# Artificial Intelligence Regulation in Indonesia: A Normative Study of Deepfake Misuse Within the Risk-Based Regulation Framework

**Fanissa Dara Safina[1], Agustianto[2]**

[1,2]Faculty of Law, Universitas Internasional Batam, Indonesia
✉ Corresponding email: 2351084.fanissa@uib.edu

**History of Article**
Submitted : June 30, 2026
Revised : July 22, 2026
Accepted : August 10, 2026
Published : September 22, 2026
DOI : https://doi.org/10.37253/conescintech.v6i1.12806

**Copyright© 2026 by Author(s).** This work is licensed under a Creative Commons Attribution-Non Commercial-Share Alike 4.0 International License. All writings published in this journal are personal views of the authors and do not represent the views of this journal and the author's affiliated institutions.

## Abstract
The rapid development of Artificial Intelligence (AI) has accelerated the emergence of deepfake technology, which offers significant benefits in fields such as entertainment, education, and digital content creation but also poses serious legal and societal risks. The misuse of deepfakes has contributed to privacy violations, identity theft, defamation, misinformation, fraud, and other forms of cybercrime. In Indonesia, the current legal framework remains inadequate because existing regulations are fragmented, reactive, and lack specific provisions governing deepfake technology. This study aims to analyze the adequacy of Indonesian legal regulations in addressing deepfake misuse and to propose the adoption of a risk-based regulation approach as a more adaptive and comprehensive regulatory framework. This research employs a normative legal method using statutory and conceptual approaches by examining relevant legislation, including the Electronic Information and Transactions Law, the Personal Data Protection Law, and the Indonesian Criminal Code. The findings indicate that Indonesia's legal framework primarily relies on broad interpretations of existing laws, resulting in legal uncertainty and limited preventive capacity. In contrast, a risk-based regulation approach classifies AI applications according to their level of risk and imposes proportional legal obligations, including transparency, accountability, and preventive risk management. The study concludes that adopting a risk-based regulatory framework would strengthen Indonesia's legal system by improving legal certainty, protecting individual rights, and ensuring a more effective balance between technological innovation and public safety.

**Keywords:** artifical Intelligence; deepfake; risk-based regulation; legal regulation; risk clasification

# Introduction

The development of Artificial Intelligence (AI) technology has brought significant transformation in various aspects of human life, including in the production and manipulation of digital content (Lestari et al., 2026). One of the most prominent forms of AI development is deepfake technology, namely technology that can create or modify audiovisual content so that it appears realistic and resembles real conditions (N. P. M. Putri et al., 2024). Deepfake enables the manipulation of a person's face, voice, and gestures in video or audio in a highly convincing manner, making it difficult to distinguish by human observers (Shabrina et al., 2026). Initially, deepfake technology was developed for creative and entertainment purposes, however, in its development, this technology has raised various serious legal and social risks (Widjaja, 2025). The misuse of deepfake can be used to spread false information, conduct digital fraud, defamation, extortion, and online sexual exploitation (Prayoga & Tuasikal, 2025). The main characteristic of deepfake, which is capable of producing hyper-realistic content, makes it a new instrument in cybercrime that not only harms individuals but also has the potential to disrupt public information stability.

In Indonesia, the phenomenon of deepfake misuse has begun to cause real impacts. One prominent case is the circulation of a deepfake video featuring a public official, namely the Governor of East Java, which appeared to promote a certain product, thereby causing harm to the public (Redaksi Tribata, 2025). This case shows that deepfake is no longer merely a technological issue but has become a real and rapidly evolving legal problem, with increasingly complex levels of risk along with the advancement of generative AI technology. However, legal regulation in Indonesia currently has not specifically accommodated the deepfake phenomenon (Basah et al., 2025). Law enforcement in Indonesia still relies on general legal instruments such as Law Number 1 of 2024 concerning the Second Amendment to Law Number 11 of 2008 on Electronic Information and Transactions (EIT Law), Law Number 27 of 2022 concerning Personal Data Protection (PDP Law), and the Indonesian Criminal Code (KUHP). Although several provisions in these regulations can be used to prosecute perpetrators of deepfake misuse, such as provisions related to defamation, dissemination of false information, and unlawful processing of personal data, all of these instruments are general in nature and are not designed to regulate the specific characteristics of AI technology (Mecca et al., 2025).

This condition raises several fundamental issues. The absence of an explicit legal definition of deepfake within the Indonesian legal system causes uncertainty in the legal classification of various forms of misuse of this technology (Nurkholisah et al., 2025b). The existing legal approach remains reactive and partial, meanwhile, the misuse of Artificial Intelligence requires a preventive, risk-based, and systemic protection approach (Azis, 2025). In a global context, the development of artificial intelligence regulation shows a paradigm shift toward risk-based regulation, namely a regulatory approach that classifies AI systems based on the level of risk they pose and imposes proportional legal obligations (Wardoyo, 2026). One of the most significant implementations of this approach is the European Union Artificial Intelligence Act (EU AI Act), which classifies

AI systems into risk categories, namely minimal risk, limited risk, and high risk (Puannandini et al., 2025). Each category is subject to different legal obligations, including transparency requirements, AI content labeling, system audits, and prohibitions on certain uses considered harmful (Nugroho & Halim, 2026).

This risk-based approach demonstrates an important shift in modern regulatory theory, namely from an approach that focuses only on enforcement after violations occur, toward a preventive approach that regulates potential risks from the design and implementation stages of technology. In the context of deepfake, this approach is relevant because of the technology's characteristics, which evolve rapidly, are difficult to detect, and have the potential for wide-ranging impacts on individual rights, information integrity, and public security (S. M. I. Putri et al., 2024). Compared to these practices, the Indonesian legal framework still shows limitations in responding to the complexity of AI technology, particularly deepfake. Dependence on general regulations such as the EIT Law, PDP Law, and KUHP indicates that existing regulations have not integrated a risk-based approach as a regulatory foundation. As a result, law functions more as a repressive instrument rather than a preventive one, and thus has not been able to anticipate potential harm before it occurs.

Based on these conditions, there is an urgency to develop a more adaptive and comprehensive regulatory approach in addressing deepfake misuse in Indonesia. The risk-based regulation approach is considered a relevant framework because it enables the classification of deepfake technology usage risks, ranging from low risk to high risk, and allows the implementation of different legal obligations according to the level of harm. For example, high-risk deepfakes such as political disinformation or non-consensual pornography may be subject to strict prohibitions, while low-risk uses may be permitted with the condition of transparency through AI-based content labeling. In addition, the risk-based approach also emphasizes the importance of transparency, accountability, and prevention through risk assessment mechanisms before technology is widely used (Nugroho & Halim, 2026). In this context, digital platform providers and AI system developers bear responsibility for identifying and mitigating potential risks from the early stages of development. This is in line with the precautionary principle in modern law, which emphasizes preventing harm before broader impacts occur in society (Syahra et al., 2025).

Several previous studies have examined AI regulatory issues in Indonesia from various perspectives. Utami, Sutrisni, & Suharyanti (2025) in "Legal Regulation of Responsibility for the Utilization of Artificial Intelligence in Indonesia" compared the legal vacuum of AI in Indonesia with the EU AI Act, concluding that although Indonesia has the EIT Law and PDP Law, both do not explicitly regulate prohibited practices, risk classification, compliance requirements, and strict administrative sanction systems. Hasanuddin, Nurfransiska, & Maryuliyanto (2026) in "Artificial Intelligence Regulation in the Perspective of Indonesian Law: Normative Analysis of Challenges, Implications, and Ideal Regulatory Models" found that AI regulation in Indonesia is sectoral and fragmented, spread across various laws not specifically designed to address the complexity of AI technology. Nicholle (2025) in "Legal Liability for AI-Based

Automated Decisions: A Comparative Study between the European Union and Southeast Asia" highlights the relevance of the EU AI Act as a risk-based framework, taking a Southeast Asian perspective overall.

Theoretically, this research contributes by developing a normative analytical framework that positions the risk-based regulation approach as an instrument to evaluate the adequacy of legal regulation regarding deepfake misuse within the Indonesian legal system. This approach allows the classification of deepfake technology risk levels and the determination of more proportional and measurable legal obligations, which have not been explicitly accommodated in regimes such as the EIT Law or KUHP. Practically, the results of this study are intended as conceptual consideration material for policymakers in formulating more comprehensive, adaptive, and anticipatory AI regulations in response to the development of digital manipulation technology. The relevance of this research is further strengthened by the increasing cases of deepfake misuse in Indonesia, which demonstrate the limitations of existing legal instruments in responding to the technical characteristics and social impacts of this technology. This research uses a normative juridical approach focusing on the analysis of positive legal norms, legal principles, and comparisons with regulatory frameworks in other jurisdictions.

This research is not intended to produce empirical findings regarding legal implementation in practice. The main focus of this research lies in the construction of normative argumentation and the formulation of an ideal regulatory model in the context of deepfake misuse in Indonesia. Thus, strengthening the legal framework based on risk-based regulation is not only intended to fill the legal vacuum in regulating deepfakes in Indonesia, but also to build a more adaptive regulatory system toward the development of artificial intelligence technology. This approach is expected to create a balance between the protection of individual rights, legal certainty, and technological innovation, so that law does not lag behind the rapidly evolving development of technology.

## Research Method

This research uses a normative legal research method with a focus on examining legal norms, principles, and concepts governing artificial intelligence. This method is chosen because the research aims to provide recommendations for strengthening artificial intelligence regulation in Indonesia based on a risk-based regulation approach from a human rights protection perspective (Disemadi, 2022). This study employs a statute approach that examines AI regulation in Indonesia, including Law Number 1 of 2024 concerning the Second Amendment to Law Number 11 of 2008 on Electronic Information and Transactions and Law Number 27 of 2022 concerning Personal Data Protection, in order to identify the extent to which these norms are able to address the risks posed by AI system and a conceptual approach to examine the legal concept of artificial intelligence through the lens of risk-based regulation.

The data sources in this research include Law Number 1 of 2024 concerning the Second Amendment to Law Number 11 of 2008 on Electronic Information and Transactions, Law Number 27 of 2022 concerning Personal Data Protection, Law Number 1 of 1946 concerning Criminal Law Regulations, along with its supporting

documents. Secondary legal materials include books, scientific journals, research reports, and legal literature relevant to AI and technology regulation. Data collection techniques are conducted through library research, namely by systematically searching, reviewing, and analyzing all relevant legal sources (Tan, 2021). The collected data are analyzed using descriptive qualitative analysis, in which legal materials are examined to understand the substance of norms, identify regulatory gaps, and evaluate the extent to which Indonesia's legal framework regulates deepfake technology.

# Results and Discussions

## Legal Regulation of Deepfake in Indonesia

The development of Artificial Intelligence (AI) technology has brought significant transformation in the way humans produce, distribute, and consume digital information (Devi, 2025). One of the most prominent manifestations of this development is deepfake technology, which employs manipulation techniques capable of reconstructing a person's face, voice, and expressions into audiovisual content that appears highly realistic (Anuar, 2026). In practice, deepfake has a dual-use nature, as it can be utilized for creative purposes such as in the film industry, education, and entertainment; however, it also has a high potential to be misused for harmful purposes, such as disinformation, fraud, defamation, and violations of privacy and identity (Usman, 2025). These characteristics position deepfake not merely as a technological innovation, but also as a complex legal issue involving ethical aspects, information security, and the protection of individual rights. This necessitates a normative response that is not only repressive in nature, but also capable of anticipating the risks arising within an ever-evolving digital ecosystem.
The legal substance in Indonesia has not yet accommodated the preventive aspects required in addressing deepfake-related crimes. Existing provisions tend to be repressive and curative in nature, focusing on punishment after the crime has occurred, while not providing effective preventive mechanisms (Alif et al., 2025). In the context of deepfake, preventive aspects are particularly important, considering that the damage caused by deepfake content can be irreversible, especially in terms of reputational harm and psychological trauma to victims. The absence of legal provisions regulating preventive obligations for digital platforms and technology developers creates a legal vacuum in efforts to prevent deepfake-related crimes (Kristiyenda et al., 2025).

Law Number 1 of 2024 concerning the Second Amendment to Law Number 11 of 2008 on Electronic Information and Transactions (EIT Law) is the primary legal instrument used in the enforcement of cybercrime law in Indonesia, including cases related to the misuse of deepfake technology. Normatively, the EIT Law does not explicitly recognize or define the term "deepfake" (Akbar et al., 2026). Therefore, any form of deepfake misuse is legally classified through analogical interpretation of the general provisions contained in the EIT Law. Although the EIT Law does not explicitly define deepfake, law enforcement authorities interpret its general provisions to prosecute offenders. The absence of such a definition also implies that legal practice is not based on specific regulation but rather depends on legal qualification through analogical

interpretation of existing general provisions. In other words, any form of deepfake misuse is not treated as a new offense but is instead forced into existing general legal constructions. This condition demonstrates that the EIT Law operates in an ex post facto manner in responding to technological developments, whereby the law is applied after the act has occurred and after the harm has arisen, rather than functioning as an instrument capable of anticipating risks from the outset.

Article 27 paragraph (1) of the EIT Law regulates the prohibition of distributing electronic information containing content that violates decency. This provision is particularly relevant to cases of deepfake pornography or manipulative content that displays a person's face in indecent situations without consent (Al-Ayoubi & Suharto, 2025). According to Constitutional Court Decision No. 50/PUU-VI/2008, content that violates decency may include images or information depicting sexual activities that contradict societal norms. Deepfake pornography that places a victim's face in sexual acts clearly constitutes obscene content and is therefore classified as violating decency (Domuari & Yudiantara, 2026). The elements of Article 27 paragraph (1) are fulfilled when the perpetrator intentionally creates and distributes deepfake pornography without the victim's consent and with the intent to harm the victim, as it involves "deliberately and unlawfully distributing decency-related content for public access" containing obscene material. In such cases, the distribution of deepfake pornography may be subject to criminal sanctions under Article 45 paragraph (1) in conjunction with Article 27 paragraph (1) of the EIT Law, carrying a maximum imprisonment of six years and/or a fine of up to IDR 1 billion.

Article 27A of the EIT Law prohibits the dissemination of electronic content containing insult and/or defamation. The criminal provision in Article 45 paragraph (1) of the EIT Law stipulates penalties of up to six years imprisonment and/or a fine of up to IDR 1,000,000,000 for violations of Article 27A. Deepfakes that manipulate a person's face or voice to depict statements or actions that harm their reputation may be classified as defamation (Yanuafriza & Eluan, 2025). A false deepfake video portraying someone as if they were attacking another person's honor or reputation through false accusations may be prosecuted under this article, as it fulfills the elements of making false imputations intentionally and without lawful justification. In addition, Article 28 paragraph (1) of the EIT Law regulates the dissemination of false and/or misleading information that results in material losses for consumers in electronic transactions. Deepfake technology can be used to create false narratives that appear visually and audibly authentic, thereby potentially causing harm to individuals or society (Fa'izah & Kurniasari, 2026). If a perpetrator distributes false video or audio content that deceives victims and results in material loss, this provision may be applied, carrying a criminal penalty of up to six years imprisonment.

Article 28 paragraph (2) of the EIT Law, which regulates the dissemination of information that incites hatred based on ethnicity, religion, race, and intergroup relations (SARA), may also be applied when deepfakes are used for social provocation. This provision prohibits electronic content that "incites, encourages, or influences" others to engage in hatred or hostility based on SARA factors (Apandi et al., 2025). Deepfakes that

manipulate a person's video for SARA-based provocation may be punished under Article 28 paragraph (2), carrying a maximum sentence of six years imprisonment. Article 28 paragraph (3) of the EIT Law regulates the dissemination of false information that may cause public disorder. Deepfakes that create large-scale false information capable of triggering public panic may be classified as hoaxes that threaten social stability if they result in unrest or societal disturbance (Fa'izah & Kurniasari, 2026).

Article 29 of the EIT Law prohibits the transmission of electronic information containing threats of violence or attempts to directly instill fear in victims. This provision is essentially intended to protect individuals from forms of digital intimidation that cause fear or psychological pressure. In the context of deepfake misuse, this article becomes relevant because the technology can be used to create content that appears to show threats from one person against another, thereby causing the victim to feel intimidated or placed under psychological pressure (Prayoga & Tuasikal, 2025). In practice, Article 29 of the EIT Law is often associated with cases of digital extortion or intimidation, for example when perpetrators of deepfake pornography threaten victims to comply with certain demands by threatening to disseminate or disclose the fabricated content without the victim's consent. For such acts, the perpetrator may be subject to a maximum imprisonment of four years.

Article 36 of the EIT Law is a provision that regulates legal liability for acts that intentionally and unlawfully cause harm to others, both material and immaterial losses. This article is relevant because the impacts caused are not limited to economic losses but also include non-economic aspects such as damage to a person's reputation, psychological distress, loss of public trust, and social stigma attached to individuals as a result of seemingly authentic manipulative content (Intani et al., 2026). These immaterial losses are increasingly significant considering the nature of deepfake technology, which is capable of realistically reproducing a person's identity, thereby strengthening public perception of information that is in fact false (Fa'izah & Kurniasari, 2026). Article 36 is further operationalized through Article 45 of the EIT Law, which provides criminal sanctions for violations resulting in the losses referred to in Article 36. Article 45 classifies violations of Article 36 as criminal offenses punishable by imprisonment and/or fines. The relationship between Article 36 and Article 45 demonstrates that the legal construction of the EIT Law places harm as the central entry point for law enforcement, thereby focusing on the consequences (result-based liability) rather than the prevention of technological risks at an early stage. In the context of deepfake, this approach illustrates that the law only becomes effective once actual harm has occurred, rather than at the stage of risk mitigation or technological control prior to misuse. This further emphasizes the repressive and ex post facto character of the EIT Law in responding to digital technological developments.

Various provisions within the EIT Law provide legal instruments to address deepfake misuse, including deepfake pornography (Article 27 paragraph (1)), defamation through deepfake (Article 27 paragraph (3)/27A), dissemination of deepfake-based hoaxes (Article 28), threats and intimidation (Article 29), and harm caused by deepfake (Article 36). However, this approach is fragmented and analogical, relying heavily on law

enforcement interpretation. Its weakness lies in the fact that the EIT Law does not yet provide an explicit definition of "deepfake" nor preventive mechanisms, thus remaining predominantly repressive in nature (Akbar et al., 2026). As a result, enforcement is often difficult and not specifically tailored to AI-related cases, and the application of these provisions remains dependent on interpretative discretion, leading to a lack of legal certainty.

Law No. 27 of 2022 concerning Personal Data Protection (PDP Law) serves as a specific legal foundation for ensuring the security and rights of personal data subjects throughout the data processing lifecycle. The increasingly complex development of digital technology places personal data as a legal element highly vulnerable to misuse, including through artificial intelligence technologies such as deepfake (Lestari et al., 2026). The PDP Law is relevant because it regulates the protection of personal data that may be used in the creation of AI-based manipulative content. Normatively, the PDP Law does not explicitly regulate or define deepfake technology (Arvitto, 2025). However, its provisions can be used to address unlawful personal data processing, including the use of biometric data such as a person's face, voice, and identity characteristics in digital systems (Susanti et al., 2026).

Article 20 paragraph (1) of the PDP Law stipulates that every personal data controller must have a lawful basis for data processing prior to carrying out data processing activities, which may include explicit consent from the data subject, fulfillment of contractual obligations, compliance with legal requirements, protection of vital interests of individuals, execution of public service duties, or other legitimate interests that must still balance the rights of the data subject. This provision emphasizes that without consent or a clear legal basis, the processing of personal data is prohibited (Kusnadi, 2021). This provision is highly relevant because the creation of deepfake content inherently involves the use of a person's biometric data, particularly facial and voice data, which constitute protected personal data. For instance, the creation and distribution of deepfake content featuring another person's face without consent constitutes unlawful personal data processing and therefore violates the PDP Law (Salsabilla, 2026). When a person's face is recorded, processed, and reconstructed using AI technology without valid consent, such action may be classified as personal data processing without a lawful basis as required under Article 20 paragraph (1).

Article 16 of Law Number 27 of 2022 concerning Personal Data Protection (PDP Law) expands the understanding of the scope of "personal data processing," which is not limited to initial technical actions such as data collection and storage, but also includes the entire range of activities related to the use, management, and distribution of such data. Under this provision, personal data processing is comprehensively understood as any operation performed on personal data, including displaying, disclosing, disseminating, modifying, deleting, and destroying personal data, whether carried out automatically or manually within electronic systems (Alfitri et al., 2024). This broad interpretation becomes relevant because a person's facial features, voice, and identity characteristics used in AI systems fundamentally constitute personal data that has been transformed into synthetic content (Shuen et al., 2026). When facial data is extracted, processed through

algorithms, and subsequently reconstructed into audio or video resembling certain actions or speech, the entire process may be construed as part of personal data processing as defined under Article 16. The dissemination of deepfake outputs is also included within the category of data disclosure, as the individual's identity remains the central element exposed to the public, even though it has been digitally manipulated.

The PDP Law explicitly prohibits any form of unlawful personal data processing (Arrasuli & Fahmi, 2023). This is regulated under Article 65 paragraphs (1), (2), and (3). Article 65 paragraph (1) stipulates that any person is prohibited from unlawfully obtaining or collecting personal data that does not belong to them with the intent of benefiting themselves or others and causing harm to the data subject. This provision establishes strict legal boundaries against any form of data acquisition that is not based on consent or a lawful legal basis. Furthermore, Article 65 paragraph (2) expands this prohibition by stipulating that any person is also prohibited from unlawfully disclosing or granting access to another person's personal data, meaning that protection covers not only data collection but also its distribution and exposure to third parties without valid authorization. Article 65 paragraph (3) reinforces this protection framework by prohibiting any unlawful disclosure of another person's personal data, thereby confirming that any act potentially violating individual privacy through data dissemination falls within the scope of legal violations. The creation of deepfakes generally involves the unauthorized collection of facial data, voice data, or biometric identity information, which is then processed into synthetic content that appears realistic, such as pornography or criminal-related depictions without the victim's knowledge (Anuar, 2026). Such facial data, although modified through artificial intelligence algorithms, remains fundamentally personal data that is unlawfully used and subsequently reproduced in audiovisual form for public dissemination. This practice not only fulfills the element of "unlawful collection" under Article 65 paragraph (1) but also constitutes "disclosure" under paragraphs (2) and (3), as the manipulated output ultimately re-exposes the individual's identity to the public without consent.

Article 66 of the PDP Law specifically prohibits any unlawful processing of personal data through the creation or falsification of personal data. It stipulates that any person is prohibited from creating false personal data or falsifying personal data with the intent to benefit themselves or others, which simultaneously results in harm to the data subject. Article 66 is highly relevant to deepfake practices because the technology fundamentally works by reconstructing, modifying, and generating a person's identity in audiovisual form that appears authentic but is in fact the result of algorithmic manipulation (Prayoga & Tuasikal, 2025). This process may be construed as a form of personal data falsification, as the facial features, voice, and expressions generated no longer represent the actual condition of the data subject but have been digitally altered to create a new illusion of reality (Anuar, 2026). The use of deepfakes for purposes such as fraud, manipulation of public opinion, or non-consensual sexual exploitation strengthens the fulfillment of the mens rea requirement under Article 66, namely the intention to benefit oneself or others accompanied by potential or actual harm to the victim. Such

falsification of identity through deepfake-generated synthetic personal data poses a significant risk of harm to the data subject.

Articles 67 and 68 of the PDP Law function as sanction provisions that impose criminal consequences for serious violations of personal data processing that result in harm to data subjects. Article 67 specifically provides that any person who intentionally violates Article 65. namely by obtaining, collecting, disclosing, or using another person's personal data without authorization may be subject to imprisonment of up to five (5) years and/or a fine of up to IDR 5,000,000,000. This provision aims to create deterrence against any form of unauthorized data misuse. Article 68 further expands the scope of sanctions to more serious conduct, namely personal data falsification, carrying a maximum imprisonment of six (6) years and/or a fine of up to IDR 6,000,000,000, thereby emphasizing that identity manipulation is considered a more severe offense than mere unauthorized data acquisition. Deepfake creation generally involves the collection of biometric data such as facial and voice data without the data subject's consent, which is then processed and manipulated into realistic synthetic audiovisual content, and therefore may be classified as a violation of either Article 67 or Article 68 depending on the nature of the act (Anuar, 2026). Deepfakes not only involve unauthorized data use but also generate false identities resembling specific individuals in a misleading manner, thus fulfilling the elements of personal data falsification under Article 68 in a substantive sense (Prayoga & Tuasikal, 2025).

In responding to the development of artificial intelligence technologies, including deepfake, the Indonesian Criminal Code (KUHP) remains a legal instrument used to prosecute offenders, even though it was not designed to address AI-based crimes. Normatively, the KUHP does not specifically recognize or regulate deepfake technology (Oktavianti et al., 2025). Therefore, any form of deepfake misuse under Indonesian criminal law must be classified through interpretation of general offenses such as falsification, fraud, and defamation (Aryawan, 2026). This demonstrates that the KUHP functions as a general criminal law that relies on analogy to address new forms of crime in the digital sphere, and it also shows that the KUHP remains reactive and has not yet adapted to developments in digital technology (Anggoro & Santoso, 2025).

One of the provisions in the Indonesian Criminal Code (Kitab Undang-Undang Hukum Pidana/KUHP) that is frequently used as a reference in addressing the misuse of digital technology is Article 263 concerning the criminal offense of document or instrument falsification. This article essentially stipulates that any person who makes a false document or falsifies a document that may give rise to rights, obligations, or debt discharge, with the intention of using it or instructing others to use it as if it were genuine and not falsified, may be criminally liable for committing falsification that harms the legal interests of others. The essential element of this provision lies in the act of manipulating a document that has evidentiary value or legal consequences, as well as the intention to treat the falsified result as authentic or legitimate (Intani et al., 2026). Although the object being manipulated is no longer in the form of a physical document as originally intended under Article 263, the fundamental principle of falsification under this article may be extended analogically when audiovisual content is used to create the impression that a

statement, action, or confession was genuinely made by a person (Prayoga & Tuasikal, 2025). Deepfake videos or audio recordings depicting individuals as if they were making certain statements, signing documents, or performing actions with legal or social implications can, in essence, be understood as a form of evidentiary fabrication that resembles an authentic document in a functional sense, although not in physical form (B. W. Y. Putri, 2025).

In addition to the offense of falsification, the provision on fraud as regulated under Article 378 of the KUHP is also highly relevant in the context of deepfake misuse, particularly when such technology is used as a means to unlawfully obtain benefits through deceptive information engineering (Silalahi & Dachi, 2026). Article 378 essentially stipulates that any person who, with the intent to unlawfully benefit themselves or others, uses a false name, false capacity, deception, or a series of lies to induce another person to surrender property, incur debt, or write off a receivable, may be punished for fraud. The core element of this provision lies in the presence of mens rea (criminal intent) to obtain unlawful gain, as well as active deceptive conduct used to influence the victim's will. In the context of deepfake, these elements are highly relevant because the technology enables perpetrators to create highly realistic visual and audio representations that can generate a strong sense of authenticity for both victims and third parties (Prayoga & Tuasikal, 2025). Simulated facial expressions, voices, or actions generated through deepfake can function as deceptive tools with strong persuasive power, thereby convincing victims to undertake actions that are detrimental to them, such as transferring money, granting data access, or making economic decisions based on false information (Sofia et al., 2026). The existence of deepfake thus expands the concept of deception under Article 378 KUHP from merely verbal statements or written documents to highly convincing multimedia manipulation in both visual and auditory forms.

Articles 310 and 311 of the KUHP are the primary provisions governing defamation and libel, which in their development have increasingly been associated with the misuse of digital technology, including deepfake. Article 310 essentially stipulates that any person who intentionally attacks another person's honor or reputation by accusing them of a certain fact with the intent that it be publicly known whether orally or in writing, may be punished for defamation. The essential element of this provision lies in the act of attacking a person's dignity in a public manner and the dissemination of accusations to the public (N. P. M. Putri et al., 2024). Article 311 strengthens this offense by stipulating that if the accusation is proven false and the perpetrator is aware of its falsity, the act is classified as libel, indicating a stronger form of mens rea compared to ordinary defamation. These provisions are highly relevant because deepfake technology enables the creation of audiovisual content that resembles reality, which can be used to fabricate scenarios in which a person appears to have performed or stated something that never actually occurred. The main strength of deepfake lies in its ability to produce visually convincing evidence, meaning that accusations are no longer merely verbal or written, but reinforced by highly persuasive visual and audio representations (Sutiyono et al., 2026). This significantly increases the harm to the victim's reputation, as society tends to trust visual evidence more easily than narrative claims alone.

Article 368 of the KUHP stipulates that any person who, with the intent to unlawfully benefit themselves, forces another person through violence or threats of violence to surrender property, incur debt, or write off a receivable may be criminally liable. Deepfake technology is often used as a tool of psychological pressure, for example by threatening victims with the dissemination of deepfake pornography unless they comply with certain demands, such as transferring money (Nirmala & Rahmania, 2025). Although such threats do not always involve physical violence, their intimidating nature can be construed as a form of coercion against the victim (Nurdin & Nugraha, 2025). Article 335 KUHP, which regulates unlawful coercion or unpleasant acts, may also be applied in certain contexts where deepfakes are used to harass, intimidate, or psychologically pressure victims. Deepfakes used for digital stalking, synthetic-image-based harassment, or personal intimidation may fall within the scope of this provision.

One of the fundamental problems in the legal regulation of deepfake misuse in Indonesia is the absence of an explicit normative definition in all applicable legislative instruments (Athallah et al., 2026). Neither the EIT Law, the PDP Law, nor the Criminal Code (KUHP) provides a conceptual boundary regarding what constitutes “deepfake” as a form of artificial intelligence-based manipulation technology. This definitional vacuum not only reflects regulatory lag in responding to technological development but also indicates the absence of a comprehensive legal understanding of deepfake’s generative, synthetic, and reality-indistinguishable characteristics (Zulfikar & Sihombing, 2025). In legal doctrine, the absence of a precise definition of a regulated object creates ambiguity in the scope of regulation, which in turn weakens the law’s capacity to address conduct that is substantively harmful but cannot be clearly qualified under existing normative texts. Without a clear definition, law enforcement authorities are placed in interpretative positions that are vulnerable to subjectivity and inconsistency across different jurisdictions (Nurkholisah et al., 2025a). Deepfake is thus understood only through academic discourse and practical interpretation by law enforcement authorities, rather than through a standardized legal construction. Consequently, law enforcement becomes heavily dependent on case-by-case interpretation, which may result in inconsistency (Kartadinata, 2026). This lack of normative clarity also leads to blurred classification of legal acts, given that deepfake may be used in a wide range of contexts, from creative expression to serious criminal offenses.

The absence of a legal framework that explicitly recognizes artificial intelligence as a distinct regulatory subject also raises several serious issues (Amboro & Komarhana, 2021). Existing regulations continue to position technology, including AI, merely as a tool in digital activities without considering its autonomous and adaptive characteristics. In fact, generative AI such as deepfake possesses the capability to produce new content that is highly realistic in both visual and audio forms, making it no longer comparable to conventional technologies. Consequently, legal issues involving AI must be forced into existing legal norms, even when they are not always substantively relevant. This creates a gap between rapidly evolving technological development and relatively stagnant legal responses. The law becomes reactive rather than anticipatory, thereby failing to provide a risk-control framework at the early stages of technological development.

The Indonesian law enforcement system's reliance on analogical interpretation of general provisions that were not designed to address the specific characteristics of deepfake ultimately results in serious and systemic legal uncertainty. The prevailing analogical approach, where provisions on defamation, falsification, fraud, or decency violations under the EIT Law and KUHP are interpreted as encompassing deepfake-related conduct, essentially represents an attempt to adapt outdated legal norms to new technological realities that far exceed the imagination of the original lawmakers (Wicaksono et al., 2026). This reliance on analogy not only creates uncertainty in the legal qualification of unlawful acts but also generates inconsistency in the application of sanctions, as different law enforcement officers may rely on different provisions for substantively similar cases (Nurkholisah et al., 2025b). Such an analogical approach demonstrates that Indonesian law remains fragmented in responding to emerging technologies, as it reacts only after harm occurs rather than anticipating risks from the outset (Ibrahim et al., 2025). In the context of a rule-of-law state, this condition is problematic because legal certainty is a fundamental principle that should ensure predictability and consistency in law enforcement. Therefore, the use of general provisions through analogical interpretation is no longer sufficient to address the complexity of deepfake as an AI-based form of crime.

In contrast to Indonesia's approach, which still relies on a fragmented, analogical, and reactive sectoral legal regime, the European Union, through the European Union Artificial Intelligence Act (EU AI Act), has established a regulatory framework that explicitly recognizes artificial intelligence as a distinct legal subject under a risk-based regulation approach (Nicholle, 2025). The EU AI Act clearly classifies AI systems into several risk categories (minimal, limited, high, and unacceptable risk), each carrying different legal consequences ranging from transparency obligations to outright prohibition (Zulkarnain & Rahmayanti, 2026). Specifically with regard to deepfakes, the regulation explicitly requires labeling of content that has been manipulated or generated by AI, as stipulated in the transparency obligations for AI systems that generate or modify image, audio, or video content resembling reality. This approach reflects a paradigm shift from ex post liability-based enforcement to ex ante risk-based prevention, where potential harm is identified and controlled at the design stage of the system. Compared to Indonesia, which still relies on general provisions under the EIT Law, PDP Law, and KUHP through analogical interpretation without a clear risk classification framework, the EU AI Act offers greater legal certainty by providing definitions, risk categories, and specific obligations directly applicable to deepfake technology. This difference demonstrates that the main issue within the Indonesian legal system is not only the absence of a definition of deepfake, but also the lack of adoption of a risk-based regulatory approach capable of accommodating the dynamic, adaptive, and potentially systemic impacts of AI on society.

## Risk-based Regulation Approach in Regulating Deepfake

The risk-based regulation approach is a legal regulatory model that places the level of risk as the primary basis for determining the form, intensity, and scope of

regulation applied to a particular technology (Prastia, 2025). In essence, risk-based regulation is a model that not only focuses on enforcement after violations occur but also emphasizes the identification of risk levels posed by a technology and the adjustment of legal instruments according to the degree of harm generated (Prasetyo, 2024). Risk-based regulation is generally understood as an approach that classifies risks into several levels, namely low risk, medium risk, and high risk (European Parliament and Council of the European Union, 2024). These levels determine the intensity of supervision, compliance obligations, and sanctions based on the degree of risk involved. This approach is increasingly relevant due to the dynamic, adaptive nature of AI systems and their high level of uncertainty (Nurhayati et al., 2026). It is particularly relevant in the context of deepfake technology, which has a dual-use nature, it can be used for legitimate purposes such as entertainment, education, and research, but also carries a high risk of misuse for fraud, disinformation, and violations of human rights (Usman, 2025).

Conceptually, risk-based regulation does not aim to prohibit a technology entirely, but rather to categorize the level of risk posed by the technology to individuals and society. This approach enables regulators to distinguish between low, medium, and high-risk uses, allowing regulatory frameworks to be designed proportionally in accordance with the potential impact of the technology (European Parliament and Council of the European Union, 2024). A risk-based approach is particularly relevant because deepfake technology has a broad spectrum of applications, ranging from entertainment and educational purposes to criminal misuse (Widjaja, 2025). This dual-use characteristic means that deepfake cannot be regulated uniformly but must instead be classified based on the level of risk posed in each specific use case (Usman, 2025). One of the most prominent global implementations of risk-based regulation is the European Union Artificial Intelligence Act (EU AI Act), which classifies AI systems into several risk categories, each of which carries different legal consequences ranging from transparency obligations to outright prohibition (European Parliament and Council of the European Union, 2024).

Within this risk classification framework, the EU AI Act provides a clear normative differentiation of AI system risk levels through category-based regulation. In the high-risk category, the regime does not impose a total ban but instead establishes a comprehensive set of obligations throughout the AI system lifecycle. Under Article 6 of the EU AI Act, an AI system is classified as high-risk if it meets either of two conditions: first, if it serves as a safety component of a product regulated under Union harmonisation legislation listed in Annex I, such as medical devices, vehicles, and industrial machinery; or second, if it falls under use cases listed in Annex III, which include biometric identification, critical infrastructure, education, employment, essential services, law enforcement, migration, and judicial administration. Once classified as high-risk, such systems are subject to extensive obligations. Article 13 requires providers to supply sufficient information to deployers, including characteristics, intended purpose, limitations, accuracy levels, robustness, and cybersecurity measures (Butt, 2024). Article 14 emphasizes human oversight, ensuring that humans can understand, interpret, and, where necessary, override or stop AI outputs. Article 15 establishes technical

requirements regarding accuracy, robustness, and cybersecurity to ensure system reliability throughout its lifecycle. Non-compliance with these obligations may result in administrative sanctions under Article 99(4) of the EU AI Act of up to €15,000,000 or 3% of global annual turnover, whichever is higher.

The limited-risk category is regulated under Article 50 of the EU AI Act, which adopts a lighter yet still mandatory transparency-based approach. Article 50 regulates five types of AI systems within this category. First, under Article 50(1), chatbots and virtual assistants that interact directly with humans in real time must inform users that they are interacting with an AI system, unless this is obvious from the context. Second, under Article 50(2), generative AI systems that produce synthetic content such as audio, images, video, or text (e.g., ChatGPT, DALL·E, or Midjourney) must ensure that outputs are marked in a machine-readable format to indicate AI-generated content. Third, under Article 50(3), emotion recognition systems and biometric categorisation systems must inform affected individuals that such systems are operating on them. Fourth, under Article 50(4), AI systems that generate or manipulate deepfake content in the form of images, audio, or video depicting real persons, places, or events must explicitly disclose that the content has been artificially generated or manipulated, with limited exceptions for law enforcement purposes and clearly identifiable artistic works. Fifth, also under Article 50(4), AI systems generating text intended to inform the public on matters of public interest, such as journalism or governmental communication, must label such content as AI-generated. Non-compliance with these transparency obligations is subject to sanctions under Article 99(4) of the EU AI Act, with administrative fines of up to €15,000,000 or 3% of global annual turnover, whichever is higher. AI systems categorized as minimal risk are not explicitly regulated under the EU AI Act and are not subject to binding legal obligations but are instead encouraged to adopt voluntary codes of conduct.

Deepfake technology may be classified as high-risk technology when used to manipulate a person's identity without consent, particularly when it affects widely impactful domains such as public reputation, financial security, and social stability (Lestari et al., 2026). However, at the same time, deepfake technology may also be used for low-risk purposes such as entertainment, education, and other activities with minimal harm potential (Widjaja, 2025). The determination of risk level does not depend solely on the technology itself, but also on the context of use, the perpetrator's intent, and the actual or potential impact generated.

Under the EU AI Act, deepfakes are explicitly categorized as AI systems requiring specific transparency obligations (European Parliament and Council of the European Union, 2024). Article 3(60) defines deepfake as "AI-generated or manipulated image, audio or video content that resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful." All AI-generated or AI-modified content must be labeled so that users are aware that the content does not originate from reality (Hermansyah, 2025). This demonstrates that the risk-based approach is not only repressive in nature but also emphasizes preventive mechanisms through information transparency. Compared to Indonesia, regulation of deepfake has not yet explicitly adopted a risk-based approach. The EIT Law, PDP Law, and KUHP still

regulate based on types of conduct rather than levels of technological risk. As a result, all forms of deepfake misuse are treated uniformly without differentiation of risk levels. The absence of risk classification in the Indonesian legal system makes regulation less adaptive to technological development. In reality, not all uses of deepfake have the same impact on society. With a risk classification framework, regulators can establish differentiated standards between low-risk and high-risk uses, thereby making regulation more proportionate and avoiding over-regulation.

The risk-based regulation approach also emphasizes the importance of the principle of transparency in the use of AI technology. In the context of deepfake, transparency can be implemented through mandatory content labeling requirements for AI-generated outputs (Hermansyah, 2025) This principle requires that content generated or manipulated by AI be clearly marked with labels such as “AI-generated” or “manipulated content.” This regulatory design demonstrates that deepfake technology is not prohibited in absolute terms, as it may still be used for legitimate purposes such as entertainment, education, and digital creativity. However, due to its significant risk potential—such as deceiving the public, damaging individual reputation, and spreading disinformation that may influence public opinion—the regulatory approach is not prohibition-based, but rather control-based through transparency mechanisms. This principle aims to enable the public to distinguish between authentic and manipulated content, thereby reducing the risk of misinformation. As an illustration, the United Kingdom, the European Union, and China have adopted obligations to label or disclose AI-generated content for transparency purposes (Respati, 2024). In Indonesia, however, there is currently no specific regulation governing AI content labeling or AI labelling, and therefore no legal obligation for platforms or developers to mark deepfake content as artificially generated. This normative gap indicates that Indonesia has not fully adopted the principle of transparency. Indonesia may adopt a similar concept by requiring platform providers to label deepfake videos or disclose AI usage at the point of content upload, in order to reduce the risk of misuse.

The risk-based regulation approach also explicitly assigns responsibility not only to end-users but also to digital platform providers and artificial intelligence developers (Puannandini et al., 2025). Within the framework of the European Union Artificial Intelligence Act (EU AI Act), this is systematically regulated through the distribution of obligations based on the roles of actors in the AI ecosystem, particularly between providers (AI system developers) and deployers (AI system users). The primary obligations of providers of high-risk AI systems are governed under Article 16 of the EU AI Act, which requires them to ensure that the AI system complies with all regulatory requirements, including risk management, data quality, technical documentation, transparency, and human oversight mechanisms (Balqis et al., 2025). In addition, Article 9 of the EU AI Act mandates the establishment of a risk management system designed to identify, analyze, and mitigate risks that may arise throughout the AI system lifecycle, including potential misuse such as deepfake generation. Risk assessment enables the evaluation of potential technological impacts before deployment, thereby minimizing risks at the earliest stage of development (Pratama et al., 2024). In this context, developers

are not only responsible for the technical functionality of AI systems but also for their potential societal impacts. Meanwhile, Article 26 of the EU AI Act imposes obligations on deployers to use AI systems in accordance with instructions for use, to conduct operational oversight, and to ensure that system usage does not create disproportionate risks to individual rights and interests. This division of responsibility is particularly relevant because misuse may originate not only from end-users but also from system design that enables highly realistic synthetic content generation.

The relevance of the risk-based regulation approach to deepfake governance in Indonesia cannot be separated from the inherent characteristics of the technology itself, which is dual-use, dynamic, and asymmetrical in its impact (Usman, 2025). Deepfake technology cannot be treated as a single entity that is uniformly harmful or entirely safe, as its usage spectrum is highly diverse and context dependent. The regulatory approach must therefore respond proportionally to this heterogeneity, rather than imposing uniform restrictions that may either hinder innovation or fail to prevent misuse. The principle of proportionality requires that regulatory burdens imposed on legal subjects be commensurate with the level of risk generated by their conduct. Deepfake used for educational or entertainment purposes should not be subject to the same legal obligations as deepfake used for political disinformation or non-consensual pornography (Widjaja, 2025).

A risk-based regulation approach enables law to function in an anticipatory manner, namely by identifying, assessing, and responding to potential harms before actual damage occurs, thereby shifting the regulatory paradigm from a purely repressive (ex post) model to a preventive (ex-ante) model. This paradigm shift is theoretically consistent with the principle of preventive regulation in modern legal theory, which prioritizes risk prevention rather than merely imposing sanctions after violations occur. Normatively, this approach is reflected in the European Union Artificial Intelligence Act (EU AI Act), which explicitly adopts a risk classification system (minimal, limited, high, and unacceptable risk) as the basis for legal obligations, where AI systems with certain risk levels are subject to transparency requirements, risk management obligations, or outright prohibition when risks are deemed unacceptable (Zulkarnain & Rahmayanti, 2026). For example, the obligation to implement a risk management system under Article 9 and the comprehensive compliance obligations for providers of high-risk AI systems under Article 16 demonstrate that regulation does not wait for harm to occur but instead governs from the design and development stage of technology. In the context of deepfake, this approach is particularly relevant because risks of identity manipulation and disinformation can already be identified at the early stages of generative system deployment, rather than only after harmful content has been disseminated.

As a normative recommendation, Indonesia should adopt a risk classification model for deepfake tailored to its national social, legal, and technological context, dividing deepfake usage into three main categories: low risk, medium risk, and high risk. The low-risk category includes the use of deepfake for entertainment, film production, animation, and educational purposes, where consent from the data subject has been explicitly obtained and the content does not significantly mislead the public, thus

requiring only AI-content labeling obligations for transparency. The medium-risk category includes the use of deepfake in political content, commercial advertising, and public campaigns, where the potential for public perception manipulation is higher, requiring stricter disclosure obligations, identity verification of content creators, and reporting mechanisms to competent authorities. The high-risk category includes the use of deepfake for non-consensual pornography, financial fraud, identity impersonation of public figures for manipulative purposes, and large-scale disinformation that may threaten social stability and democratic processes, where legal obligations must be both preventive and repressive, including explicit prohibitions, mandatory content removal, and severe criminal sanctions. The logic underlying this classification is proportionality: the greater the potential harm to individual rights, information integrity, and social security, the heavier the obligations imposed on creators, disseminators, and platform providers. This model also enhances legal certainty, as legal subjects can understand ex ante the risk category of their activities and the associated legal consequences, thereby reducing reliance on analogical interpretation that is prone to inconsistency (Orwat et al., 2022).

Prior to the publication or commercial distribution of a generative AI system with deepfake capabilities in Indonesia, developers and platform providers are required to undergo a series of comprehensive pre-deployment procedures as an integral component of implementing a risk-based regulation approach. These procedures include impact assessment and watermarking obligations. Impact assessment refers to a systematic evaluation process of the potential impacts that an AI system may generate on individual rights, public information stability, and social security, whereby developers are required to transparently document risk-related usage scenarios as well as the mitigation measures designed within the system (Khair & Ahyar Wiraguna, 2025). Meanwhile, the watermarking obligation, namely the insertion of an indelible or difficult-to-remove digital marker on AI-generated content, serves as a transparency instrument that enables users, law enforcement authorities, and regulatory bodies to technically identify the origin of synthetic content.

The differentiation of sanctions based on risk categories constitutes one of the most decisive elements in ensuring that risk-based regulation operates effectively as a legal instrument that is both preventive and proportionate. A uniform sanctioning system, in which all deepfake-related violations are subject to identical legal consequences regardless of the level of harm caused, not only creates legal injustice but also fails to provide appropriate incentives for compliance. Within the proposed framework, violations in the low-risk category, such as failure to include AI labeling on entertainment or educational content that does not cause significant harm, should be subject only to administrative sanctions, including warnings, corrective orders, and proportionate administrative fines based on the scale of the entity involved. Violations in the medium-risk category, such as the dissemination of political or commercial deepfake content without adequate disclosure that has the potential to mislead the public, may be subject to more severe administrative sanctions, including mandatory immediate content removal, higher monetary penalties, and possible suspension of platform operating licenses. Meanwhile, violations in the high-risk category, such as the creation or

dissemination of non-consensual deepfake pornography, identity-based fraud, or large-scale disinformation capable of undermining social stability and democratic processes, must be subject to severe criminal sanctions that reflect the seriousness of the offense against human dignity and public order, including significant imprisonment terms and deterrent-level fines. This differentiation of sanctions is not merely a legal instrument, but a normative statement regarding the values that the Indonesian legal system seeks to protect in the era of artificial intelligence, namely that the protection of individual dignity, information integrity, and democratic stability occupies the highest hierarchy of legal interests and therefore requires the strongest level of legal protection.

## Conclusion

Overall, the legal framework in Indonesia addressing the misuse of deepfake technology remains lacking in specific and comprehensive normative provisions. Existing legal instruments, namely the EIT Law, the PDP Law, and the Criminal Code (KUHP), essentially provide only general regulatory foundations through the expanded interpretation of existing provisions, such as prohibitions on defamation, dissemination of false information, unlawful personal data processing, as well as falsification and fraud. However, none of these regulations explicitly govern deepfake as a form of artificial intelligence technology with unique characteristics, particularly its ability to generate highly realistic synthetic content. This condition indicates limitations in normative adequacy, as the existing legal framework is not yet capable of addressing the complexity of problems posed by deepfake technology in a systemic manner. In addition, the current regulatory framework does not incorporate a risk-based preventive mechanism that can anticipate potential misuse at an early stage. In this context, the risk-based regulation approach becomes relevant as an alternative regulatory framework that is more adaptive, proportional, and preventive, as it enables the classification of deepfake usage based on risk levels and the determination of differentiated legal obligations according to the degree of harm involved. Accordingly, it can be concluded that the implementation of risk-based regulation is necessary to fill the existing legal gap and to strengthen the Indonesian legal system so that it becomes more responsive to the development of artificial intelligence technology in the digital era.

## References

Akbar, B. M., Yulia, R., & Romdoni, M. (2026). Urgensi Reformulasi Undang-Undang Informasi dan Transaksi Elektronik terhadap Penyalahgunaan Kecerdasan Buatan Deepfake dalam Perspektif Pembaharuan Hukum Pidana. *ALADALAH : Jurnal Politik, Sosial, Hukum Dan Humaniora*, *4*(2), 1–20.

Al-Ayoubi, S. J., & Suharto, M. A. (2025). Pemidanaan Kepada Pelaku Konten Pornografi Menggunakan Aplikasi Deepfake Pada Peraturan Perundang-Undangan di Indonesia. *UNES Law Review*, *8*(1), 163–173. https://doi.org/10.31933/unesrev.v8i1.2464

Alfitri, N. A., Rahmawati, & Firmansyah. (2024). Perlindungan Terhadap Data Pribadi di Era Digital Berdasarkan Undang-Undang Nomor 27 Tahun 2022. *Journal Social Society*, *4*(2), 92–111. https://doi.org/10.54065/jss.4.2.2024.511

Alif, S. S. Al, Rindiani, A., & Marhayati, C. (2025). Konstruksi Hukum Pidana Dalam Penanggulangan Kejahatan Siber Berbasis Teknologi Deepfake di Indonesia. *Legal Standing: Jurnal Ilmu Hukum*, *9*(5), 1169–1183. https://news.detik.com/berita/d-3567290/polling-58-masyarakat-puas-kinerja-kpk

Amboro, F. Y. P., & Komarhana, K. (2021). Prospek Kecerdasan Buatan sebagai Subjek Hukum Perdata di Indonesia. *Law Review*, *XXI* (2), 193–217.

Anggoro, T. C., & Santoso, B. (2025). Politik Kriminal Integratif dalam Penanggulangan Perjudian Onlinedi Indonesia. *Jurnal USM Law Review*, *8*(3), 1643–1664. https://doi.org/10.26623/julr.v8i3.12541

Anuar, M. A. (2026). *Pengaturan Hukum Terhadap Penyalahgunaan Artificial Inteligence (AI) Dalam Konteks Deepfake Pornografi Di Indonesia*. Skripsi: Universitas Jambi.

Apandi, M., Rahayu, K., Agus Prayugo, W., & Ariany, L. (2025). Kekaburan Norma dalam Kebebasan Berekspresi di Era Digital: Analisis Undang-Undang Nomor 1 Tahun 2024 tentang Perubahan Atas Undang-Undang Nomor 11 Tahun 2008 tentang Informasi dan Transaksi Elektronik. *Jurnal Hukum Lex Generalis*, *5*(12). https://doi.org/10.56370/jhlg.v5i12.1007

Arrasuli, B. K., & Fahmi, K. (2023). Perlindungan Hukum Positif Indonesia Terhadap Kejahatan Penyalahgunaan Data Pribadi. *Unes Journal of Swara Justisia*, *7*(2), 369–392.

Arvitto, R. S. (2025). Implikasi Hukum Deepfake: Telaah terhadap UU ITE dan UU PDP. *Jurnal Ilmiah Hukum Dan Hak Asasi Manusia*, *4*(2), 73–82. https://doi.org/10.35912/jihham.v4i2.3937

Aryawan, I. P. G. (2026). Kekosongan Norma Hukum dalam Penanganan Deepfakesebagai Alat Kejahatan Siber: Studi Yuridis di Wilayah Polri Klungkung. *Menulis : Jurnal Penelitian Nusantara*, *2*(4), 295–307.

Athallah, N. B. P., Candra, A., & Sanusi, H. A. (2026). Urgensi dan Kecukupan Pengaturan Deepfake dalam Undang- Undang Informasi dan Transaksi Elektronik di Indonesia. *Judge: Jurnal Hukum*, *06*(09), 1665–1676.

Azis, A. (2025). Rekonstruksi Hukum Penyalahgunaan Artificial Intelligence(AI) Dalam Konten Kesusilaan Dan Pencemaran Nama Baik. *Rewang Rencang : Jurnal Hukum Lex Generalis*, *6*(7), 1–30.

Balqis, A. N., Ramadhani, S. N., & Sumual, A. K. (2025). Risiko Operasional dalam Teknologi Informasi: Kajian Literatur Mengenai Mekanisme Identifikasi, Penilaian, dan Dampaknya terhadap Manajemen Risiko. *Journal of Economics, Management, and Accounting*, *1*(1), 67–77. https://doi.org/10.65310/8hq2hc60

Basah, D. A. Y., Wijaya, A., & Januardy, I. (2025). Kriminalisasi Pelanggaran Protokol Digital: Tinjauan Hukum Pidana Terhadap Penyebaran Deepfake di Media Sosial. *Innovative: Journal Of Social Science Research*, *5*(4), 386–398.

Butt, J. S. (2024). Analytical Study of the World's First EU Artificial Intelligence (AI) Act, 2024. *International Journal of Research Publication and Reviews*, *5*(3), 7343–7364.

Devi, B. C. (2025). Etika Komunikasi dan Tantangan Regulasi atas Konten AI-Generated di TikTok : Studi Disinformasi Visual di Indonesia. *Seminar Nasional Fakultas Hukum, Ilmu Sosial, Dan Ilmu Politik*, 1–9.

Disemadi, H. S. (2022). Lenses of Legal Research: A Descriptive Essay on Legal Research Methodologies. *Journal of Judicial Review*, *24*(2), 289–304. https://doi.org/10.37253/jjr.v24i2.7280

Domuari, A., & Yudiantara, I. G. N. N. K. (2026). Penyalahgunaan Deepfake Dalam Bentuk Pornografi Anak Menurut Hukum Positif Indonesia. *Jurnal Media Akademik*, *4*(1), 1–16.

European Parliament and Council of the European Union. (2024). *European Union Artificial Intelligence Act (EU AI Act) - Final Text*. European Parliament and Council of the European Union.

Fa'izah, F. N., & Kurniasari, R. J. (2026). Deep Fake Crime On Social Media: A Juridical Analysis Of The Spread Of Hoaxes And The Manipulation Of Public Opinion In State Security. *Yurijaya : Jurnal Ilmiah Hukum*, *7*(3), 522–550. https://doi.org/10.51213/yurijaya.v7i3.238

Hasanuddin, M., Nurfransiska, F., & Maryuliyanto. (2026). Pengaturan Artificial Intelligence (AI) Dalam Perspektif Hukum Indonesia: Analisis Normatif Atas Tntangan, Implikasi, Dan Model Regulasi Ideal. *Judge: Jurnal Hukum*, *6*(6), 1890–1897.

Hermansyah, M. G. R. N. (2025). Penggunaan Artificial Intelligence dan Deep Fake Analisis Kasus Pidana Video Deepfake yang Mencatut Presiden Prabowo Subianto Tahun 2025. *JIC: Jurnal Hukum Dan Konstitusi*, *1*(3), 132–146.

Ibrahim, M. K., Hafidz, M., & Atmaja, Z. F. H. (2025). Urgensi Pembaruan Hukum Dalam Menghadapi Kejahatan Yang Melibatkan Teknologi Kecerdasan Buatan AI. *Journal of Legal, Political, and Humanistic Inquiry*, *1*(2), 106–115. https://doi.org/10.65310/2dk3wh05

Intani, N. A., Widyastuti, I., & Simamora, H. S. (2026). Tinjauan Hukum Terhadap Penyalahgunaan Artificial IntelligenceDalam Pembuatan Konten Video Bermuatan Pelecehan Seksual Terhadap Perempuan. *Al-Zayn: Jurnal Ilmu Sosial & Hukum*, *4*(1), 173–183.

Kartadinata, A. (2026). Rekonstruksi Delik Pidana dalam Kejahatan Deepfake : Tantangan Pembuktian dan Perlindungan Korban. *Jurnal Ilmiah Hukum Dan Administrasi Publik (MUARAKUM)*, *2*(1), 36–48.

Khair, F., & Ahyar Wiraguna, S. (2025). Data Protection Impact Assessment (DPIA) sebagai Instrumen Kunci Menjamin Kepatuhan UU PDP 2022 di Indonesia. *Politika Progresif : Jurnal Hukum, Politik Dan Humaniora*, *2*(2), 246–254. https://doi.org/10.62383/progres.v2i2.1821

Kristiyenda, Y. S., Faradila, J., & Basanova, C. (2025). Pencegahan Kejahatan Deepfake: Studi Kasus terhadap Modus Penipuan Deepfake Prabowo Subianto dalam

Tawaran Bantuan Uang. *ALADALAH: Jurnal Politik, Sosial, Hukum Dan Humaniora*, *3*(2), 149–164. https://doi.org/10.59246/aladalah.v3i2.1263

Kusnadi, S. A. (2021). Perlindungan Hukum Data Pribadi Sebagai Hak Privasi. *AL WASATH Jurnal Ilmu Hukum*, *2*(1), 9–16. https://doi.org/10.47776/alwasath.v2i1.127

Lestari, P. A., Saputri, Z. R., Jamilah, R., Febiyanti, F., Fadilah, M. F., Firnanda, M. F., & Koeswara, T. S. N. (2026). Konten Ilegal Kecerdasan Buatan Berbasis Deepfake dan Upaya Penanggulangannya di Indonesia. *Jurnal Ilmiah Sistem Informasi*, *5*(1), 26–35. https://doi.org/10.51903/r5xwq018

Mecca, A. S. P., Aznul Hidaya, W., & Tuasikal, H. (2025). Pemanfaatan Teknologi Kecerdasan Buatan (Artificial Intelligence) Dalam Sistem Peradilan Pidana Di Indonesia. *Jurnal Sosial Teknologi*, *5*(6), 1730–1746. https://doi.org/10.59188/jurnalsostech.v5i6.32207

Nicholle, M. (2025). Pertanggung-jawaban Hukum atas Keputusan Otomatis Berbasis Kecerdasan Buatan: Studi Komparatif antara Uni Eropa dan Asia Tenggara. *Jaksa : Jurnal Kajian Ilmu Hukum Dan Politik*, *3*(4), 65–79. https://doi.org/10.51903/ejr3ep34

Nirmala, A. Z., & Rahmania, N. (2025). Transformation of Sexual Harassment in the Age of Artificial Intelligence: A Review of Indonesian Law. *Unizar Law Review*, *8*(1).

Nugroho, Y., & Halim, A. (2026). Rekonstruksi Tanggung Jawab Hukum Penyalahgunaan Artificial Intelligence di Indonesia Dalam Perspektif Regulasi Global. *FENOMENA*, *20*(01), 174. https://doi.org/10.36841/fenomena.v20i01.8262

Nurdin, S. W., & Nugraha, I. F. (2025). Ancaman Deepfake Dan Disinformasi Berbasis Ai: Implikasi Terhadap Keamanan Siber Dan Stabilitas Nasional Indonesia. *JIMR: Journal Of International Multidisciplinary Research*, *4*(vol), 84. http://azramedia-indonesia.azramediaindonesia.com/index.php/JIMR/index DOI: https://doi.org/10.62668/jimr.v4i01.1551

Nurhayati, S., Eddy Soeryanto Soegoto, & Tri Utomo Wiganarto. (2026). Model Technopreneurship Berbasis Explainable Artificial Intelligence (Xai) untuk Meningkatkan Kepercayaan dan Intensi Adopsi Teknologi pada Startup Digital. *EKONOMIKA45 : Jurnal Ilmiah Manajemen, Ekonomi Bisnis, Kewirausahaan*, *13*(2), 1744–1756. https://doi.org/10.30640/ekonomika45.v13i2.6141

Nurkholisah, S., Rismana, D., Nugroho, A. E., Munjiyah, A., & Ayunisa, Q. (2025a). Tantangan Kriminalisasi Deepfake dalam Hukum Pidana Indonesia Challenges in Criminalizing Deepfakes under Indonesian Criminal Law. *Jurnal Usm Law Review*, *8*(3), 1–12. https://journals.usm.ac.id/index.php/julr/article/view/13060

Oktavianti, R. G., Muqsith, A., & Nadia, I. (2025). Penyalahgunaan Deepfake Era Canggihnya Teknologi Memicu Kejahatan Media: Misuse of Deepfake in The Era of Advanced Technology Is Triggering Media Crime. *IUS: Jurnal Ilmiah Fakultas Hukum*, *13*(1), 114–122.

Orwat, C., Bareis, J., Folberth, A., Jahnel, J., & Wadephul, C. (2022). Normative Challenges of Risk Regulation of Artificial Intelligence and Automated Decision-Making. *SSRN Electronic Journal*, *01*. https://doi.org/10.2139/ssrn.4274828

Prasetyo, E. (2024). Kesesuaian Sistem Pengawasan Perizinan Berbasis Risiko di Sektor Lingkungan Hidup Pasca UU Cipta Kerja Dengan Prinsip Perlindungan dan Pengelolaan Lingkungan Hidup di Indonesia. *Prosiding Seminar Hukum Aktual Fakultas Hukum Universitas Islam Indonesia*, *2*(6), 84–102.

Prastia, Y. (2025). Keseimbangan Inovasi dan Kepatuhan: Implementasi Pelindungan Data Pribadi dalam Ekosistem Finansial Teknologi dan Aset Digital. *Journal of Financial Sector Law & Policy*, *1*(2), 262–302.

Pratama, F., Avini, T., Saputra, I., Putri, M. K., & Imamfajri, S. (2024). Analisis Risiko Keamanan Pada Pengembangan Perangkat Lunak Berbasis Cloud. *Jurnal Kecerdasan Buatan Dan Teknologi Informasi*, *3*(2), 61–67. https://doi.org/10.69916/jkbti.v3i2.130

Prayoga, H., & Tuasikal, H. (2025). Penyebaran Konten Deepfake Sebagai Tindak Pidana: Analisis Kritis Terhadap Penegakan Hukum Dan Perlindungan Publik Di Indonesia. *Abdurrauf Law and Sharia*, *2*(1), 22–38. https://doi.org/10.70742/arlash.v2i1.194

Puannandini, D. A., Fabian, R. M., Firdaus, R. A. P., Mustopa, M. Z., & Herdiyana, I. (2025). Liabilitas Produk Ai Dalam Sistem Hukum Indonesia: Implikasi Bagi Pengembang, Pengguna, Dan Penyedia Layanan. *Luris Studia: Jurnal Kajian Hukum*, *6*(1), 24–33.

Putri, B. W. Y. (2025). Perlindungan Hukum Hak Publisitas Terhadap Eksploitasi Teknologi Deepfake Ai Dalam Konteks Komersial Menurut Hukum Positif Di Indonesia. *NUSANTARA: Jurnal Ilmu Pengetahuan Sosial*, *12*(9), 3968–3978.

Putri, N. P. M., Hartono, M. S., & Yudiawan, I. D. G. H. (2024). Analisis Reformulasi Pertanggungjawaban Pidana Pengguna Teknologi Deepfake Dalam Tindak Pidana Pencemaran Nama Baik Berbasis Artificial Intelligence. *Jurnal Pacta Sunt Servanda*, *5*(2), 120–129. https://doi.org/10.23887/jpss.v5i2.5807

Putri, S. M. I., Salsabila, N., & Hosnah, A. U. (2024). Kriminalisasi Penggunaan Deepfake dalam Tindak Pidana Penipuan dan Pencemaran Nama Baik: Tantangan dan Solusi Hukum. *Jurnal Hukum Legalita*, *6*(2), 83–90. https://doi.org/10.47637/legalita.v6i2.1453

Redaksi Tribata. (2025). *Polisi Berhasil Menangkap Pelaku Sindikat Penipuan Modus Video Deepfake*. Tribata News.

Respati, A. A. (2024). Reformulasi UU ITE terhadap Artificial Intelligence Dibandingkan dengan Uni Eropa dan China AI Act Regulation. *Jurnal USM Law Review*, *7*(3), 1737–1758. https://doi.org/10.26623/julr.v7i3.10578

Salsabilla, A. (2026). Perlindungan Hukum terhadap Korban Perekaman Tidak Sah dan Penyebaran Konten Pribadi di Platform Digital Berdasarkan Undang-Undang Nomor 27 Tahun 2022 tentang Perlindungan Data Pribad. *Al-Zayn: Jurnal Ilmu Sosial & Hukum*, *4*(3), 3769–3781.

Shabrina, A. N., Naila, G. S., Nuryansyah, G. P., & Amanda, R. (2026). Bahaya Deepfake Dalam Penyalahgunaan Grok AI Di Platform X Sebagai Bentuk Disinformasi. *Integrative Perspectives of Social and Science Journal (IPSSJ)*, *3*(1), 195–207.

Shuen, Y. Y., Mariana, L., & Sihombing, L. (2026). Perlindungan Hukum Hak Atas Potret Publik Figur Terhadap Penggunaan Teknologi Generative AI oleh Penggemar. *Al-Zayn: Jurnal Ilmu Sosial & Hukum*, *4*(3), 8303–8311.

Silalahi, W., & Dachi, F. N. (2026). Pertanggung jawaban Hukum Platform Media Sosial terhadap Penyebaran Konten Deepfake Berbasis Artificial Intelligence dalam Kasus Penipuan Berkedok Ritual Uang Gaib Ujang Busthomi. *Politika Progresif: Jurnal Hukum, Politik Dan Humaniora*, *3*(2), 22–36.

Sofia, Shahrullah, R. S., & Situmeang, A. (2026). Tanggung Jawab Hukum atas Penyalahgunaan Teknologi AI: Studi Kasus Deepfake Prabowo Subianto dalam Modus Bantuan Uang. *Jurnal Media Hukum*, *26*(1), 165–191.

Susanti, D. K. E. D., Yuliartini, N. P. R., & Mangku, D. G. S. (2026). Analisis Yuridis Pertanggungjawaban Pidana Penyelenggara Sistem Elektronik Artificial Intelligence Atas Manipulasi Data Biometrik Orang Lain Dalam Penipuan Transaksi Elektronik. *Al-Zayn: Jurnal Ilmu Sosial & Hukum*, *4*(3), 125–138.

Sutiyono, H., Ihsan, M., Yahya, M., Syauqi, M., Puspita, I., Fitriyani, A., Fachrurozi, M., & Rahman, R. (2026). *Teori Belajar Cybernetik dalam Perspektif Islam & Etika Kejujuran di Era Teknologi Informasi* (N. L. Mahmudah & Y. N. Hadi (eds.)). Semarang: PT. Naya Sasmita Pustaka.

Syahra, A., Mecca, P., Hidaya, W. A., Tuasikal, H., & Sorong, U. M. (2025). Pemanfaatan Teknologi Kecerdasan Buatan (Artifical Intelligence) dalam Sistem Peradilan Pidana di Indonesia. *Jurnal Sosial Teknologi*, *5*(6), 1–17.

Tan, D. (2021). Metode penelitian hukum: Mengupas dan mengulas metodologi dalam menyelenggarakan penelitian hukum. *NUSANTARA: Jurnal Ilmu Pengetahuan Sosial*, *8*(8), 2463–2478.

Usman, S. O. (2025). *The Dual-Use Dilemma: Deepfakes as a Tool for Innovation and Misinformation*. ResearchGate.

Utami, N. P. G. S., Sutrisni, N. K., & Suharyanti, N. P. N. (2025). Pengaturan Hukum Tanggung Jawab Pemanfaatan Artificial Intelligence di Indonesia. *Jurnal Hukum Mahasiswa*, *5*(2), 315–325. https://doi.org/10.36733/jhm.v5i2

Wardoyo, H. (2026). Analisis Yuridis Penggunaan Artificial Intelligencedalam Aktivitas Bisnis Digital di Indonesia: Tantangan Regulasi dan Kepastian Hukum. *Jurnal Serambi Hukum*, *19*(2), 133–148.

Wicaksono, D. T., Prasetya, F., & Lestari, N. P. (2026). Penegakan Hukum Terhadap Kejahatan Deepfake dalam Perspektif Hukum Positif Indonesia. *Jurnala Fakta Hukum*, *4*(2), 63–70.

Widjaja, G. (2025). Deepfake dan Masa Depan Kebenaran: Implikasi Etis Dan Sosial. *Berajah Journal: Jurnal Pembelajaran Dan Pengembangan Diri*, *5*(2), 147–156.

Yanuafriza, A. Y., & Eluan, F. V. (2025). Pertanggungjawaban Perdata atas Penyalahgunaan Teknologi Deepfake dalam Perspektif Hukum Indonesia. *Indonesian Journal of Social Sciences and Humanities*, *5*(2), 294–304.

Zulfikar, & Sihombing, P. V. A. E. (2025). Urgensi Model Kerangka Yuridis Adaptif terhadap Deepfake AI (Artificial Intelligence) dalam Perspektif Hukum Siber.

*Jurnal Pendidikan Sejarah Dan Riset Sosial Humaniora (KAGANGA)*, *8*(6), 915–926.

Zulkarnain, R., & Rahmayanti. (2026). Kedudukan Artifical Intelligence dalam Sistem Hukum Indonesia Analisis Normatif Antara Subjek Hukum dan Instrumen Teknologi. *Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi Dan Pendidikan*, *5*(5), 2482–2490.

## Acknowledgments

The author wishes to express sincere gratitude to all parties who provided support and contributed to the success of this activity. Heartfelt thanks are extended to the Institute for Research and Community Service of Universitas Internasional Batam for the institutional support that enabled the activity to run smoothly. Deep appreciation is also conveyed to the supervising lecturers, Agustianto, S.H., MKn., for his guidance, direction, and encouragement throughout the process from beginning to end.

## Competing Interest

The authors state that there is no conflict of interest in the publication of this article.

## Generative AI Statement

During the preparation of this manuscript, the author(s) used a generative artificial intelligence, such as ChatGPT, Claude, and Perplexity as assistant to help translate, paraphrase, and refine the language of the article on the basis of the underlying community service report, and to improve overall readability. The author(s) critically reviewed, edited, and validated the output to ensure accuracy and scientific integrity. The author(s) maintain full accountability for the final content. The tool was used solely as a supportive aid and is not credited with authorship. No original data analysis or interpretation was performed by the AI without human oversight.