reason to abandon accountability, and treating mere proximity to an AI system as a sufficient basis for criminal liability.

A second difficulty concerns digital evidence. A conventional chain-of-custody inquiry is indispensable for documenting seizure, handling, transfer, storage, and examination. Yet synthetic media creates a prior question: whether the object was authentic before it entered official custody. A convincingly rendered voice or video may be synthetic from inception or altered before acquisition. The evidentiary inquiry must consequently include the origin and production history of the object, relevant transformations, metadata, system interactions, and the basis for associating it with a person. This article calls that expanded inquiry a chain of provenance. It supplements rather than displaces established forensic preservation practices, including the identification, collection, acquisition, and preservation principles contained in ISO/IEC 27037 (International Organization for Standardization [ISO], 2012). Third, AI-mediated conduct

may be transnational by design. The user, victim, model provider, hosting infrastructure, payment channel, and stored evidence may be located in different jurisdictions. Criminal justice therefore depends on lawful access to electronic evidence, timely preservation, mutual legal assistance, cooperation with service providers, and compatible procedural safeguards. The United Nations Convention against Cybercrime, adopted in 2024, provides an emerging framework for international cooperation and electronic evidence in serious crime investigations. However, its legal status must be distinguished from its adoption: the Convention should not be described as being in force while its entry-into-force threshold remains unmet (United Nations, 2024).

Indonesia's legal setting has undergone material change. The Criminal Code under Law No. 1 of 2023 and the Criminal Procedure Code under Law No. 20 of 2025 took effect on 2 January 2026 (Republic of Indonesia, 2023, 2025). Law No. 1 of 2026 on Criminal Penalty Adjustment also forms part of the updated penal framework (Republic of Indonesia, 2026). Alongside these instruments, Law No. 1 of 2024 amended the Electronic Information and Transactions Law, while Law No. 27 of 2022 governs personal data protection (Republic of Indonesia, 2022, 2024). These instruments provide relevant legal foundations, but their coexistence does not by itself guarantee coordinated investigation, forensic validation, cross-border evidence handling, or rights-based use of AI by public authorities. Existing scholarship is substantial but distributed across partially separate conversations. AI-crime scholarship maps emerging threats (King et al., 2020); AI governance instruments emphasise trustworthy systems, accountability, transparency, safety, and human rights (Organisation for Economic Co-operation and Development [OECD], 2024; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2021); digital-forensic standards address identification, acquisition, preservation, and integrity (ISO, 2012); and adaptive-governance scholarship explains institutional learning under complex and changing conditions (Folke et al., 2005). These bodies of work do not, individually, supply a complete criminal justice architecture that links offence definition, individualised attribution, synthetic evidence, institutional coordination, transnational cooperation, and constitutional safeguards.

The research gap is therefore architectural rather than simply legislative. The question is not only whether existing offences can cover conduct involving AI. It is whether criminal justice institutions can detect, attribute, investigate, validate, coordinate, adjudicate, review, and learn from AI-mediated events while remaining constrained by legality and fundamental rights. This article addresses three research questions: (1) what structural limitations does AI-driven crime expose in the Indonesian criminal justice system; (2) how should responsibility and digital-evidence governance be analysed when relevant conduct is distributed across multiple actors and systems; and (3) how can institutions adapt without weakening legality, culpability, due process, proportionality, privacy, and judicial oversight? The article proposes Adaptive Criminal Justice Governance (ACJG). Its novelty is not the invention of a new offence category or a claim that AI itself should be treated as a criminal subject. Rather, it reconstructs the unit of analysis from an isolated offence or technology to the criminal justice governance ecosystem. ACJG integrates five structural deficits, three governance layers, six operational pillars, and a continuous institutional learning cycle. Its central proposition is that criminal justice must remain normatively stable while becoming institutionally adaptive.

II. METHODS

This article employs normative legal research to examine the coherence, interpretation, and institutional adequacy of legal norms governing artificial intelligence (AI)-mediated crime in Indonesia. Normative legal research is appropriate because the principal concern of this study is the examination of legal rules, principles, doctrines, and institutional arrangements rather than the empirical measurement of AI-related criminal activity. The analysis focuses on how existing legal frameworks can accommodate technological developments while maintaining legality, individual culpability, due process, and fundamental human rights. Three complementary approaches are applied: the statutory approach, the conceptual approach, and the functional-comparative approach. First, the statutory approach examines the Indonesian legal framework governing criminal responsibility, criminal procedure, electronic information, and personal

data protection. The principal instruments include Law No. 1 of 2023 concerning the Criminal Code, Law No. 20 of 2025 concerning the Criminal Procedure Code, Law No. 1 of 2026 concerning Criminal Penalty Adjustment, Law No. 1 of 2024 concerning the Second Amendment to the Electronic Information and Transactions Law, and Law No. 27 of 2022 concerning Personal Data Protection (Republic of Indonesia, 2022, 2023, 2024, 2025, 2026).

These instruments are examined to identify the legal foundations, regulatory intersections, and potential institutional gaps arising from AI-mediated criminal conduct. Second, the conceptual approach analyses the legal and theoretical concepts underlying criminal responsibility and adaptive criminal justice governance. The analysis covers criminal attribution, corporate criminal responsibility, culpability, digital evidence, chain of provenance, adaptive governance, proportionality, and rights-based accountability. In particular, the concept of adaptive governance is used to understand how institutions respond to complex and changing environments through coordination, institutional learning, and adjustment of governance arrangements (Folke et al., 2005). This conceptual analysis provides the foundation for developing the proposed Adaptive Criminal Justice Governance (ACJG) framework. Third, the functional-comparative approach examines selected international and national governance instruments to identify institutional practices relevant to the Indonesian criminal justice system.

The European Union is examined through its risk-based regulatory approach under the Artificial Intelligence Act (European Parliament and Council of the European Union, 2024). Singapore is considered through its Model AI Governance Framework, particularly its emphasis on organisational accountability, transparency, and practical implementation (Infocom Media Development Authority & Personal Data Protection Commission, 2020). The United States is examined through the National Institute of Standards and Technology Artificial Intelligence Risk Management Framework, which provides a lifecycle-based approach organised around the functions of Govern, Map, Measure, and Manage (National Institute of Standards and Technology [NIST], 2023).

International governance principles are further considered through the OECD Recommendation on Artificial Intelligence and the UNESCO Recommendation on the Ethics of Artificial Intelligence, particularly their emphasis on human-centred values, transparency, accountability, safety, and human rights (Organisation for Economic Co-operation and Development [OECD], 2024; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2021).

The comparative analysis is functional rather than transplantation. The selected jurisdictions and instruments are not treated as legal models that can be adopted wholesale in Indonesia. Instead, their respective governance functions are assessed in relation to Indonesia's constitutional principles, criminal-law framework, institutional mandates, and procedural safeguards. This approach enables the study to identify transferable governance principles without disregarding differences in legal systems and institutional structures. The research proceeds through four analytical stages. First, it maps the applicable national and international legal frameworks governing AI-related criminal conduct, criminal responsibility, digital evidence, and institutional accountability. Second, it identifies five interconnected structural deficits: normative fragmentation, distributed attribution, synthetic-evidence vulnerability, jurisdictional dispersion, and institutional adaptation lag. Third, it compares selected international governance approaches to identify relevant regulatory, institutional, and technological practices. Fourth, it reconstructs the findings into the Adaptive Criminal Justice Governance (ACJG) framework, consisting of three governance layers, six operational pillars, and a continuous institutional learning cycle.

The proposed framework is subsequently assessed against fundamental principles of criminal justice, including legality, individualised culpability, the presumption of innocence, due process, proportionality, privacy, equality, human dignity, and judicial supervision. Particular attention is given to ensuring that technological capabilities do not become independent grounds for criminal liability or unrestricted investigative authority. Accordingly, the four dimensions of distributed attribution—control, knowledge, foreseeability, and contribution are employed as analytical instruments for investigation and legal reasoning. They do not establish new criminal offences, replace statutory elements, or eliminate the requirement to prove individual culpability under applicable criminal law. This study is exclusively normative in nature. It

does not claim to have conducted interviews, surveys, field observations, or empirical evaluations of institutional performance. Its principal output is a conceptual and normative reconstruction of criminal justice governance designed to strengthen institutional responsiveness to AI-driven crime while preserving the rule of law and fundamental rights.

III. RESULTS AND DISCUSSION

3.1. The Evolving Landscape of AI-Driven Crime: Operational Transformations and Structural Challenges

Artificial intelligence (AI)-driven crime is best understood as a transformation in the conditions under which criminal conduct is planned, executed, concealed, and investigated, rather than as a single homogeneous category of criminal offences. The development of generative and agentic AI systems has expanded the technical capabilities available to individuals and organised groups, particularly in generating synthetic content, automating interactions, imitating identities, and processing information at scale. These developments create new opportunities for the commission of unlawful conduct while simultaneously increasing the complexity of criminal investigation and attribution. King et al. (2020) explain that AI-related crime requires an interdisciplinary understanding because technological capabilities may generate foreseeable threats that cannot be adequately addressed through conventional approaches to criminal enforcement alone. The transformation of AI-driven crime can be identified through five interconnected operational characteristics. First, scale refers to the ability of AI systems to generate, reproduce, and distribute harmful content or fraudulent communications to a large number of targets. Second, speed concerns the compression of the time required to plan, execute, and disseminate criminal conduct. Third, sophistication refers to the reduction of technical barriers that previously limited the ability of individuals to conduct complex forms of digital offending. Fourth, synthetic credibility concerns the capacity of AI-generated outputs to imitate trusted individuals, documents, voices, images, or events, thereby increasing the potential for deception and impersonation. Fifth, opacity relates to the difficulty of reconstructing the processes through which an output emerges from interactions among users, models, datasets, platforms, and supporting digital infrastructure.

These characteristics should not be interpreted as independent technological phenomena. Their significance lies in the way they interact and reshape the relationship between criminal conduct, actors, evidence, and institutional enforcement. For example, the combination of scale and speed may increase the number of potential victims before an incident is detected. Synthetic credibility may make fraudulent communications more persuasive, while opacity may complicate the reconstruction of the technical processes and human decisions underlying the harmful result. Consequently, the investigation of AI-mediated crime requires more than identifying the immediate user of a technological system. It also requires examining the relevant technical processes, organisational relationships, and legally significant contributions of the actors involved. From a criminal justice perspective, these operational transformations expose five interconnected structural deficits. First, normative fragmentation arises when a single AI-mediated incident engages several legal regimes, including criminal law, electronic information regulation, personal data protection, procedural law, and sector-specific legislation.

The coexistence of these regimes may create interpretative and institutional challenges, particularly where different authorities have overlapping responsibilities or apply different procedural standards. The development of trustworthy AI governance principles by the OECD (2024) and UNESCO (2021) demonstrates the importance of integrating accountability, transparency, human rights, and institutional responsibility into technological governance. However, these governance principles do not independently establish criminal liability, which must remain grounded in applicable statutory provisions. Second, distributed attribution emerges when relevant conduct is distributed among multiple actors, including users, developers, providers, deployers, corporate decision-makers, platforms, and infrastructure operators.

Their respective involvement may differ in terms of knowledge, control, intention, and contribution. Accordingly, technical participation in an AI system cannot automatically be equated with criminal responsibility. The central challenge is to distinguish between actors whose conduct is legally relevant to a prohibited offence and those whose involvement is merely technical or incidental. This distinction is

essential to preserving individualised culpability and preventing the overextension of criminal liability (King et al., 2020). Third, synthetic-evidence vulnerability arises when the authenticity and reliability of digital objects cannot be assumed merely because they have been collected and preserved through formal evidentiary procedures. AI-generated or AI-manipulated images, audio, video, and documents may have been altered before their acquisition by law enforcement authorities. Therefore, conventional chain-of-custody procedures must be complemented by an examination of the origin, creation, transformation, and transmission history of digital objects. International digital-forensic guidance provides an important foundation for identifying, collecting, acquiring, and preserving digital evidence (International Organization for Standardization [ISO], 2012). Nevertheless, the emergence of synthetic media requires additional attention to provenance and the technical limitations of authenticity-detection methods.

Fourth, jurisdictional dispersion occurs when the relevant actors, victims, AI service providers, data centres, financial transactions, and electronic evidence are distributed across different territorial jurisdictions. This situation complicates the identification of applicable jurisdiction, the preservation of electronic evidence, and the coordination of investigative authorities. International cooperation therefore becomes an important component of effective criminal justice governance. The United Nations Convention against Cybercrime (2024) provides an emerging international framework for cooperation and electronic evidence in serious crime investigations. Its contribution must nevertheless be understood in accordance with its treaty status and applicable legal requirements. Fifth, institutional adaptation lag refers to the gap between the rapid development of AI technologies and the capacity of criminal justice institutions to understand, investigate, evaluate, and respond to their consequences. Investigators, prosecutors, forensic experts, judges, and oversight institutions require appropriate technical competence, procedural standards, and coordination mechanisms.

Adaptive governance scholarship emphasises the importance of institutional learning, collaboration, and adjustment in responding to complex and changing environments (Folke et al., 2005). In the context of AI-driven crime, this perspective highlights the need for criminal justice institutions to develop capabilities that are responsive to technological change without abandoning established legal safeguards. The five structural deficits are mutually reinforcing rather than isolated problems. Weak provenance documentation may obscure the attribution of digital conduct. Fragmented institutional responsibilities may delay the preservation of electronic evidence. Jurisdictional dispersion may restrict access to relevant information, while limited institutional expertise may undermine the assessment of forensic reliability. These interactions demonstrate that the challenges posed by AI-driven crime cannot be adequately addressed through isolated legislative amendments or technology-based enforcement measures. Accordingly, the structural transformation of AI-driven crime requires a governance approach that connects substantive criminal law, procedural safeguards, digital evidence management, institutional coordination, and international cooperation. The central concern is not merely whether existing criminal offences can accommodate conduct involving AI, but whether the criminal justice system possesses the institutional capacity to identify legally relevant conduct, establish individual responsibility, validate digital evidence, and protect fundamental rights throughout the enforcement process. This provides the analytical foundation for the proposed Adaptive Criminal Justice Governance (ACJG) framework.

3.2 Indonesia's Criminal Law Framework in the AI Era: Regulatory Fragmentation and Institutional Integration Challenges

The effective date of Law No. 1 of 2023 concerning the Criminal Code is particularly significant for assessing Indonesia's capacity to address AI-mediated crime. The analysis must be grounded in the current national penal framework rather than mechanically extending assumptions derived from the former Criminal Code. The 2023 Criminal Code provides the general principles governing criminal responsibility and recognises corporate criminal liability. This recognition is relevant because AI-mediated offences may involve organisational structures, commercial activities, and decisions distributed across several levels of authority. Nevertheless, the involvement of a corporation or organisation does not automatically establish criminal liability. Liability must remain grounded in an applicable statutory provision, legally relevant conduct, the relevant rules of attribution, and the mental element or other conditions required by the offence

concerned (Republic of Indonesia, 2023). This distinction becomes particularly important when an AI system is involved in the commission of an offence. A corporation may develop, provide, deploy, or operate an AI system without necessarily participating in the unlawful conduct subsequently committed by a user. Conversely, organisational involvement may become legally relevant where evidence demonstrates that particular individuals or corporate structures exercised relevant control, possessed legally material knowledge, or made contributions connected to the prohibited conduct. Accordingly, the analysis of AI-related criminal responsibility should not be based solely on an actor's technical position within the AI ecosystem. It must identify the legally relevant relationship between the actor's conduct and the elements of the offence. This approach preserves individualised culpability while allowing investigators to examine organisational and technological relationships that may otherwise remain concealed.

The procedural framework is equally important. Law No. 20 of 2025 concerning the Criminal Procedure Code provides the procedural basis through which substantive criminal responsibility is investigated, prosecuted, and adjudicated. Its reforms address procedural rights, investigative and prosecutorial powers and coordination, coercive measures, and pretrial review (Republic of Indonesia, 2025). These matters have direct implications for AI-related investigations because the collection and examination of digital evidence may involve personal devices, electronic communications, platform records, cloud infrastructure, and information controlled by service providers. The use of technological capabilities in investigation must therefore remain subject to lawful authority, procedural safeguards, and the possibility of judicial scrutiny. For this reason, the procedural framework should not be interpreted as a general authorisation for unrestricted access to electronic data or the uncritical use of algorithmic outputs in criminal proceedings. Investigators must be able to explain the legal basis for obtaining relevant information, the methods used to acquire and preserve it, and the limitations of the technical procedures applied. Prosecutors must also be able to disclose material information concerning the reliability of digital evidence, while courts must retain responsibility for evaluating its admissibility and probative value. Technical complexity cannot justify reducing the procedural rights of the accused or transferring judicial judgment to forensic software and automated systems. Law No. 1 of 2026 concerning Criminal Penalty Adjustment must also be considered in conjunction with the Criminal Code and other relevant penal provisions. Its position within the broader penal framework makes consistency in the application of sanctions an important consideration, particularly where AI-mediated conduct may simultaneously engage general criminal provisions and sector-specific legislation (Republic of Indonesia, 2026). The existence of multiple potentially applicable provisions requires careful examination of the legal basis of prosecution, the relationship between offences, and the applicable sentencing framework. Without such coordination, the response to technologically mediated crime may generate uncertainty concerning the applicable offence and the proportionality of criminal sanctions.

The Electronic Information and Transactions Law, as amended by Law No. 1 of 2024, remains relevant where AI-mediated conduct involves electronic systems, information, documents, communications, or prohibited digital activities (Republic of Indonesia, 2024). Its relevance lies in the regulation of conduct and activities occurring through electronic environments. However, the involvement of AI does not automatically transform every harmful digital activity into a separate AI-specific offence. The legal assessment must identify the conduct prohibited by the applicable provision and establish the required elements of criminal responsibility. This distinction is necessary to prevent technological novelty from becoming an independent and insufficient basis for criminalisation. The Personal Data Protection Law provides another relevant component of the regulatory framework. Law No. 27 of 2022 is particularly significant where AI-enabled conduct involves unlawful processing of personal data, identity misuse, profiling, or the exploitation of information relating to identifiable individuals (Republic of Indonesia, 2022). AI systems may process or generate information derived from personal data, while synthetic identities and manipulated personal information may be used to facilitate fraud, impersonation, or other unlawful activities. Nevertheless, the application of personal data protection provisions must be distinguished from the determination of criminal liability under other legislation. Each legal basis requires an independent assessment of its scope, prohibited conduct, responsible actors, and applicable legal consequences. Taken together, these instruments demonstrate that Indonesia's regulatory weakness should not be characterised

simply as the absence of laws governing AI-related crime. The more precise problem is an integration and institutional capability gap. Existing legal provisions address different dimensions of the problem, including criminal conduct, procedural authority, electronic information, personal data, and criminal sanctions. However, the existence of these separate regulatory components does not by itself establish an integrated mechanism for detecting AI-mediated offences, attributing responsibility, preserving synthetic or manipulated evidence, coordinating investigations, and ensuring judicially reviewable procedures.

This gap becomes visible in several operational questions. First, responsibility for preserving evidence held by AI providers, digital platforms, and other service providers requires clear procedures and lawful coordination. Second, investigators need methods capable of examining the origin, authenticity, and transformation history of potentially synthetic digital objects. Third, prosecutors require sufficient technical information to assess the reliability and limitations of forensic methods before relying on their results. Fourth, courts must be equipped to evaluate technical evidence without treating algorithmic assessments as conclusive determinations of authenticity or criminal responsibility. Finally, the acquisition and use of personal data must remain consistent with applicable legal authority and procedural safeguards. These challenges demonstrate that the effectiveness of criminal law depends not only on the formulation of substantive offences but also on the institutional capacity to implement them. A fragmented response may result in delayed evidence preservation, inconsistent investigative practices, uncertainty in responsibility attribution, and inadequate protection of individual rights. Equally, an excessive reliance on technological detection tools may produce wrongful attribution where synthetic content is treated as authentic or where technical association is mistaken for criminal participation. Within this context, Adaptive Criminal Justice Governance (ACJG) positions Indonesia's principal challenge as the integration of existing legal norms with institutional, evidentiary, and technological capabilities. The framework does not propose the automatic creation of new AI-specific criminal offences. Instead, it emphasises coordinated legal interpretation, individualised responsibility assessment, reliable digital-evidence procedures, inter-institutional cooperation, and continuous institutional learning. Its central objective is to ensure that the criminal justice system can respond to technological change without compromising legality, culpability, due process, proportionality, privacy, and judicial supervision. Accordingly, the Indonesian regulatory challenge is fundamentally architectural. The issue is how to connect the Criminal Code, the Criminal Procedure Code, criminal penalty regulation, electronic information law, and personal data protection within a coherent system of criminal justice governance. ACJG addresses this challenge by treating legal norms, responsible actors, digital evidence, institutional authority, and technological infrastructure as interconnected components. Such an approach allows institutional adaptation while maintaining the normative stability required by the rule of law.

3.3 International AI and Cybercrime Governance: Normative Standards, Institutional Accountability, and Cross-Border Cooperation

International AI and cybercrime instruments provide important normative and institutional references for developing a criminal justice governance framework capable of responding to technologically mediated offences. These instruments address different dimensions of AI governance, including accountability, transparency, human rights, risk management, and international cooperation. However, they do not establish a uniform system of criminal responsibility. Their relevance to Indonesia must therefore be assessed according to their legal character, regulatory scope, and institutional functions rather than through an assumption that international AI governance standards automatically determine domestic criminal liability. The Organisation for Economic Co-operation and Development (OECD) Recommendation on Artificial Intelligence, as amended in 2024, establishes principles for the development and responsible use of trustworthy AI. These principles include inclusive growth and sustainable development, human-centred values and fairness, transparency and explainability, robustness and safety, and accountability (Organisation for Economic Co-operation and Development [OECD], 2024). From a criminal justice perspective, these principles provide a governance foundation for assessing the responsibilities of organisations that develop, deploy, or operate AI systems. Transparency and accountability are particularly relevant when AI systems are used in activities that may affect individual rights, the production of digital evidence, or decisions within

law enforcement institutions. Nevertheless, the OECD Recommendation is a policy instrument rather than a criminal statute. Its principles can inform institutional standards and regulatory development, but they cannot independently establish criminal offences or replace the statutory requirements for criminal responsibility. UNESCO's Recommendation on the Ethics of Artificial Intelligence similarly places human dignity, human rights, inclusion, transparency, accountability, and human oversight at the centre of AI governance (United Nations Educational, Scientific and Cultural Organization [UNESCO], 2021). Its significance for criminal justice lies in the recognition that AI systems must remain subject to meaningful human responsibility, particularly where their application may affect fundamental rights. In criminal investigations and adjudication, this principle requires caution when automated systems are used for profiling, risk assessment, facial recognition, evidence analysis, or other activities capable of influencing the treatment of individuals. Human oversight should not be reduced to a formal approval procedure. It requires the capacity to understand, question, review, and, where necessary, reject technological outputs. Accordingly, UNESCO's framework supports the ACJG proposition that institutional adaptation must remain constrained by human rights and procedural accountability. A more specific regulatory approach is reflected in Regulation (EU) 2024/1689, commonly known as the Artificial Intelligence Act. The EU AI Act establishes a binding, risk-based framework that differentiates regulatory obligations according to the nature and potential risks of AI systems (European Parliament and Council of the European Union, 2024). Its approach is relevant to criminal justice because certain AI applications in law enforcement and the administration of justice may have significant implications for fundamental rights. The Act demonstrates how regulatory obligations can be differentiated according to system function, risk level, and the role of relevant actors. This provides a functional lesson for Indonesia: governance obligations should reflect the actual risks associated with an AI system rather than treating all systems as equivalent. However, the EU AI Act should not be interpreted as directly governing Indonesian criminal proceedings. Its relevance is comparative and conceptual, particularly in illustrating the relationship between risk-based regulation, institutional responsibility, and rights protection.

The Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law provides another important international reference. The Convention adopts a lifecycle-oriented approach to AI governance and connects the development, deployment, and use of AI systems with the protection of human rights, democratic principles, and the rule of law (Council of Europe, 2024). Its contribution to criminal justice governance is the emphasis on institutional accountability throughout the lifecycle of AI systems. This is particularly relevant where public authorities use AI in investigation, evidence assessment, or other exercises of state power. The Convention reinforces the proposition that technological efficiency cannot be treated as an overriding justification for measures that undermine legality, equality, procedural fairness, or effective remedies. For ACJG, this approach supports the integration of rights-based safeguards into institutional procedures rather than treating human rights review as an external or subsequent consideration. In contrast to the broader AI governance instruments, the United Nations Convention against Cybercrime focuses more directly on international cooperation in combating crimes committed through information and communications technology systems and on the sharing of electronic evidence in serious crime investigations. Adopted by the United Nations General Assembly in 2024, the Convention provides an emerging framework for cooperation in the investigation and prosecution of cybercrime (United Nations, 2024). Its significance for AI-mediated crime arises from the transnational character of many digital activities. A single incident may involve a victim in one country, a user in another, a model provider in a third jurisdiction, and electronic evidence stored across several territories. Effective criminal justice responses therefore depend on lawful evidence preservation, cooperation between competent authorities, and procedures that accommodate differences between national legal systems. As of 8 November 2026, the United Nations Treaty Collection recorded 96 signatories and three parties to the Convention, while its entry into force remained subject to the deposit of the required forty instruments (United Nations, 2024). This status is important because adoption of an international convention does not, by itself, mean that all its provisions have become operational treaty obligations for every state. Its contribution to ACJG should therefore be understood in terms of its developing normative and cooperative

framework, while the applicable legal consequences must be assessed according to its entry into force and the relevant rules governing treaty participation.

Taken together, these international instruments reveal two complementary dimensions of AI-related criminal justice governance. The first concerns the governance of AI systems themselves, including risk management, transparency, accountability, human oversight, and rights protection. The second concerns the investigation and prosecution of offences involving digital technologies, including international cooperation and access to electronic evidence. These dimensions are related but legally distinct. Regulatory obligations imposed on AI developers, providers, or deployers may contribute to prevention, accountability, and evidence preservation. However, a breach of an AI governance standard does not automatically establish criminal liability. Criminal responsibility must remain based on applicable domestic law, legally attributable conduct, and proof of the elements of the relevant offence. For Indonesia, the international frameworks examined above provide comparative and normative guidance rather than a substitute for domestic legislative authority. Their collective contribution lies in demonstrating that effective governance requires more than technical compliance or the expansion of criminal prohibitions. It requires coordination between risk-based regulation, institutional accountability, reliable digital evidence, international cooperation, and enforceable human rights safeguards. ACJG incorporates these lessons through an integrated governance architecture in which normative stability is maintained while institutional capabilities are continuously adapted. This distinction is fundamental: the criminal justice system must be capable of responding to technological change without allowing technological risk assessments or governance standards to displace the principles of legality, individualised culpability, presumption of innocence, due process, proportionality, and judicial supervision.

3.4 Distributed Criminal Attribution in AI-Driven Crime: Reconstructing Individualised Culpability Across the AI Ecosystem

AI-mediated criminal events may involve multiple actors, including users, developers, providers, deployers, corporate decision-makers, digital platforms, and infrastructure operators. These actors may participate at different stages of the technological process, from system development and configuration to deployment, distribution, and data processing. However, their technical or organisational connection to an incident does not establish equivalent criminal responsibility. The identification of an actor's role within an AI ecosystem must therefore be distinguished from the legal determination of that actor's culpability. Criminal responsibility requires an assessment of conduct, the applicable statutory offence, the legally recognised mode of participation, and the relevant mental element or other conditions of liability prescribed by law (Republic of Indonesia, 2023).

This distinction is particularly important because AI-mediated conduct may distribute relevant actions, information, and decision-making authority across several organisational and technological layers. A user may initiate an unlawful instruction, a developer may design the model, a provider may operate the infrastructure, a deployer may integrate the system into a commercial service, and a platform may distribute the resulting content. Although these activities may form part of the same causal sequence, they do not necessarily constitute equivalent contributions to the prohibited conduct. A defensible criminal justice response must avoid two symmetrical errors: treating technological complexity as a reason to abandon accountability, and treating mere proximity to an AI system as sufficient grounds for criminal liability. To address this problem, this article proposes distributed attribution as an analytical framework consisting of four dimensions: control, knowledge, foreseeability, and contribution. First, control refers to an actor's factual and legally relevant capacity to direct, configure, restrict, supervise, or intervene in the operation of an AI system. Control should be assessed in relation to the actor's actual authority and technical capacity, rather than inferred merely from a formal organisational position. An actor who possesses administrative access to a system, for example, may have a different degree of control from an actor who can only use its publicly available interface.

Second, knowledge concerns the actor's actual awareness of relevant circumstances, including the unlawful use of a system, the nature of the conduct being facilitated, or information that may be legally material to the offence under examination. Knowledge must be assessed in accordance with the applicable

statutory requirements. General awareness that an AI model could potentially be misused cannot automatically be equated with knowledge of a particular criminal plan or an intention to facilitate a specific offence. This distinction protects the principle of individualised culpability and prevents criminal responsibility from being inferred solely from an actor's professional or technological involvement. Third, foreseeability concerns whether the relevant consequence could reasonably have been anticipated from the actor's position, available information, and circumstances at the material time. This dimension helps distinguish a general technological risk from a specific and legally relevant risk associated with particular conduct. Nevertheless, foreseeability is not an independent substitute for the mental element required by a criminal offence. The fact that a harmful outcome could have been anticipated does not, by itself, establish criminal intent, negligence, participation, or any other statutory basis of liability. Fourth, contribution examines the actor's conduct or omission and its legally relevant relationship to the prohibited conduct or result. The analysis must identify whether the contribution constitutes conduct recognised by the applicable criminal law, including a legally relevant form of participation or corporate involvement where the relevant statutory conditions are satisfied. A technical contribution, commercial relationship, or organisational connection is not sufficient in itself. The legal significance of an act or omission must be established through the elements and attribution requirements of the offence concerned (Republic of Indonesia, 2023).

These four dimensions are complementary analytical tools rather than independent grounds of guilt. They structure the investigation and legal reasoning by helping authorities distinguish the respective positions of actors within a distributed technological environment. Their application must remain offence-specific and subject to the principles of legality, individualised culpability, and the presumption of innocence. They cannot create a new offence, expand criminal liability beyond statutory limits, or replace the prosecution's obligation to establish the elements of the alleged offence through admissible and reliable evidence. The practical implications can be illustrated through several situations. A developer's general awareness that a model could be misused does not, without additional legally relevant circumstances, establish an intention to facilitate a particular crime. Similarly, a provider's technical capacity to suspend a service does not automatically establish a legal duty to suspend it in every situation. The existence and scope of such a duty must be determined by reference to applicable law and the circumstances of the case. A platform's receipt of a notice concerning unlawful content may be relevant to assessing its knowledge, but receipt of that notice does not automatically demonstrate participation in the underlying offence. By contrast, an actor who knowingly configures or deploys a system specifically to facilitate a defined criminal act occupies a materially different position from a provider supplying a neutral service without the legally required knowledge or contribution.

The same approach applies to corporate and organisational responsibility. The involvement of a company in developing, deploying, or operating an AI system does not eliminate the need to identify the relevant organisational conduct and establish the statutory conditions for corporate criminal liability. The inquiry should distinguish individual conduct from organisational decision-making while examining how authority, knowledge, control, and contribution are distributed within the relevant structure. Corporate involvement must therefore be assessed under the applicable rules of Indonesian criminal law rather than inferred solely from the fact that an offence was committed through a company's technological infrastructure (Republic of Indonesia, 2023). Within the ACJG framework, distributed attribution connects the normative layer, which defines the applicable offence and liability requirements, with the institutional layer, which investigates and evaluates the conduct of relevant actors, and the technological-network layer, which contains records and information necessary to reconstruct their respective roles. This connection requires investigators to identify relevant actors and their authority, preserve evidence of system configuration and interaction, document organisational decision-making where legally relevant, and distinguish verified information from technical inference. The resulting attribution analysis must remain open to challenge by the defence and subject to judicial assessment under the applicable procedural guarantees (Republic of Indonesia, 2025). Accordingly, distributed attribution seeks to prevent both an accountability vacuum and over-attribution of criminal responsibility. It recognises that evidence of AI-mediated conduct may be dispersed across individuals, corporations, platforms, and technical infrastructures, while maintaining that

criminal liability must ultimately attach only to conduct and actors for whom the applicable legal requirements have been established. Its principal contribution is therefore methodological: it provides a structured way to investigate complex technological relationships without transforming technological involvement, risk awareness, or organisational proximity into automatic criminal guilt.

3.5. Beyond Chain of Custody: Establishing Digital Evidence Provenance and Reliability in AI-Driven Crime

The reliability of digital evidence in AI-mediated criminal cases requires an evidentiary approach that extends beyond conventional chain-of-custody procedures. Chain of custody records the handling of an evidentiary object after its collection, including its acquisition, transfer, storage, examination, and preservation. It remains essential for demonstrating continuity, maintaining integrity, and protecting evidence against unauthorised alteration. International forensic guidance, particularly ISO/IEC 27037:2012, establishes principles for the identification, collection, acquisition, and preservation of digital evidence (International Organization for Standardization [ISO], 2012). These procedures provide an important foundation for ensuring that digital evidence is handled systematically and can subsequently be examined and challenged. However, AI-generated and AI-altered materials introduce an additional evidentiary problem. The authenticity of a digital object may already be compromised before it enters official custody. An image, audio recording, video, or electronic document may have been synthetically generated, manipulated, or processed through multiple systems before being obtained by investigators. Consequently, an unbroken chain of custody does not necessarily establish that the original object was authentic. It demonstrates the documented handling of the item after acquisition, but it cannot, by itself, establish the item's origin, production history, or relationship to the event under investigation. To address this limitation, this article proposes chain of provenance as a complementary evidentiary framework. Rather than replacing chain of custody, chain of provenance expands the inquiry to include the lifecycle of a digital object, from its creation or initial generation to subsequent modification, processing, storage, transmission, and acquisition. The framework is organised around four interconnected questions.

First, authenticity examines whether the digital object represents what it purports to represent. The inquiry concerns whether an audio recording genuinely represents the voice attributed to a particular person, whether a video depicts the event it claims to document, or whether an electronic document originated from the person or institution identified as its source. In AI-mediated cases, apparent visual or auditory credibility cannot independently establish authenticity. Second, integrity examines whether the object has been altered and whether any alteration can be identified or reconstructed. Technical mechanisms such as cryptographic hashes may assist in determining whether a digital file has changed between documented stages of acquisition and examination. Nevertheless, a hash value establishes the identity of a particular digital state; it does not independently prove that the file was authentic before it was collected or that its contents accurately represent an actual event. Third, provenance concerns the origin and history of the digital object. Relevant information may include original files, timestamps, metadata, acquisition context, available generation or editing records, platform logs, model or version information where obtainable, and documented transformation history. Such information helps investigators reconstruct how an object was created, processed, modified, stored, and transmitted. The completeness of this record may vary according to the technical environment, the availability of provider-held information, and the circumstances in which the evidence was obtained. Fourth, attribution examines whether the digital object, its creation, or its relevant transformations can be reliably connected to an individual or organisation involved in the alleged offence. This inquiry is particularly important where several actors participate in the same technological process. A file's presence on a device, account, or platform does not automatically establish that a particular person created, modified, uploaded, or knowingly distributed it. Attribution therefore requires the examination of relevant technical records alongside other admissible evidence and the applicable legal requirements for individualised criminal responsibility.

The four questions are interdependent but should not be treated as interchangeable. Authenticity concerns what the object represents; integrity concerns whether it has changed; provenance concerns its origin and transformation history; and attribution concerns its connection to a relevant actor. An object may

retain its integrity after acquisition while remaining inauthentic from its inception. Similarly, a file may have a traceable origin without sufficient evidence to establish who was legally responsible for its creation or distribution. Chain of provenance therefore provides a structured method for evaluating evidentiary reliability rather than a presumption that a digital object is true. Several technical limitations must also be recognised. Metadata may be absent, manipulated, removed, or misleading. AI-generated content detection tools may produce false-positive and false-negative results. Watermarking and content credentials may not be available, may be removed, or may not provide conclusive evidence of authenticity. Model and version information may also be inaccessible where relevant processing occurs through proprietary or third-party systems. These limitations mean that no single technical indicator should be treated as conclusive proof of authenticity or manipulation. Forensic conclusions should instead be evaluated in relation to the methods used, their validation conditions, known error rates where available, the completeness of the chain of custody, alternative explanations, and the possibility of independent or adversarial testing.

The judicial assessment of AI-related digital evidence must consequently distinguish technical findings from legal conclusions. Forensic experts may explain the methods used to examine an object, the available provenance records, the limitations of detection tools, and the degree of technical confidence supported by the examination. However, the determination of evidentiary admissibility, probative weight, and criminal responsibility remains a judicial function. Technical complexity must not transfer the court's legal assessment to an expert or an automated detection system. Equally, uncertainty concerning the origin of digital evidence must be addressed through disclosure, examination, and the opportunity to challenge the evidence, rather than by imposing an obligation on the accused to prove that an object is synthetic. Within the Indonesian criminal justice framework, these requirements must operate consistently with applicable procedural guarantees governing the acquisition, examination, presentation, and evaluation of evidence (Republic of Indonesia, 2025). The prosecution retains the burden of establishing the alleged offence through legally admissible and sufficiently reliable evidence. A dispute over whether an image, recording, or document is AI-generated cannot, by itself, reverse that burden or create a presumption of authenticity against the accused. Within ACJG, chain of provenance connects the technological-network layer with the institutional and normative layers. The technological-network layer provides relevant records and technical information, including available system logs, metadata, generation histories, and provider-held information. The institutional layer ensures that investigators, forensic experts, prosecutors, and courts apply documented procedures and critically evaluate technical findings. The normative layer establishes the legal boundaries for obtaining, disclosing, testing, and using digital evidence while protecting privacy, due process, and the presumption of innocence. Accordingly, chain of provenance represents a conceptual extension of conventional digital-evidence governance. Its contribution is not to guarantee the truth of AI-related digital material, but to make its origin, integrity, transformation history, and attribution more transparent and contestable. By supplementing chain of custody with lifecycle-based evidentiary analysis, ACJG seeks to strengthen the reliability of digital evidence without weakening procedural fairness or shifting the burden of proof. The central principle is that technological sophistication may expand the methods of evidentiary examination, but it cannot diminish the legal standards governing proof and criminal responsibility. **Table 1. Comparative Institutional and Regulatory Frameworks Informing the ACJG Model**

Jurisdiction / instrument	Governance feature	Functional lesson for ACJG
European Union – Artificial Intelligence Act	Binding, risk-based obligations and differentiated duties according to AI system risk and the roles of relevant actors (European Parliament and Council of the European Union, 2024).	Risk classification can guide prevention, institutional oversight, and accountability. However, it cannot replace the requirements of criminal-law proof and individualised culpability.
Singapore – Model AI Governance Framework	Practical organisational guidance addressing AI governance structures, clearly assigned responsibilities, transparency, human oversight, and accountability (Infocomm Media	Effective governance requires operational guidance, clearly assigned institutional responsibilities, and iterative review of AI governance practices.

	Development Authority & Personal Data Protection Commission, 2020).	
United States – NIST AI Risk Management Framework	Voluntary lifecycle risk-management functions organised around Govern, Map, Measure, and Manage (National Institute of Standards and Technology [NIST], 2023).	A common risk-management vocabulary and systematic documentation can strengthen technical accountability, risk assessment, and institutional coordination.
International – United Nations Convention against Cybercrime	International cooperation and electronic-evidence mechanisms for addressing serious crimes involving information and communications technology systems (United Nations, 2024).	Cross-border evidence cooperation must be lawful, rights-protective, procedurally accountable, and consistent with the Convention's actual treaty status.

The comparison is functional rather than transplantational. The EU demonstrates the value of differentiated obligations; Singapore illustrates implementation guidance; NIST supplies a lifecycle risk-management vocabulary; and the UN Convention highlights the importance of international cooperation. None of these instruments, on its own, provides a complete criminal justice model. ACJG connects their relevant functions to domestic criminal-law safeguards and institutional responsibilities.

3.6. Reconstructing Criminal Justice Governance in the AI Era: The Adaptive Criminal Justice Governance (ACJG) Architecture

ACJG is organised into three connected governance layers. The normative layer defines the boundaries of criminalisation, culpability, investigative authority, evidence, and rights, thereby ensuring that technological developments do not displace the principles of legality, proportionality, and due process. The institutional layer concerns investigators, prosecutors, courts, forensic bodies, regulators, and oversight institutions, whose respective mandates must remain clearly defined and accountable. The technological-network layer recognises the practical influence of AI providers, platforms, cloud and telecommunications infrastructure, financial intermediaries, and international partners.

These actors may hold information or capabilities relevant to a case, but their involvement does not itself confer public coercive authority. This distinction reflects the importance of maintaining state responsibility, competent-authority requirements, and procedural safeguards in cross-border criminal investigations (United Nations General Assembly, 2024, arts. 23, 35, and 47; Council of Europe, 2001, art. 15). Six operational pillars connect the layers: (1) adaptive regulation, which reviews genuine normative gaps while avoiding reflexive criminalisation; (2) anticipatory institutional capacity, which develops expertise before risks become systemic; (3) technology-responsive investigation, which uses validated tools and reviewable methods; (4) integrated digital evidence governance, which standardises authenticity, integrity, provenance, attribution, preservation, and disclosure; (5) collaborative criminal justice networks, which coordinate competent institutions and lawful private-sector cooperation; and (6) rights-based accountability, which ensures that technological effectiveness remains subordinate to legality and judicial control. These pillars are consistent with the OECD's emphasis on human rights, transparency, robustness, traceability, accountability, and risk-based governance throughout the AI system lifecycle (OECD, 2024). The evidentiary dimension also draws upon established digital forensic principles concerning validated tools, repeatability, preservation of information integrity, and chain of custody (Kent et al., 2006). The model operates through a recurring cycle: DETECT → ATTRIBUTE → VALIDATE → COORDINATE → REVIEW → LEARN → ADAPT. Detection identifies the event and preserves potentially relevant information. Attribution maps actors and their legally relevant relationships. Validation tests evidence and technical methods. Coordination allocates responsibilities and secures lawful cooperation. Review subjects coercive measures and evidentiary claims to independent scrutiny. Learning evaluates failures, error patterns, and institutional performance. Adaptation updates protocols, training, and, where necessary, legal rules. The cycle is continuous, but every stage remains bounded by law. This iterative structure reflects the principles of

continuous risk management, traceability, institutional accountability, and international cooperation recognised in contemporary AI governance and cybercrime frameworks (OECD, 2024; United Nations General Assembly, 2024).

3.7. Rights-Based Adaptive Criminal Justice Governance: Institutional Accountability, Evidentiary Integrity, and Implementation in Indonesia

Adaptability can itself create risks. Predictive tools may reproduce historical and structural biases embedded in training data; automated triage may conceal the underlying rationale of institutional decisions; broad data collection may interfere with privacy and informational self-determination; and opaque forensic systems may weaken the ability of defendants to examine and challenge the reliability of evidence. These risks demonstrate that technological efficiency cannot automatically be equated with procedural fairness or substantive justice. Research on algorithmic accountability has shown that automated decision-making may reproduce discriminatory patterns and create accountability gaps when the mechanisms underlying computational outputs remain insufficiently transparent (Barocas et al., 2019; Selbst et al., 2019). Similarly, the NIST Artificial Intelligence Risk Management Framework emphasises the need to identify, assess, manage, and continuously monitor risks associated with AI systems throughout their lifecycle (National Institute of Standards and Technology [NIST], 2023).

ACJG therefore adopts constrained adaptability: institutions may change their methods, but they may not relax legality, necessity, proportionality, equality, the presumption of innocence, access to counsel, judicial supervision, or the right to challenge evidence. These safeguards are consistent with international human rights standards, particularly the International Covenant on Civil and Political Rights, which protects equality before the law, liberty, fair trial guarantees, and the presumption of innocence (United Nations, 1966, arts. 9, 14, and 17). Human review must be meaningful rather than a formal signature attached to an automated output. Accordingly, institutional decision-makers must possess sufficient competence, access to relevant information, and actual authority to question, reject, or override algorithmic recommendations. This requirement reflects the OECD's emphasis on human agency, oversight, transparency, and accountability in AI governance (OECD, 2024). For Indonesia, implementation should begin with an inter-agency protocol rather than an assumption that a new AI-specific criminal offence is always required. The protocol should identify lead and supporting authorities, preservation request procedures, minimum provenance documentation, forensic validation requirements, disclosure obligations, expert competence, privacy safeguards, and escalation routes for cross-border evidence. The protocol should be aligned with the Indonesian Criminal Code (Law No. 1 of 2023), the Criminal Procedure Code (Law No. 20 of 2025), the Electronic Information and Transactions Law as amended by Law No. 1 of 2024, the Personal Data Protection Law (Law No. 27 of 2022), and applicable sectoral legislation. This approach recognises that the existing legal framework already provides important foundations for criminal responsibility, criminal procedure, electronic information, and personal data protection, while institutional coordination and technological implementation require further operational development (Republic of Indonesia, 2022, 2023, 2024, 2025).

Second, investigative and prosecutorial capacity should include training on generative systems, synthetic media, model limitations, platform logs, metadata, and the interpretation of forensic outputs. Such training should develop not merely technical familiarity but also the ability to distinguish algorithmic recommendations from independently established evidentiary facts. Third, forensic laboratories should publish validated procedures and limitations, including known error conditions, reproducibility requirements, and circumstances in which particular tools may produce unreliable results. These requirements are consistent with established digital forensic guidance concerning evidence acquisition, preservation, examination, documentation, and integrity verification (Kent et al., 2006). Fourth, courts should receive technically intelligible expert evidence while retaining the exclusive responsibility to determine legal admissibility, evidentiary weight, and guilt. Expert testimony should therefore explain the methodology, assumptions, limitations, error rates where available, and alternative interpretations of forensic findings. Fifth, cooperation with private providers should be documented, purpose-limited, and subject to lawful authority and review. This is particularly important where relevant information is controlled by platforms,

cloud service providers, telecommunications operators, or other private intermediaries. Finally, institutional learning should be measured through indicators such as preservation timeliness, completeness of provenance records, forensic reproducibility, disclosure compliance, successful lawful cooperation, reversal and error patterns, and rights complaints. These indicators translate general accountability principles into measurable institutional performance criteria (NIST, 2023; OECD, 2024). These proposals do not presume that all agencies currently possess the same technical capability or that a single national platform can solve the problem. The recommended sequence is incremental: establish minimum procedural standards; map institutional mandates; build specialised capacity; pilot interoperable provenance records; evaluate safeguards; and revise protocols through transparent review. This sequencing reflects the adaptive governance principle that institutional learning and policy adjustment should proceed through continuous assessment rather than premature technological centralisation (Duit & Galaz, 2008; NIST, 2023). This is how adaptive governance can be made operational without creating an unbounded discretion to investigate. Ultimately, the legitimacy of ACJG depends not on how extensively institutions deploy AI, but on whether technological adaptation strengthens institutional accountability, evidentiary reliability, and the protection of fundamental rights. Table 2. Conceptual Innovations and Normative Foundations of the Adaptive Criminal Justice Governance (ACJG) Model

Conceptual contribution	Analytical shift	Limiting principle
Ecosystem-centred criminal justice	From isolated offence or AI system to the connected legal-institutional-technical environment	The ecosystem is an analytical unit, not a collective criminal subject.
Distributed attribution	From simple actor identification to structured analysis of control, knowledge, foreseeability, and contribution	No dimension replaces statutory elements or culpability.
Chain of provenance	From post-acquisition custody alone to lifecycle tracing of digital objects	Provenance is a reliability inquiry, not a presumption of truth.
Constrained adaptability	From unrestricted flexibility to institutionally adaptive action within rights and rule-of-law limits	No efficiency rationale displaces legality, due process, or proportionality.
Continuous institutional learning	From episodic reform to recurring detection, review, evaluation, and adaptation	Learning must be transparent, reviewable, and rights-compliant.

The theoretical novelty of ACJG lies in combining these five movements into one criminal justice governance architecture. Ecosystem-centred analysis identifies the relevant relationships; distributed attribution structures the responsibility inquiry; chain of provenance addresses the lifecycle of digital evidence; constrained adaptability limits institutional discretion; and continuous learning supplies the feedback mechanism. The concepts are complementary but not interchangeable. In particular, distributed attribution concerns legal reasoning about actors, whereas chain of provenance concerns the reliability and history of evidentiary objects. The principal proposition criminal justice must remain normatively stable while becoming institutionally adaptive expresses a dual requirement. Normative stability protects citizens against shifting or technologically opportunistic penal power. Institutional adaptability prevents that stability from becoming institutional inertia. ACJG therefore seeks to prevent both an accountability vacuum and technological over-criminalisation. Its contribution is a framework for governing the relationship between legal norms, actors, evidence, institutions, technical networks, and rights.

IV. CONCLUSION

AI-driven crime exposes structural limitations that cannot be resolved solely by adding technology-specific offences. The normative analysis identifies five connected deficits: normative fragmentation, distributed attribution, synthetic-evidence vulnerability, jurisdictional dispersion, and institutional adaptation lag. These deficits demonstrate that effective criminal justice depends on the interaction of substantive law,

procedure, evidence, institutional competence, cooperation, and rights protection. The article's principal contribution is Adaptive Criminal Justice Governance. Its three governance layers and six operational pillars connect the stable normative core of criminal law with adaptive institutional and technological-network capabilities. Distributed attribution offers four analytical dimensions—control, knowledge, foreseeability, and contribution—without replacing offence elements or culpability. Chain of provenance supplements chain of custody by tracing the lifecycle of potentially synthetic evidence. Constrained adaptability and continuous institutional learning ensure that institutional responsiveness remains subject to legality and fundamental rights. For Indonesia, the immediate priority is to integrate the Criminal Code, the 2025 Criminal Procedure Code, the 2026 penalty-adjustment framework, electronic information regulation, and personal data protection through operational protocols, validated forensic methods, trained personnel, lawful provider cooperation, and judicially reviewable safeguards. The model should be tested through institutional pilots and subsequent empirical research; this normative article does not claim to have measured implementation outcomes. Criminal justice must remain normatively stable while becoming institutionally adaptive.

ACKNOWLEDGMENTS

The authors acknowledge the academic support of the Master of Law Program, Postgraduate Program, Sahid University, Jakarta, Indonesia. The authors are responsible for the analysis and conclusions expressed in this article.

REFERENCES

[1]. King, T. C., Aggarwal, N., Taddeo, M., & Floridi, L. (2020). Artificial intelligence crime: An interdisciplinary analysis of foreseeable threats and solutions. Science and Engineering Ethics, 26, 89–120. https://doi.org/10.1007/s11948-018-00081-0

[2]. International Organization for Standardization. (2012). ISO/IEC 27037:2012: Information technology—Security techniques—Guidelines for identification, collection, acquisition and preservation of digital evidence. ISO.

[3]. United Nations. (2024). United Nations Convention against Cybercrime: Strengthening international cooperation for combating certain crimes committed by means of information and communications technology systems and for the sharing of evidence in electronic form of serious crimes. A/RES/79/243. United Nations Treaty Collection, status as at 8 November 2026.

[4]. Republic of Indonesia. (2023). Law No. 1 of 2023 concerning the Criminal Code. State Gazette of the Republic of Indonesia Year 2023 No. 1, Supplement No. 6842.

[5]. Republic of Indonesia. (2025). Law No. 20 of 2025 concerning the Criminal Procedure Code. State Gazette of the Republic of Indonesia Year 2025 No. 188, Supplement No. 7149.

[6]. Republic of Indonesia. (2026). Law No. 1 of 2026 concerning Criminal Penalty Adjustment. State Gazette of the Republic of Indonesia Year 2026 No. 1.

[7]. Republic of Indonesia. (2024). Law No. 1 of 2024 concerning the Second Amendment to Law No. 11 of 2008 concerning Electronic Information and Transactions. State Gazette of the Republic of Indonesia Year 2024 No. 1, Supplement No. 6905.

[8]. Republic of Indonesia. (2022). Law No. 27 of 2022 concerning Personal Data Protection. State Gazette of the Republic of Indonesia Year 2022 No. 196, Supplement No. 6820.

[9]. Organisation for Economic Co-operation and Development. (2024). Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449, amended 3 May 2024). OECD.

[10]. United Nations Educational, Scientific and Cultural Organization. (2021). Recommendation on the ethics of artificial intelligence. UNESCO.

[11]. European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union.

[12]. Folke, C., Hahn, T., Olsson, P., & Norberg, J. (2005). Adaptive governance of social-ecological systems. Annual Review of Environment and Resources, 30, 441–473. https://doi.org/10.1146/annurev.energy.30.050504.144511

[13]. Council of Europe. (2024). Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225).

[14]. Infocomm Media Development Authority & Personal Data Protection Commission. (2020). Model artificial intelligence governance framework (2nd ed.). Singapore.
[15]. National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1). U.S. Department of Commerce. https://doi.org/10.6028/NIST.AI.100-1.