



DOI: <https://doi.org/10.38035/jlhr.v1i2>
<https://creativecommons.org/licenses/by/4.0/>

Legal Accountability for Deepfake Abuse Against Women and Children in Indonesia: A Regulatory Challenge in the Era of Artificial Intelligence

Wahyu Kurniadi¹

¹Mahasiswa Fakultas Hukum (Alumni) Universitas Jambi, Indonesia, salisah0908@gmail.com

Corresponding Author: Kurniadiwahyu1511@gmail.com¹

Abstract: *The development of artificial intelligence (AI) has given rise to deepfake technology capable of realistically manipulating images, sounds, and videos. While this technology offers benefits, it also poses new threats in the form of digital identity abuse, often targeting women and children, as vulnerable groups. This study aims to analyze legal liability for deepfake abuse and formulate an adaptive regulatory model in Indonesia. The study uses a normative juridical method with a statutory, conceptual, and comparative regulatory approach. The results indicate that existing legal regulations do not specifically regulate deepfake crimes, creating a legal vacuum in victim protection and law enforcement. This study proposes a regulatory model that integrates the responsibilities of the state, digital platforms, and AI developers by strengthening oversight mechanisms, early detection, content removal, and victim rights restoration. This model is expected to create adaptive legal protection for the development of digital technology in Indonesia.*

Keyword: *Legal Liability, Deepfake, Artificial Intelligence, Women, Children.*

INTRODUCTION

The rapid digital transformation has transformed various aspects of human life, including communication, economics, education, and law enforcement. One innovation that has developed significantly is deepfake technology, a subset of artificial intelligence (AI). This technology enables the realistic manipulation of a person's face, voice, and video, making them difficult to distinguish from the original content (Chesney & Citron, 2019). The development of deepfake technology poses serious threats to digital security, privacy, and human rights. Women and children are the most vulnerable groups, as their identities are often misused to create sexually exploitative content, online harassment, fraud, and the spread of false information (UNESCO, 2023). This phenomenon is growing with increasing access to AI applications that can be freely used by the public. Perpetrators no longer need advanced technical skills to create manipulative videos or images. This situation increases the potential for digital crimes, which can impact the psychological, social, economic, and reputational harm

to victims (Floridi, 2019). Indonesia already has several regulations that can serve as a basis for legal protection, including Law Number 1 of 2024 concerning the Second Amendment to the Electronic Information and Transactions Law, Law Number 27 of 2022 concerning Personal Data Protection, and Law Number 35 of 2014 concerning Child Protection. However, these regulations do not specifically address deepfake crimes and the accountability of parties involved in the creation and distribution of manipulative content.

The absence of comprehensive regulations creates a legal gap that has the potential to hinder victim protection. Furthermore, technological developments move faster than the legal process, necessitating an adaptive regulatory approach based on AI governance. Previous research has generally addressed cybersecurity, AI ethics, or personal data protection separately. There is little research that integrates victim protection, digital platform liability, and AI developer liability within a single, comprehensive legal framework.

Research Objectives

This research aims to:

1. Analyze the forms of legal accountability for the misuse of deepfake technology against women and children in Indonesia.
2. Identifying weaknesses in existing regulations in providing legal protection to victims.
3. Formulating a regulatory model that adapts to developments in artificial intelligence technology.

Supporting Theories

This research utilizes several theories, namely:

1. Philipus M. Hadjon's Legal Protection Theory, which emphasizes preventive and repressive protection of community rights (Hadjon, 1987).
2. Gustav Radbruch's Legal Certainty Theory, which places legal certainty as one of the primary goals of law, alongside justice and utility (Radbruch, 1946).
3. Cyber Law Theory, which explains that the development of digital technology requires regulatory updates to anticipate new forms of crime (Lessig, 2006).
4. AI Governance Theory, which emphasizes the importance of oversight, accountability, transparency, and human rights protection in the use of artificial intelligence (Floridi, 2019).

METHOD

This research uses a normative juridical method, namely legal research conducted by reviewing primary, secondary, and tertiary legal materials to analyze legal issues related to legal accountability for the misuse of deepfake technology against women and children in Indonesia.

The approaches used in this research include:

1. Statute Approach

This approach is carried out by reviewing relevant regulations such as Law Number 35 of 2014 concerning Child Protection, Law Number 27 of 2022 concerning Personal Data Protection, and Law Number 1 of 2024 concerning the Second Amendment to the Electronic Information and Transactions Law (ITE).

2. Conceptual Approach

This approach is used to analyze the concepts of legal protection, criminal liability, and artificial intelligence governance (AI governance).

3. Comparative Approach

This approach is used to compare deepfake regulations in several countries, such as the European Union and the United States, as a basis for comparing legal formation in Indonesia.

Legal sources consist of:

- a. Primary legal materials: laws and regulations and court decisions.
- b. Secondary legal materials: books, scientific journals, and academic articles.
- c. Tertiary legal materials: legal dictionaries and legal encyclopedias.

Data analysis was conducted qualitatively and descriptively, interpreting existing legal materials to address the legal issues under study.

RESULT AND DISCUSSION

Transformation of Digital Crime in the Era of Artificial Intelligence

The development of artificial intelligence (AI) has significantly changed the characteristics of crime in the digital space. While previously dominated by hacking, online fraud, and the distribution of simple illegal content, new, more complex forms of crime have emerged: AI-based manipulation crimes such as deepfakes. Deepfakes are a technology that uses machine learning algorithms, specifically deep learning, to create or manipulate images, videos, and sounds to appear highly realistic. This technology allows someone to "fake reality" with such a high degree of accuracy that it is difficult for ordinary humans to distinguish. In a legal context, this transformation has created a major shift from crime without technology to technology-generated crime, where technology is not only a tool but also part of the structure of the crime itself. This complicates the legal evidence process because AI-generated digital evidence has characteristics that cannot always be verified conventionally. According to Chesney and Citron (2019), deepfakes pose a threat not only to individuals but also to social, political, and legal systems because they can undermine public trust in visual evidence. In the Indonesian context, this is even more serious because the public's digital literacy is still relatively diverse.

Forms of Deepfake Abuse Against Women and Children

The analysis shows that deepfake abuse against women and children can be categorized into several main forms:

- a. Non-consensual pornographic deepfakes

This is the most common form, involving the manipulation of a woman's or child's face into pornographic content without consent. Victims often have no direct relationship with the perpetrator but are instead random targets whose data is harvested from social media. This phenomenon is known as image-based sexual abuse, which in the context of deepfakes is even more dangerous because the victim never actually participated in the creation of the content, yet still experiences the same, if not more severe, social and psychological impacts.

- b. AI-Based Cyber Grooming

Perpetrators use chatbots or AI systems to interact with children in a manipulative manner. AI can be used to slowly build the victim's trust before leading them to digital sexual exploitation.

- c. Digital Identity Forgery

Deepfakes are also used to create fake identities that resemble the victim, which are then used for fraud, blackmail, or the dissemination of false information.

- d. Cyber Harassment and Blackmail

Victims are often threatened with the distribution of deepfake content if they do not comply with the perpetrator's demands, including requests for money or certain actions.

These four forms demonstrate that deepfakes are not simply a technological issue but have become a highly complex new tool of exploitation.

Weaknesses of the Indonesian Legal System in Responding to Deepfakes

1. Absence of Explicit Norms

One of the main findings of this study is the lack of explicit legal norms governing deepfakes in the Indonesian legal system. Neither the ITE Law nor the Personal Data Protection Law specifically mentions the terms deepfake or AI-generated content. As a result, law enforcement officials often use general articles such as defamation or the distribution of obscene content, which do not fully reflect the complexity of deepfake crimes.

2. Limitations of Digital Evidence

In criminal procedure law, proof relies heavily on the authenticity of the evidence. However, in deepfake cases, digital evidence can be manipulated in such a way that it is difficult to distinguish between genuine content and AI manipulation. This creates a serious challenge to the principle of due process of law, as there is a risk of errors in determining material truth.

3. Unclear Subjects of Liability

There are three main actors in the deepfake ecosystem:

- a. Content creators
- b. Content distributors
- c. Digital platforms

However, Indonesian law does not clearly divide the responsibilities of these three. As a result, legal uncertainty arises in law enforcement, especially when content is distributed through global platforms outside of Indonesian jurisdiction.

4. Slow Regulatory Response to Technology

This phenomenon indicates a phenomenon of law lagging behind technology, a condition where technological developments outpace legal development. According to Floridi (2019), this condition can create a legal vacuum that can be exploited by criminals.

Analysis of Legal Liability from an Indonesian Perspective

1. Criminal Liability of Individual Perpetrators

In the Indonesian legal system, perpetrators of deepfake acts can be subject to several legal provisions, such as:

- a. The Electronic Information and Transactions Law (ITE Law) regarding the distribution of illegal electronic content
- b. The Criminal Code (KUHP) regarding defamation and slander
- c. The Pornography Law regarding the distribution of immoral content

However, this approach is still reactive, not preventative. This means that the law only comes into play after the crime has occurred, rather than preventing it from occurring.

2. Liability of Digital Platforms

Digital platforms play a crucial role in content distribution. Under the theory of intermediary liability, platforms can be held liable if they fail to remove illegal content within a reasonable time. However, in Indonesia, this mechanism has not been regulated in detail, so there is still reliance on the platform's internal policies (community guidelines), rather than national law.

3. Liability of AI Developers

In the AI governance approach, technology developers also have an ethical and legal responsibility to ensure the system is not misused. This concept is known as "responsible AI development." However, in practice, developers are often beyond the reach of national laws due to the transnational nature of the technology.

Social, Psychological, and Legal Impacts on Victims

Research shows that the impact of deepfakes on women and children is extensive, including:

a. Psychological Impacts

Victims experience trauma, anxiety, depression, and a loss of sense of security due to the digital misuse of their identities.

b. Social Impacts

Victims often experience social stigma, isolation, and even loss of social relationships due to the spread of fake content.

c. Legal Impacts

Victims often struggle to obtain justice due to limited evidence and unclear regulations.

According to UNICEF (2023), victims of child digital exploitation often do not report their cases due to fear of stigma and lack of trust in the legal system.

Comparison of International Regulations

1. European Union

The European Union, through the EU AI Act, has regulated the use of AI based on risks, including a prohibition on the use of AI systems that pose a high risk to human rights.

2. United States

Several states, such as California, have already enacted bans on deepfakes in the context of elections and non-consensual pornography.

3. South Korea

South Korea has even established strict criminal sanctions against the distribution of pornographic deepfakes.

This comparison shows that Indonesia is still at an early stage in regulating AI, necessitating accelerated and adaptive regulation.

AI Governance Model as a Legal Solution

This research proposes a legal protection-based AI Governance model consisting of four pillars:

a. Regulatory framework

Establishment of specific deepfake regulations in national law.

b. Technological control

Obligation for digital platforms to use AI detection systems.

c. Institutional collaboration

Collaboration between the government, digital platforms, and child protection agencies.

d. Victim-centered approach

Focus on victim recovery, not just punishment of perpetrators.

This model aligns with the principle put forward by Floridi (2019) that AI must be developed ethically, transparently, and responsibly.

Regulatory Gap Analysis (Legal Gap)

The research results indicate three main gaps:

1. Normative gap: there are no specific legal norms for deepfakes.

2. Institutional gap: there is no specific agency for AI oversight.

3. Enforcement gap: law enforcement still uses outdated approaches.

These gaps demonstrate the need for comprehensive digital law reform. 9. Implications for the Indonesian Legal System

The implications of this research are the need for:

- a. Technology-based criminal law reform
- b. Strengthening the ITE Law based on AI
- c. Establishing national AI regulations
- d. Improving public digital literacy

Without these reforms, the law will continue to lag behind technological developments.

CONCLUSION

This study concludes that the misuse of deepfake technology in the context of artificial intelligence poses a serious and emerging legal challenge in Indonesia, particularly in relation to the protection of women and children as vulnerable groups. The findings indicate that current Indonesian legal frameworks, including the Electronic Information and Transactions Law, the Child Protection Law, and the Personal Data Protection Law, are still insufficient to comprehensively regulate deepfake-related offenses. The absence of explicit legal norms governing deepfake creates a significant legal vacuum, especially regarding the classification of offenses, identification of legal responsibility, and enforcement mechanisms. As a result, law enforcement agencies tend to rely on general provisions that do not fully capture the complexity and technological nature of AI-based crimes. Furthermore, this study finds that legal responsibility in deepfake cases is fragmented among individuals, digital platforms, and AI developers. However, Indonesian law has not yet clearly defined the extent of liability for each actor, resulting in uncertainty in legal enforcement and victim protection. Therefore, this research proposes the development of an AI governance-based legal framework that integrates preventive, repressive, and restorative approaches. Such a framework should include stricter platform accountability, rapid content removal mechanisms, and stronger victim-centered protection policies. In conclusion, legal reform is urgently required to ensure that Indonesian law remains adaptive to technological advancements and capable of providing effective protection against AI-driven digital crimes.

REFERENCES

- Arendt, H. (1958). *The Human Condition*. University of Chicago Press.
- Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *California Law Review*, 104(3), 671 – 732.
- Chesney, R., & Citron, D. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 107(6), 1753 – 1820.
- Maisharoh, T., & Ali, H. (2020). Faktor-faktor yang mempengaruhi Infrastruktur Teknologi Informasi: Keuangan, Fleksibilitas TI, dan Kinerja Organisasi. *Universitas Mercu Bauna (UMB) Jakarta*, 1-9.
- Cohen, J. E. (2019). *Between Truth and Power: The Legal Constructions of Informational Capitalism*. Oxford University Press.
- Crawford, K. (2021). *Atlas of AI*. Yale University Press.
- Floridi, L. (2019). Establishing the rules for building trustworthy AI. *Nature Machine Intelligence*, 1(6), 261 – 262.
- Djojo, A., & Ali, H. (2012). Information technology service performance and client's relationship to increase banking image and its influence on deposits customer banks loyalty (A survey of Banking in Jambi). *Archives Des Sciences*, 65(8), 65-76.
- Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems*, 14(3), 330 – 347.
- Gasser, U., & Almeida, V. A. F. (2017). A layered model for AI governance. *IEEE Internet Computing*, 21(6), 58 – 62.

- Hadjon, P. M. (1987). *Perlindungan Hukum bagi Rakyat Indonesia*. Surabaya: Bina Ilmu.
- Lessig, L. (2006). *Code: Version 2.0*. Basic Books.
- Zahran, R., & Ali, H. (2020). Faktor-Faktor Yang Mempengaruhi Sistem Informasi: Sumber Daya Manusia, Bisnis, Teknologi dan Metode. *Jurnal Akuntansi Universitas Mercubuana*, 1, 21.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms. *Big Data & Society*, 3(2).
- NIST. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. National Institute of Standards and Technology.
- OECD. (2019). *OECD Principles on Artificial Intelligence*. OECD Publishing.
- Mansur, S., & Ali, H. (2017). Impact of customer engagement to reputation of the Bri Syariah Indonesia. *International Journal of Economic Research*, 14(19), 425-431.
- Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.
- Solove, D. J. (2007). *Understanding Privacy*. Harvard University Press.
- UNESCO. (2023). *Guidance for Generative AI in Education and Research*. United Nations Educational, Scientific and Cultural Organization.
- UNICEF. (2023). *Child Online Protection Report*. United Nations Children' s Fund.
- United Nations. (2021). *General Comment No. 25 on children' s rights in relation to the digital environment*.
- Masruri, M., Ali, H., & Rosadi, K. I. (2021). Pengelolaan keuangan dalam mempertahankan kualitas pondok pesantren selama pandemi COVID-19. *Jurnal Ilmu Manajemen Terapan*, 2(5), 644-657.
- Undang-Undang Republik Indonesia Nomor 35 Tahun 2014 tentang Perlindungan Anak.
- Undang-Undang Republik Indonesia Nomor 27 Tahun 2022 tentang Perlindungan Data Pribadi.
- Undang-Undang Republik Indonesia Nomor 1 Tahun 2024 tentang Perubahan Kedua atas UU ITE.
- Whittaker, M. (2021). *The AI Now Report*. AI Now Institute, New York University.
- Zuboff, S. (2019). *The Age of Surveillance Capitalism*. PublicAffairs.
- European Union. (2024). *Artificial Intelligence Act (EU AI Act)*.
- California Legislative Information. (2022). *AB-602: Deepfake Regulation Law*.
- Citron, D. K. (2014). *Hate Crimes in Cyberspace*. Harvard University Press.