

Criminalising AI Harms in ASEAN

How can ASEAN and its member states implement effective accountability for malicious AI-enabled harms, without producing a chilling effect on innovation?

Author: Rusydi Farhan

UMEA AI Policy Lab

1.Introduction

The exponential growth of generative artificial intelligence (GenAI) in Southeast Asia has produced both unprecedented economic opportunities and novel categories of digital harm. The region's digital economy is projected to reach USD 1 trillion by 2030, having attracted substantial investment from global technology firms and businesses (WEF, 2023). Parallel with this growth, GenAI has been weaponised to perpetrate harms at scale, most notably through non-consensual deepfake pornography (NCDP). Various research indicates that the overwhelming majority of deepfake content online is pornographic, with women and children comprising the vast majority of victims, who are subjected to harassment, blackmail, and psychological abuse (Citron & Chesney, 2019). Beyond the harm to direct victims, NCDP erodes trust in visual media and the integrity of the information ecosystem more broadly.

The regulatory response to these harms has been fragmented. The Grok AI incident, in which the xAI system was exploited to generate 'nudifier' content from clothed images, illustrates this divergence (Watkins, 2026). The European Union (EU) initiated formal data protection investigations, while Malaysia and Indonesia responded through platform access restrictions and monitoring, revealing the absence of harmonised ASEAN-level governance standards (O'Regan, 2026). Sophisticated actors can take advantage of exploitable gaps created by the regulatory disparity across nations.

2.Research Question and Scope

This paper examines a fundamental tension at the intersection of economic development and human rights: how can ASEAN and its member states implement effective criminal action for malicious AI-enabled harms without producing a chilling effect on innovation? The paper focuses on Singapore and Indonesia as comparative reference points, analysing how their evolving regulatory approaches and landscapes may inform a coherent supranational framework. In addition, the doctrinal approach revolves on the attribution of criminal responsibility, the standard of culpability, and the allocation of liability across the development and deployment chain, and issues brought forward due to the probabilistic and complex nature of AI systems.

2.1 Methodology

A doctrinal and comparative analysis will be employed to assess the respective frameworks. The doctrinal component examines foundational criminal law principles and their application to AI systems within ASEAN and EU frameworks. The comparative component identifies structural convergences and divergences in how these frameworks are able to conceptualise accountability and criminalisation. Primary sources include legislations such as Indonesia's Government Regulation No. 17 of 2025, Singapore's Online Safety (Relief and Accountability) Bill 2025 and Regulation (EU) 2024/1689 (Artificial Intelligence Act), and relevant soft law instruments. Secondary sources comprise of plethora of peer-reviewed articles addressing issues central to NCDP and the intent to criminalise AI-enabled harms, international organisation reports, and technological policy analyses.

3. Defining the Threat

3.1 Synthetic Media and Inauthentic Material

A legally operative distinction between legitimate synthetic media and malicious content is foundational to any sound criminalisation framework. Synthetic media have substantial and legitimate applications across film production and medical imaging (Koetzier et. al, 2024; Tsiavos & Kitsios, 2025). The critical distinction lies in what may be termed a 'functional transformation': a system operated within legitimate bounds when deployed consistently according to its design and uses, and the reasonable expectations of interested parties (KPMG, 2024). In terms of the creation of NCDP, human actors exploit the generative capability of AI to create materials which causes abuse to others. The Online Safety (Relief and Accountability) Bill 2025 defines "inauthentic material" in s.16(2) as conveying false or misleading information about a victim which a reasonable person would believe is likely to cause "harassment, alarm, distress or humiliation". Similarly, the EU AI Act shares the same sentiment whereby the AI-generated content closely resembles an actual person. Criminal liability attaches at the point where the system is utilised for extortion, sexual abuse and defamation, transforming a creative tool into a weapon (Citron & Chesney, 2019).

3.2 Categories of Harm

NCDP constitutes a distinct and particularly severe category of AI-enabled harm. Unlike traditional intimate image offences, which presumes the existence of an authentic underlying recording, NCDP is created at the outset: explicit content can be generated without any

identifiable authentic recording, thereby placing evidentiary burden on the current statutes concerning sexual abuse, harassment and data security (Suki, 2024). Along with the difficult experience of sexual abuse faced by the victim due to NCDP, they face a secondary violation arising from disbelief or minimisation of their harm (Citron & Chesney, 2019). Accordingly, with the increasing volume of such content disseminated across social media platforms, victims may suffer from severe anxiety, depression, and loss of personal autonomy among affected individuals, the majority of whom are women and young people already subject to gender-based stigma (Park et al., 2023). Following this, law reform must therefore explicitly criminalise the creation and distribution of NCDP regardless of whether there is an existing recording, or purely synthetic.

4. Criminalisation without Hinder Innovation

4.1 ‘Chilling Effect’ Argument

Southeast Asian nations are actively competing to position themselves as regional AI hubs. Singapore’s and Indonesia’s National AI Strategy reflect significant state investment in AI as a pillar of economic competitiveness (Chua et al., 2025). Ensuring smooth economic transactions will include the role of compliance with national legislation. Criminal legislation that are too expansive run the risk of disincentivising research and development efforts (Panattoni, 2024). Further, in comparison to jurisdictions with a more flexible approach to innovation, overregulation can make a jurisdiction less competitive thus, preventing small and medium companies from development (Abdelaziz, 2025). The appropriate response is a tiered liability framework that reserves criminal sanctions for intentional misconduct and channels cases which does not meet the criminalisation threshold into civil litigation, seeking for statutory tort redress (Sai & Wang, 2026). Additionally, platform providers that make GenAI systems publicly available without intending their misuse should not face criminal liability when intent or reasonable foreseeability of a criminal act, known as *mens rea*, is absent; explicit legislative safe harbour provisions for good-faith developers are therefore essential (Suki, 2024; Chua et al., 2025).

4.2 Intent-based Criminalisation

The foundation of sound criminalisation must contain the aspect of anthropocentricity (Anggara, 2026). Criminal liability must remain anchored to the culpable human actor possessing *mens rea*. AI systems lack moral agency, the capacity for genuine intention, and the

ability to internalise criminal sanctions; they are instruments deployed by human actors (Anggara, 2026; Sai & Wang, 2026). Accordingly, the operator who knowingly uses AI systems to create and distribute NCDP bears criminal liability, as does the developer who designs a system for malicious use. By contrast, a developer of an AI system would not bear criminal liability for a third party's misappropriation of that system due to the absence of knowledge, inducement, or recklessness. Singapore and Indonesia's law instruments – Code of Practice for Online Safety (2023) and Government Regulation No. 17 of 2025 - made provisions to oblige social media platforms to implement reasonable safeguards to filter and remove harmful content, and shift the focus to the perpetrator's malicious intent.

4.3: Safe Harbours

To operationalise the balance between innovation and accountability, legislative frameworks should establish explicit safe harbour provisions for developers adhering to responsible practices (Panattoni, 2024). The EU AI Act mandates AI systems which functions out of good faith shall not be imposed with fines. Moreover, the regulatory sandbox approach exemplifies this approach, permitting organisations to test AI systems under strict compliance requirements in exchange for rigorous documentation and commitment to implementing safeguards prior to deployment at scale. Safe harbour provisions should protect developers who implement security-by-design practices consistent with established guidelines, such as those issued by Cyber Security Agency of Singapore, and who respond promptly to reports of misuse and cooperate with law enforcement. These provisions must define qualifying practices with precision, create mechanisms for withdrawal of protection in cases of persistent non-compliance.

5. Protecting Fundamental Human Rights

5.1: Anthropocentricity as a Pillar

Effective governance of AI-enabled harms must be grounded in a foundational commitment to human-centricity; AI systems must be designed and deployed in ways that preserves fundamental moral values such as but not limited to human autonomy and dignity (ASEAN, 2024). Importantly, AI must be prevented operating as an unsupervised social actor as it is able to cause disruptions to employment and reduced social benefits (ASEAN, 2024). In concrete terms, this requires meaningful human control over AI systems, defined as the capacity of human actors to understand system operation, intervene when harm occurs, and exercise

oversight over consequential automated decisions (Anggara, 2026). AI governance which has sufficient and consistent human oversight could bring about accountability with less friction by ensuring AI systems respond to human reasons and values. Further, authorised persons of the organisations shall have adequate knowledge and moral awareness to understand the system's limitations (Santoni de Sio & Mecacci, 2021). Human-centricity further demands attention to the distribution of harms. A system generating economic value while systematically harming a group of victims is not justifiable, and governance frameworks must prioritise protections of those who are prone to harm (ASEAN, 2024).

5.2: Protecting Vulnerable Groups

AI-enabled harms are not uniformly distributed amongst society. Women constitute the overwhelming majority of NCDP victims, and children are increasingly targeted through child sexual abuse materials which are generated with AI (De Bolle, 2026). Indonesia's Government Regulation No. 17 of 2025, also known as PP Tunas, reflects one protective model imposing mandatory content moderation obligations on Electronic System Providers (ESPs) to detect and remove child sexual exploitation material, and requiring cooperation with child protection authorities. More broadly, AI systems trained on unrepresentative data systematically disadvantage minority communities, generating offensive or inaccurate content about groups whose experiences are unfairly reflected in training datasets (Seghid et al., 2026). Protective frameworks should mandate diverse and representative data collection and prevent autocratic control where it is likely to impact particular vulnerable groups.

By way of example, Indonesia's Pentahelix Consortium, consisting of government agencies, academic institutions, technology companies, civil society organisations, and media in joint AI governance, offers a replicable collaborative model that reduces the risk of governance frameworks reflecting only technical or commercial interests to the exclusion of affected populations (Puspitasari & Pradana, 2025). This approach eliminates threats to vulnerable populations and enhance transparency and data resilience and security in a sustainable way.

6. Comparative Analysis

The European Union's AI Act represents a paradigmatic shift towards comprehensive, hard-law regulation of artificial intelligence systems. This prescriptive framework establishes categorical risk classification such as high-risk and minimal-risk, along with specific technical requirements, documentation standards, and mandatory compliance mechanisms. This leaves

limited discretion in implementation as this approach puts forth legal certainty and uniformity across all member states. The model assumes that harmonised rules can effectively govern heterogeneous landscape of AI application while upholding consumer protection and fundamental rights.

An example of the AI Act's characterisation is that it mandates that all high-risk AI systems are subjected to conformity assessment before they are released for use by the public. On the other hand, the ASEAN governance framework such as 'ASEAN Guide on AI Governance and Ethics – Generative AI' was designed as a voluntary framework, build to guide member states in building and deploying AI responsibly. Moreover, the ASEAN Responsible AI Roadmap categorises member states by their level of readiness in order for member states of varying levels to focus on different aspects of a responsible AI governance, in respect to their level of readiness.

As shown previously, ASEAN's governance framework adopts a deliberately flexible approach. Rather than defining AI systems through categorical risk levels with corresponding regulatory obligations, ASEAN focuses on outcome-based objectives, while allowing member states considerable authority in how to achieve them. This methodological distinction reflects a recognition that prescriptive, although provides clarity, may inadvertently hinder innovation and overlook contextual variations across diverse ecosystems. In result, the approach which ASEAN takes acknowledges that AI systems are continually evolving. Hence, ASEAN prioritises adaptability and proportionality, allowing regulatory measures to scale with demonstrated harms however it lacks explicit guidelines on the attribution of harm initiated especially by human actors.

6.1 Harmonisation and Unification: Pluralism in ASEAN AI governance

The deliberate distinction between regulatory harmonisation and unification consists of a foundational principle of ASEAN's AI governance architecture. Harmonisation seeks to align substantive objectives, technical standards, and safety protocols across member states while preserving discretion in implementation mechanisms, institutional design, and enforcement strategies (Erasmus, 2025). In comparison, unification demands the imposition of standardised regulations across jurisdictions, therefore constraining the ability of individual states to tailor approaches to local regulatory landscape, institutional capacities, and developmental priorities (Ponomareva, 2022).

This commitment to harmonisation over unification reflects ASEAN's enduring commitment to the principle of national sovereignty within a multilateral framework. Member states are given the authority to create and enforce laws in which is befitting to their respective technological infrastructure and regulatory capacity. Furthermore, regulatory pluralism within a harmonised framework permits members states to serve as 'natural laboratories' for policy experimentation, allowing successful innovation to diffuse organically across borders while permitting rapid course correction when approaches prove ineffective (ASEAN, 2024). Therefore, enforcing member states to attribute criminal liability on the creation of NCDP materials and its likes will benefit from such an approach moving forward.

6.2: Anthropocentricity in Criminal Law

In the current landscape of criminal law, albeit it is evolving with the trends, the laws enacted is centered behind the motivation of the human perpetrator. The legal responsibility presupposes human agency and intentionality (Panattoni, 2025). The foundations of criminal liability requires two essential elements: a human act (*actus reus*) along with the culpable intent (*mens rea*) or criminal negligence. This foundation assumes that the perpetrator possesses the cognitive capacity for conscious decision-making, moral deliberation, and accountability for their actions. With the emergence of nascent technological innovations like artificial intelligence, it presents a significant challenge to the traditional foundation as these technologies work through mechanisms different from human cognition.

6.3: Doctrinal Conflict

The divergence between human agency and artificial intelligence systems is vulnerable to operational contradiction within existing criminal laws. Artificial intelligence systems function through probabilistic calculations and algorithmic computation whereas humans have the cognitive capacity to consciously make decisions and reason with morals (Yeung, 2018). It lacks the subjective experience, intentionality, and moral awareness where it is not able to attribute blame, or compute wrongdoings similar to humans. When a system produces harmful outcomes, through discrimination or misclassification, the absence of a 'guilty mind' renders it legally incompatible with conventional concepts of culpability. Hence, the incapability is a conflict between the nature of artificial intelligence and the philosophical core of criminal responsibility.

6.4: Culpability Gap

Entities which lack consciousness and moral agency such as artificial intelligence create the inability to assign moral blame. Incrimination of an unlawful act becomes challenging when there is an absence of consciousness of a natural person. A phenomenon called ‘black box’ problem where the developers are unable to explain the system’s decision-making processes in creating an output causes difficulty in establishing the causal relationship between human conduct and harmful AI-generated content. Taking this into account, this would give perpetrators the opportunity to exploit the traditional legal requirements of intention and control therefore, making it difficult to hold perpetrators accountable. For instance, section 377BE(5)(b) of Singapore’s Penal Code excluded media as evidence of deepfake pornography when “so altered that no reasonable person would believe that it depicts” the victim. Further, perpetrators could lack culpability for non-consensual generated materials when they claim to had no reason to believe their actions would cause “humiliation, alarm, or distress”.

In the regions ASEAN and EU, both had discussed the overall theme of accountability as an aspect of risk management. At this moment of writing, ASEAN prescribes accountability as one of the guiding principles for member states to comply with, by including human oversight and sufficient monitoring mechanism (ASEAN, 2024). The EU AI Act, in contrast, necessitates AI systems, especially those who are classified as high-risk, to legally oblige with technical documentation, logging, human oversight. It is apparent that culpability gap represents not merely a technical challenge for law enforcement but a fundamental philosophical problem regarding how to attribute moral responsibility in a system where decision-making processes exceed human comprehension and conscious intent (Santoni de Sio & Mecacci, 2021).

6.5: Preserving Human Agency

Due to the apparent challenge in attributing blame to AI systems for its ability to generate harmful content, academics and lawmakers iterates for the focus on the human factor in the criminal act. Since AI systems does not possess moral reasoning like humans do, lawmakers and academics mentions to keep the criminal liability to the human agents responsible for designing, deploying, and overseeing AI systems (Anggara, 2026; AI Act, 2024). The human-centered liability model clears doubt as to whom lies the liability on such harmful content. If the deployers used the AI system not for its intended purpose, it is highly likely that the deployer will receive the highest culpability in the data chain. The problem before the

emergence and recognition of the model was placing accountability on to whom in the data chain (Bovens, 1998). People who faced problems with incorrect data had a hard time finding who is the right person of contact in companies handling their personal data. During the early stages of AI adoption, it became more severe. By following the model, it preserves the essential connection between criminal responsibility and human agency by recognising that harms attributed to AI systems ultimately result from human decisions regarding development methodologies, deployment strategies, oversight mechanisms, and the contexts in which these AI systems were used for (Anggara, 2026).

As it stands, ASEAN AI governance frameworks encourages for member states to establish legal frameworks inculcating the guiding principles. The ASEAN Guide on AI Governance and Ethics addressed non-consensual generated materials as one emerging risk whilst promoting the need for a robust cyber security framework. Singapore, in its Online Safety (Relief and Accountability) Act 2025 target the perpetrator's intent to mimic realism or cause humiliation, even if the image is purely synthetic. This alignment between ASEAN's governance principles and Singapore's criminal provisions demonstrates that regardless of regional approach, whether aspirational or prescriptive, criminal law must ultimately focus on the human decisions that weaponise AI systems, not the technology itself.

6.6: Criminalisation in the Region

Despite the philosophical limitations on attribution moral agency to AI systems, multiple jurisdictions have implemented criminal prohibitions against specific AI-enabled harms. This legislative trend reflects recognition that regardless of the absence of moral agency, the intentional or reckless use of AI systems by human actors to cause harm requires criminal sanction. These jurisdictional examples demonstrate convergence on the principle that criminal law should focus on the malicious intent of human perpetrators utilising AI tools rather than on the AI systems.

Singapore has emerged as a jurisdictional leader in implementing Ai-specific criminal provisions. The Elections (Integrity of Online Advertising) (Amendment) Bill 2024 explicitly criminalises the publication of realistic deepfakes depicting election candidates during polling periods, recognising that such content poses a direct threat to democratic integrity and electoral legitimacy. Furthermore, the proposed Criminal Law (Miscellaneous Amendments) Bill 2025 extends criminal liability to the production of "completely synthetic" non-consensual intimate images, representing a significant expansion of criminal protection even for victims of AI-

generated abuse where no actual recorded material exists (Withers, 2026). This development acknowledges that the harm principle upholding criminal law applies equally to synthetic representations, as victims suffer psychological trauma and reputational damage regardless of whether the media originated from a real recording.

Similarly, Indonesia has adopted a hybrid approach to addressing criminal liability for AI-enabled harms through the implementation of PP Tunas and Circular Letter No. 9 of 2023, which together shift the regulatory paradigm from purely reactive criminal enforcement toward preventive platform accountability and ethical governance standards. PP Tunas established binding legal duties to Electronic System Providers to actively identify, filter, and remove content that are deemed in contravention of human rights, including sexual violence and exploitation which could be facilitated by AI. However, Indonesian lawmakers have identified a “sanction gap” whereby the current penalties remain primarily administrative rather than substantive criminal or financial sanctions (PS Engage, 2026). Complementing the PP Tunas, the soft law instrument Circular Letter No. 9 of 2023 established the ethical foundations for assessing human culpability such as safety, accountability, and inclusivity. This instrument provides the normative benchmarks necessary for courts to determine negligence or malicious intent when AI systems produce harmful outputs, particularly by directing developers and users to prevent acts which would cause bias or abusive generated content.

These two ASEAN member states exemplify dual regulatory models for addressing AI-enabled harms within the region’s shared governance framework: Singapore, as the first to codify criminal prohibitions, establishes direct criminal liability for malicious deployment of AI uses in various contexts, while Indonesia contributes to a preventive governance model mandating platform accountability and meaningful human control throughout the AI lifecycle, though the effectiveness of this complementary approach remains constrained by the current absence of enforceable criminal sanctions proportional with the severity of AI-enabled harms.

6.7: Shared Responsibility Framework

From initial conception and design through development, deployment and operation, the diffusion of responsibility across multiple actors throughout the AI lifecycle creates the structural problem. A problem occurs when one actor of the value chain does not have complete authority over, or visibility, into the entire AI process in producing harmful outputs. The problem of diffusing responsibility is centra to AI governance frameworks, with both ASEAN and the EU working to develop an increasingly sophisticated models for demarcating liability

among developers, service providers, deployers, and end-users. With the emergence of a shared responsibility framework, a significant shift from a binary liability assignment toward multi-party accountability structures that recognise the interdependence of actors within the AI value chain.

The EU AI Act established a legally binding demarcation of responsibilities among various actors. The legislation defines specific categories of actors and assigns distinct legal obligations based on their position within the value chain. For example, the providers bear primary responsibility for ensuring that systems comply with applicable requirements prior to market entry whilst, the deployers assume responsibility for appropriate use, ongoing monitoring, and intervention when systems produce suboptimal outcomes. By this approach, it aims to strengthen the capacity to trace accountability to specific actors and minimise lexical ambiguity. Most importantly, the framework preserves the conventional anchoring of criminal liability in human and corporate actors, safeguarding the ethical foundations of punishment while remaining responsive to the regulatory complexities introduced by emerging technologies. For ASEAN, it has yet to inculcate the similar model into their AI governance by organising forums for various actors of the AI value chain to discuss risk management measures. The Guide on AI Governance and Ethics aimed to set normative foundation on a voluntary framework for all actors whilst aligning with international standards such as the Hiroshima Process International Code of Conduct for Organisations Developing Advanced AI Systems. However, it is upon ASEAN's discretion to strongly encourage enforceability of such a framework, as implementation rigor could cause significant variation across member states and organisations.

Policy Recommendation

8.1 Implementation of intent-based criminalisation and harmonisation of measures

Criminal liability should be confined to natural or legal persons acting with evident malicious intent, while unintentional technical failures and emergent AI behaviours are addressed through civil or statutory tort mechanisms (Sai and Wong, 2026). Singapore has shown that the anthropocentric approach, whereby its definition of synthetic non-consensual materials was included, and made illegal of non-consensual generated materials depicting a child, regardless if it is an imitation or based on a real person. ASEAN should adopt this framework as a regional template with explicit safe harbour provisions for good-faith developers.

Definitional harmonisation is equally important. ASEAN should adopt an uniform definition of ‘inauthentic material’ as any digitally generated or altered content constituting a false or misleading material to deceive a reasonable person. This aligns with ASEAN vision of reducing regulatory fragmentation as a precondition for sustainable digital trade growth.

8.2 Statutory Torts as a Middle Ground for Redress

Criminal proceedings are unsuitable as the primary redress mechanism for AI-facilitated harms given their high evidentiary burden and protracted timelines. The Online Safety (Relief and Accountability) Act 2025 established a dual-track model: an Online Safety Commission empowered to issue binding content-removal directions, complemented by statutory torts granting victims the right to pursue civil proceedings for compensatory damages and injunctions directly against perpetrators and platform intermediaries. Member states should adopt equivalent authority which deems fit in their legal landscape, covering deepfake and non-consensual generated materials abuse. Accordingly, this would distribute accountability across communicators, administrators, and platforms to create structural incentives for preventive action.

8.3 ‘Lead-and-Leverage’ Model

Due to uneven regulatory landscape across ASEAN, a uniform rules-based framework similar to the EU AI Act is not feasible. Under a ‘Lead-and-Leverage’ approach, advanced member states such as Singapore and Indonesia should pilot hard-law instruments, while emerging economies focus on foundational data privacy and cybersecurity measures. The ASEAN AI Safety Network provides the institutional architecture for this model, enabling resource-constrained states to access shared safety benchmarks and regional regulatory templates without the cost of independent research (ASEAN, 2024).

8.4 Formation of a Taskforce

On frontier risks, ASEAN member states have been structurally excluded from high-level international AI safety forums, leaving the region as a norm-recipient rather than a norm-setter. ASEAN should establish a taskforce within the ASEAN AI SAFE Network, mandated to monitor agentic AI and generative AI trajectories and issue regional risk assessments that national governments can localise. This would ensure ASEAN’s participation in global forums with a coherent regional position, advancing the transition from passive policy adoption to active standard-setting.

Conclusion

ASEAN member states confront a critical challenge at the intersection of AI-enabled sexual violence and digital economic opportunity. The weaponisation of generative AI for non-consensual generated materials and other harms has created an urgent need for criminal accountability. The analysis presented above demonstrated that this tension is not ungovernable. ASEAN's guidelines on ethics shows motivation towards a responsible future for AI. Legislation from Singapore with Indonesia's evolving legal framework provide concrete institutional and doctrinal precedents that establish the feasibility of the balance between protection and innovation. Furthermore, the policy recommendations operationalises this balance through the four mechanisms discusses above: intent-based criminalisation and harmonisation of measures; statutory tort for redress; lead-and-leverage model; and the formation of a taskforce.

A framework which directly addresses the question posed in this paper will demonstrate effective criminalisation of AI-enabled harms and the preservation of conditions favourable to innovation are complementary, rather than contradictory objectives. By anchoring liability to culpable human conduct, providing explicit safe harbours, establishing civil remedies, implementing tiered governance structures, and ensuring that the region maintains a voice in international AI governance, ASEAN and its member states can protect the fundamental rights and dignity of individuals harmed by generative AI. The window for proactive, principled governance is narrow. Jurisdictions that defer such frameworks risk being forced into reactive, punitive measures that are less calibrated and ultimately less conducive to innovation and human flourishing.

Bibliography

1. Abdelaziz, DKA. (2025). Criminal liability for the misuse and crimes committed by AI: A comparative analysis of legislation and international conventions. *Journal of Infrastructure, Policy and Development*. 9(1), 10722. <https://doi.org/10.24294/jipd10722>.
2. Anggara, I. (2026). Legal Liability of Artificial Intelligence in Criminal Law: A Comparative Study between Indonesia and European Union. *Indonesian Journal of Law and Justice*, 3(3), 12. <https://doi.org/10.47134/ijlj.v3i3.5565>.
3. ASEAN. (2024). *ASEAN Responsible AI Roadmap (2025-2030)*. ASEAN.
4. ASEAN. (2024). *Expanded ASEAN Guide on AI Governance and Ethics – Generative AI*. ASEAN Secretariat.
5. Bovens, M. (1998). *The Quest for Responsibility: Accountability and Citizenship in Complex Organisations*. Cambridge University Press.
6. Chadha, S. (2023, December 12). How Southeast Asia can become a \$1 trillion digital economy. World Economic Forum. <https://www.weforum.org/stories/2023/12/how-southeast-asia-can-become-trillion-digital-economy/>.
7. Chesney, R., & Citron, D. (2019). Deepfakes and the New Disinformation War: The Coming Age of Post-Truth Geopolitics. *Foreign Affairs*, 98(1), 147–155. <https://www.jstor.org/stable/26798018>.
8. Chua, L., Tham, P., Okolo, C. T. (2025) *AI SAFETY GOVERNANCE, THE SOUTHEAST ASIAN WAY*. AI Safety Asia and Brookings Institution. <https://www.brookings.edu/articles/ai-safety-governance-the-southeast-asian-way/>.
9. Cyber Security Agency of Singapore. (2024). *GUIDELINES ON SECURING AI SYSTEMS*. Cyber Security Agency of Singapore.
10. De Bolle, C., (2026, March 26). A legal vacuum on CSAM detection puts children at greater risk. Europol. <https://www.europol.europa.eu/media-press/newsroom/news/combating-child-sexual-exploitation-statement-catherine-de-bolle>.
11. Erasmus, E. G. (2025). Harmonization, Convergence and Divergence of Accounting Standards: A Literature Synthesis. *Journal of Accounting and Financial Management*, 11(5), 43-62. <http://doi.org/10.56201/jafm.vol.11.no5.2025.pg43.62>.

12. European Parliament: Directorate-General for Internal Policies of the Union, Park, K., Ging, D., Murphy, S., & McGrath, C. (2023). *The impact of the use of social media on women and girls*, European Parliament. <https://data.europa.eu/doi/10.2861/521265>
13. Government Regulation No. 17 of 2025.
14. Info-communications Media Development Authority. (2023). *CODE OF PRACTICE FOR ONLINE SAFETY*.
15. Koetzier, L. R., Wu, J., Mastrodicasa, D., Lutz, A., Chung, M., Koszek, W. A., Pratap, J., Chaudhari, A. S., Rajpurkar, P., Lungren, M. P., & Willemink, M. J. (2024). Generating Synthetic Data for Medical Imaging. *Radiology*, 312(3), e232471. <https://doi.org/10.1148/radiol.232471>
16. KPMG. (2023). *AI & generative AI in functional transformation*. KPMG. <https://assets.kpmg.com/content/dam/kpmgsites/uk/pdf/2024/06/ai-and-generative-ai-in-functional-transformation.pdf>.
17. O'Regan, E. (2026, February 17). Ireland launches 'large-scale inquiry' into Musk's AI bot Grok. *Politico*. <https://www.politico.eu/article/ireland-inquiry-elon-musk-grok-privacy-artificial-intelligence-ai/>.
18. Panattoni, B. (2025). Generative AI and criminal law. *Cambridge Forum on AI: Law and Governance*, 1(9), 1-14. <https://doi.org/10.1017/cfl.2024.9>.
19. Ponomareva, D., S. (2022). Harmonisation and unification and their influence on the evolution of national legal systems. *Journal of the Belarusian state University*. 2. 39–43. <https://doi.org/10.33581/2521-6848-2022-2-39-43>.
20. PS Engage. (2026, January 13). *PP Tunas: Indonesia's New Digital Regulation to Protect Children Online*. PS Engage. <https://ps-engage.com/pp-tunas-indonesias-new-digital-regulation-to-protect-children-online/>.
21. Puspitasari, C., Pradana, A., H., (2025). NAVIGATING THE DIGITAL THREAT: A PENTAHELIX APPROACH TO AI-DRIVEN DATA SECURITY IN SMART CITIES. *Jurnal Informatika Polinema*, 12(1), 29-38.
22. Sai, S., & Wang, Z. (2026). Criminal Regulatory Approaches to Deepfake-Related Offenses: Focusing on the Crime of Fraud. *International Journal of Asian Social Science Research*, 3(1), 20-31. <https://doi.org/10.70267/ijassr.260301.2031>.
23. Santoni de Sio, F., Mecacci, G. (2021). Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them. *Philos. Technol.* **34**, 1057–1084. <https://doi.org/10.1007/s13347-021-00450-x>.

24. Seghid, N., Iqbal, F., Al-Room, K., and MacDermott, Á. (2026). Emerging threats in AI: a detailed review of misuses and risks across modern AI technologies. *Front. Commun. Netw.* 6:1727425. [http://doi.org/ 10.3389/frcmn.2025.1727425](http://doi.org/10.3389/frcmn.2025.1727425).
25. Suki, A.C. (2024, November 20). Andrea Christine Suki: Law Reform Paper on Criminal Liability and Generative Artificial Intelligence. *Lawtech.asia*. <https://lawtech.asia/andrea-christine-suki-law-reform-paper-on-criminal-liability-and-generative-artificial-intelligence/>
26. Tsiavos V., Kitsios, F. (2025) The digital transformation of the film industry: How Artificial Intelligence is changing the seventh art. *Telecommunications Policy*, 49(8). <https://doi.org/10.1016/j.telpol.2025.103021>.
27. Watkins, A., (2026, January 11). Malaysia and Indonesia Block Access to Grok Because of Sexually Explicit Content. *The New York Times*. <https://www.nytimes.com/2026/01/11/world/asia/malaysia-indonesia-grok-ban.html>.
28. Withers. (2026, January 26). Deepfakes in Singapore: legal recourse, regulatory powers, and building long-term corporate resilience. Withers Worldwide. <https://www.withersworldwide.com/en-gb/insight/read/deepfakes-in-singapore-legal-recourse-regulatory-powers-and-building-long-term-corporate-resilien>.
29. Yeung, K. (2018). A Study of the Implications of Advanced Digital Technologies (Including AI Systems) for the Concept of Responsibility Within a Human Rights Framework. *MSI-AUT*, 05. <https://ssrn.com/abstract=3286027>.