

Criminal Liability for Deepfake Content that Damages a Person's Reputation: A Perspective Under Indonesian Criminal Law

Saupi Hasbi

Universitas Pembangunan Nasional Veteran Jakarta, Indonesia

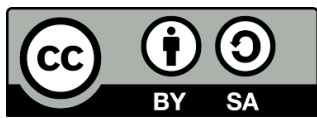
Email: saupihhasbi.law@gmail.com

Keywords:

artificial intelligence; deepfake;
defamation; criminal liability.

Abstract

The development of Artificial Intelligence (AI) has given rise to deepfake technology, which can generate or manipulate images, audio, and videos to make them appear authentic. The misuse of this technology may be employed to attack a person's honor and reputation. This study aims to analyze the criminal law regulation of deepfake content that harms a person's reputation, as well as the criminal liability of its creators and disseminators. This research employs a normative legal method using statutory, conceptual, and case approaches. The results of the study indicate that the misuse of deepfake technology may be subject to criminal charges under the provisions of the National Criminal Code and Article 27A of the Electronic Information and Transactions Law (ITE Law) if the elements of a criminal offense are fulfilled. Criminal liability is imposed on parties who commit the unlawful act and possess criminal intent or negligence, whereas AI developers cannot be held automatically liable without evidence of direct involvement and fault. Therefore, strengthening criminal law policies and enhancing digital forensic capabilities are necessary to effectively address the misuse of deepfake technology.



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

Copyright © 2026 by Author. Published by
International Journal Labs

INTRODUCTION

The development of information and communication technology has brought fundamental changes to patterns of human interaction, including the processes of creating, disseminating, and receiving information. One of the most significant technological developments is the presence of artificial intelligence (AI). Artificial intelligence technology is basically developed to assist humans in performing various tasks more quickly, effectively, and efficiently. However, these developments also have legal consequences because technology can be used as a means to commit acts that harm the legal interests of others. One form of misuse of artificial intelligence that is increasingly developing is deepfake technology.

Deepfake technology is an artificial intelligence-based technology capable of producing or manipulating images, sounds, videos, and other forms of digital content so that they appear realistic and depict events that never actually occurred. Through this technology, a person's face can be superimposed onto another person's body, a person's voice can be digitally imitated, or a video can be created to make it appear as though someone is doing or saying something that they have never actually done or said (Hameleers et al., 2024).

The ability of deepfakes to produce synthetic content with a high level of realism poses serious challenges from a criminal law perspective. If the technology is used without authorization to create content that depicts someone committing a disgraceful act, making a particular statement, or engaging in an action that never actually occurred, such content has the potential to damage the honor and reputation of the person who is the subject of manipulation.

The harm caused can become even greater when such content is distributed through social media and digital platforms because its dissemination can occur rapidly, extensively, and be difficult to control (Avalle et al., 2024; Jafar et al., 2023; Jaff, 2023; Park & Rohatgi, 2024; Surjatmodjo et al., 2024).

In criminal law, honor and reputation are legal interests that receive protection. Moeljatno (2008) explained that criminal law is part of the overall legal system that determines acts that should not be committed, establishes the conditions under which a person may be subject to criminal sanctions, and determines the penalties that may be imposed on offenders. Therefore, technological development does not automatically eliminate the applicability of criminal law principles. If technology is used as a means to commit acts that fulfill the elements of a criminal offense, the focus of criminal law remains directed toward human actions, the perpetrator's fault, and the consequences arising from such actions.

The problem arises because the characteristics of deepfakes differ from conventional forms of defamation. In conventional defamation, perpetrators directly convey accusations or attacks against a person's honor through certain words, writings, or media. In deepfake content, attacks on reputation can be carried out through the construction of a false reality. The perpetrator does not always need to explicitly write an accusation but can create images, videos, or synthetic audio that cause the public to believe that the victim has actually committed a particular act.

Thus, the legal issue that arises is whether deepfake content can be qualified as a form of attack against honor or reputation. Sarah Amanda Uly Sijabat and Diana Lukitasari, in their research on deepfake pornographic image and video content, argue that deepfake content can be categorized as a form of defamation if it fulfills the elements of defamation, particularly when the content presents an image that degrades a person's honor and does not clearly indicate its manipulative nature (Du et al., 2023; Noggle, 2025; Wang et al., 2025).

In the context of Indonesia's positive law, one of the important provisions is Article 27A of Law Number 1 of 2024 concerning the Second Amendment to Law Number 11 of 2008 concerning Electronic Information and Transactions. This provision regulates the prohibition against any person attacking the honor or reputation of another person by making accusations with the intention of making them known publicly through an Electronic System. This provision is relevant to the dissemination of deepfake content because the technology can be used through electronic systems to create and distribute false representations of an individual (Afanasyeva & Yumasheva, 2022; Alanazi et al., 2024; Doss et al., 2023; Gramigna, 2024; Ivanova, 2025; Perriello, 2026).

In addition to the ITE Law, regulations concerning defamation are also contained in the national Criminal Code, namely Law Number 1 of 2023 concerning the Criminal Code. The Criminal Code provides a legal basis for acts that attack a person's honor or reputation through certain accusations. The existence of both the Criminal Code and the ITE Law demonstrates that the protection of honor is not only relevant in direct social interactions but must also be applied to forms of attacks conducted through digital technology.

However, the application of these provisions to deepfakes is not always straightforward. The first issue is determining the form of conduct that can be categorized as a criminal offense. In a particular case, the party who creates the content, the party who orders or requests the creation of the content, and the party who distributes the content may be different individuals.

Therefore, an analysis of the role of each party based on the principle of criminal liability is required.

The second issue concerns the proof of fault. One of the fundamental principles of criminal law is the principle of *geen straf zonder schuld*, which is generally interpreted as “no punishment without fault.” A person cannot be convicted merely because a harmful consequence occurs; rather, it must be proven that there is a relationship between the act committed and the element of fault required under criminal law. In cases involving deepfakes, proof of intent is essential, particularly to distinguish between the legitimate use of technology and the use of technology that was intended from the outset to attack another person's honor or reputation.

Criminal liability relates to whether a person can be held responsible for a criminal offense that has been committed. Such responsibility fundamentally requires the existence of a criminal act and fault on the part of the perpetrator (Hasbi et al., 2023; Pasat, 2022; Titahelu, 2022; Widyaningrum et al., 2024). Based on this approach, deepfake technology cannot be positioned as an entity that automatically possesses criminal fault. Within the framework of Indonesian criminal law, accountability must be traced to the legal subject who commits, orders the commission of, participates in, or contributes through a specific role to the occurrence of a criminal offense.

Research on deepfakes also demonstrates that Indonesia's legal system continues to face challenges related to regulatory fragmentation. Herson Djuanda, Erikson Sihotang, and Komang Edy Dharma Saputra, for example, examine criminal liability for the misuse of deepfake videos in defamation cases through the relationship between Article 27A of the ITE Law, provisions concerning the manipulation of electronic information or electronic documents, and other relevant regulations. This indicates that the resolution of deepfake cases cannot always rely solely on a single criminal provision.

The position of deepfakes as a technological development requires a re-analysis of the construction of criminal liability and criminal law policy. This technology is capable of creating highly realistic audio-visual representations, while the criminal law system was fundamentally developed to address forms of conduct that were previously more conventional. Based on this description, it is necessary to conduct an in-depth analysis of how Indonesian criminal law can provide a basis for accountability for creators and parties who distribute deepfake content that harms a person's reputation. This analysis is important to ensure the protection of a person's honor and reputation without disregarding the principles of legality, fault, and legal certainty.

Therefore, this study adopts the title “Criminal Liability for Deepfake Content That Damages A Person’s Reputation: A Perspective Under Indonesian Criminal Law.”

Problem Formulation

1. How is the criminal law regulation of *deepfake* content that harms a person's good name in Indonesia's positive law?
2. What is the criminal liability for those who create and/or disseminate *deepfake* content that harms someone's reputation?

Research Objectives

1. To analyze the criminal law regulation of *deepfake* content that harms a person's good name based on Indonesia's positive law.

2. To analyze the form of criminal liability against the party who creates and/or spreads *deepfake* content that harms a person's good name.

Research Uses

1. Theoretical Uses

This research is expected to contribute to the development of legal science, especially criminal law and cyber law, regarding the construction of criminal liability for the misuse of artificial intelligence technology. This research is also expected to enrich academic studies on the application of the principles of error, criminal responsibility, and the protection of one's honor in the face of the development of digital technology.

2. Practical Uses

Practically, this research is expected to serve as a consideration for law enforcement officials in identifying the elements of criminal offenses and determining criminal liability in cases of deepfake misuse. In addition, this research is expected to provide input for lawmakers in formulating criminal law policies that are more adaptive to developments in artificial intelligence technology.

Criminal liability is a concept that determines whether a person can be held accountable and convicted for an act that they have committed. In criminal law doctrine, criminal acts and criminal liability are two interrelated but distinguishable concepts. Criminal acts relate to conduct that fulfills the elements of a prohibition stipulated by law, while criminal liability concerns whether the perpetrator can be personally held responsible for such conduct (Hameleers et al., 2024).

Criminal liability is a continuation of the existence of a criminal act. However, not every person who commits a criminal act can automatically be convicted because a legal basis is still required to attribute responsibility for the act to the perpetrator. Therefore, in determining criminal liability for deepfake perpetrators, it must first be proven that an act fulfills the elements of a criminal offense, followed by an analysis of the perpetrator's fault.

Basically, fault in criminal law is related to intent and negligence in accordance with the formulation of the criminal offense. In cases involving deepfake content that attacks a person's reputation, intent can be assessed from the perpetrator's knowledge and willingness when creating or distributing the content. If a person knows that the content is the result of manipulation, knows that the subject of the content is another person, and intends to distribute the content with the purpose of degrading the victim's honor, then the element of fault has a strong basis for proof.

Criminal offenses are acts prohibited by legal provisions and accompanied by criminal sanctions for anyone who violates such prohibitions. This definition emphasizes that determining an act as a criminal offense must be based on a legal provision. In relation to deepfakes, the technology itself does not constitute a criminal offense. The object of criminal law assessment is the human use of the technology. Thus, the creation of deepfake technology for research, education, filmmaking, or other lawful purposes cannot automatically be classified as a criminal act. An act becomes relevant under criminal law only when the technology is used to fulfill the elements of conduct prohibited by laws and regulations.

Defamation is an act that attacks a person's honor or reputation. Protection of honor is fundamentally intended to maintain a person's status and standing in public life. In the digital context, attacks on reputation may have broader consequences because information can be

replicated and disseminated rapidly. Deepfake content has a distinctive characteristic because attacks on reputation can be constructed through visual, audio, or audiovisual manipulation. For example, a person can be portrayed as saying something that they have never said or as committing an act that they have never performed. If such content is perceived by the public as factual and causes degradation of the victim's honor or reputation, the act may be examined based on criminal law provisions protecting honor and reputation.

Artificial intelligence (AI) is a technology that enables computer systems to perform processes related to information processing, pattern recognition, learning, and generating outputs based on data and instructions. One of the developments in AI involves the use of machine learning and deep learning to produce synthetic content. Deepfakes can generally be understood as digital content that is manipulated or generated using artificial intelligence technology to create realistic representations. Such content may take the form of images, audio, videos, or combinations of various media formats. The primary concern regarding deepfakes is their ability to create an illusion of authenticity, making it difficult for individuals to distinguish between genuine and manipulated content.

From a criminal law perspective, this manipulative characteristic becomes legally relevant when it is used to create a false representation concerning an individual. Indonesian legal studies on deepfakes indicate that this technology has the potential to be misused for defamation, non-consensual pornography, fraud, identity falsification, and other forms of cybercrime.

Legal certainty is one of the fundamental objectives of law enforcement. In criminal law, legal certainty is closely related to the principle of legality. A person may only be convicted for an act that has been regulated as a criminal offense under applicable law. In deepfake cases, the principle of legal certainty is necessary to prevent the arbitrary expansion of criminal liability. Law enforcement officials cannot impose criminal sanctions solely because a person uses emerging technology. Instead, they must clearly establish the applicable legal provisions that have been violated, the fulfillment of the elements of the criminal offense, the existence of fault, and the relationship between the perpetrator and the harmful content distributed to the victim.

METHOD

Types of Research

This research employed normative legal research or doctrinal legal research. The analysis focused on relevant legal norms, principles, theories, and statutory provisions related to criminal liability for the creation and dissemination of deepfake content that harms a person's reputation.

This research used a prescriptive approach by analyzing how legal provisions should be applied to address criminal liability arising from the creation and dissemination of deepfake content that damages a person's reputation.

Research Approach

This research uses several approaches, namely:

1. Statute Approach

This approach is carried out by reviewing laws and regulations related to defamation, electronic information, criminal liability, and the use of a person's data or identity. The

regulations analyzed include the National Criminal Code, the ITE Law, the Personal Data Protection Law, and other relevant regulations.

2. Conceptual Approach

A conceptual approach is used to analyze the concepts of criminal liability, criminal acts, mistakes, intentionality, honor, good name, artificial intelligence, and *deepfakes*.

3. Case Approach

The case approach is used in a limited way to understand the character of legal problems that can arise in the practice of manipulative digital content abuse and its relevance to the application of criminal norms.

Legal Ingredients

The legal materials in this study consist of:

1. Primary legal materials, in the form of laws and regulations, especially Law Number 1 of 2023 concerning the Criminal Code, Law Number 1 of 2024 concerning the Second Amendment to Law No. 1 of 2024, Law Number 27 of 2022 concerning Personal Data Protection, and other related regulations.
2. Secondary legal materials are in the form of criminal law books, expert opinions, journal articles, research results, and scientific papers that discuss defamation, cybercrime, AI, and *deepfakes*.
3. Tertiary legal materials, in the form of legal dictionaries, encyclopedias, and other sources that help provide explanations of the terms used in research.

Legal Material Collection Techniques

The collection of legal materials is carried out through *library research*. Legal materials are collected by studying laws and regulations, books, journals, and scientific documents relevant to the object of research.

Legal Materials Analysis Techniques

Legal materials are analyzed qualitatively using the deductive reasoning method. The analysis starts from legal norms and general concepts regarding criminal acts and criminal liability, then is applied to answer specific problems regarding *deepfake* content that harms a person's good name.

Criminal Law Regulations on Deepfake Content That Harms a Person's Good Name Deepfakes as a Form of Technology-Based Crime Development

The development of technology has caused the form of criminal acts to change. Crimes are no longer only committed through physical acts or direct communication, but also through electronic systems. In this context, *deepfakes* are one of the technologies that has the potential to expand the way perpetrators carry out attacks on someone's legal interests. In principle, criminal law does not punish technology. Criminal law regulates and assesses human actions that use technology as a means to commit criminal acts. Therefore, the use of *deepfakes* must be seen based on the purpose, form of act, consequences, and mistakes of the perpetrator.

If someone uses deepfake technology to create videos that depict the victim as if they are committing a crime, committing an immoral act, accepting bribes, having an affair, or making certain statements that never actually happened, then the content can have legal consequences if it is used to attack the victim's honor or good name. Research on deepfake pornographic images and videos shows that this form of content can meet the elements of defamation if it creates false representations that attack the victim's honor. However, a concrete

assessment must be made based on the content itself, the context of the spread, the perception caused, and the presence or absence of an explanation of the manipulative nature of the content (Sijabat & Lukitasari, 2024).

Application of Article 27A of Law No. 1 of 2024

Article 27A of Law No. 1 of 2024 is one of the main provisions in analyzing defamation carried out through electronic systems. This provision basically prohibits everyone from deliberately attacking the honor or good name of others by accusing something with the intention of making it known to the public through electronic systems.

There are several important elements. First, there is a subject of "everyone". Second, there is an element of intentionality. Third, there is an act of attacking honor or good name. Fourth, it is done by accusing something. Fifth, there is an intention for the accusation to be known to the public. Sixth, the act is carried out through an electronic system.

In the context of *deepfakes*, the element of "alleging something" does not always have to be understood narrowly as an accusation that is stated in writing. Video or audio content that depicts someone as if they are committing a certain act can have the same communicative function as the submission of accusations. For example, if the perpetrator makes a *deepfake* video that shows the victim as if he received a bribe and then distributes it to make the public believe that the event is true, the substance of the communication can be seen as conveying an accusation about the victim's actions.

Nevertheless, the application of Article 27A must still be carried out carefully. The public prosecutor must prove that the content is really intended or is actually used to attack the victim's honor. Not every content that is automatically manipulated constitutes defamation.

Regulations in Law No. 1 of 2023 (National Criminal Code)

The National Criminal Code also provides a legal basis for the crime of assault on honor or good name. In the context of this study, the provisions regarding defamation become relevant if the elements of acts committed through *deepfakes* meet the formulation of criminal acts as stipulated in the Criminal Code. The use of new technology does not mean that the acts committed are outside the criminal law. If the core of the act is to attack someone's honor through accusations or the submission of false information that degrades the victim, then *deepfake* technology can be positioned as a mode or means used to realize criminal acts.

Thus, the digital character of *deepfakes* does not eliminate legal protection of honor. Precisely the characteristics of these technologies can be an important factor in assessing how actions are carried out and the consequences they cause.

Provisions Regarding the Manipulation of Information or Electronic Documents

In addition to the provisions regarding defamation, a *deepfake* act in certain circumstances can be related to other criminal provisions, especially if the content is created through the manipulation of electronic data and used as if it were authentic information. Therefore, law enforcement does not always have to use one provision separately. Law enforcement officials can analyze the entire set of acts. For example, a person can manipulate content, use the identity of another person, create misleading electronic information, and disseminate it to bring down the victim's good name. Each act must be tested based on the elements of the criminal act in question. It shows the importance of looking at the relationship between defamation provisions, electronic information manipulation, and other relevant provisions when dealing with *deepfake* abuse (Djuanda et al., 2026).

Personal Data Protection Act

In *deepfake* practices, faces, sounds, images, and data that can identify a person can be used as materials to create synthetic content. Therefore, in certain cases, *deepfake* abuse not only has the potential to attack reputation, but can also be related to violations of personal data protection.

However, accountability under the Personal Data Protection Law must still be tested based on the elements of a specially regulated criminal act. Not every use of a person's photo or voice automatically meets all the elements of a criminal act in the PDP Law. Law enforcement must distinguish between violations of honor, violations of personal data, and other criminal acts that may occur simultaneously.

Emptiness and Fragmentation of Settings

One of the main problems is that there is no specific criminal act that explicitly formulates "*deepfake* abuse" as a separate offense. As a result, the handling of cases must use various existing legal provisions based on the character of the act and its consequences.

The research also identified the problem of norm fragmentation, the proof of the *actus reus* element for synthetic content, the proof of *mens rea*, and the issue of accountability of those who provide or develop technology (Gunawan, 2026). This condition does not mean that all *deepfake* acts are in a legal vacuum. On the contrary, certain acts can still be charged based on applicable provisions if the elements are met. However, there is a need to clarify criminal law policies so that law enforcement provides more certainty.

Criminal Liability for Deepfake Content Creators and Spreaders

Criminal Liability of Content Creators

The party who creates *deepfake* content can basically be held criminally liable if it is proven that they have committed an act that meets the elements of a criminal act and has a mistake.

For example, someone deliberately inserts the victim's photos and voices into a *deepfake* application, makes a video depicting the victim committing a humiliating act, and then uploads or hands over the content to another party with the aim of humiliating the victim. In such a situation, there is a direct relationship between the perpetrator's actions and the content used to attack the victim's honor. The element of intentionality can be analyzed from several circumstances, including:

- a. the perpetrator knows that the content is the result of manipulation;
- b. the perpetrator knows the identity of the person who is the object of manipulation;
- c. the perpetrator knows that the description in the content does not correspond to reality;
- d. the perpetrator wants or is at least aware of the possibility that the content will harm the victim's honor; and
- e. there is an act of dissemination or purpose for the content to be known by other parties.

If these circumstances can be proven, then the principle of criminal liability can be applied to content creators.

Criminal Liability of Content Spreaders

The party who spreads *deepfake* content can also be held criminally liable if it meets the elements of a criminal act. However, the responsibility must be analyzed individually. A person who creates content for the first time certainly has a different position than someone who only receives and forwards it. However, the recipient of the content who knows that the

content is fake and deliberately spreads it to bring down the victim's honor can have their own mistakes. Thus, redistribution cannot always be seen as a neutral act. If a person knows the manipulative nature of a content, understands that the content is detrimental to the victim, and then still deliberately distributes it for the public to know, then his actions can be analyzed as a form of involvement in a criminal act in accordance with applicable regulations.

Distinction between Maker, Enforcer, and Parties Involved

In the practice of using *deepfakes*, there is a possibility of division of roles. One person can give instructions to others to create content. The second person has the technical ability to create *deepfakes*. The third person is in charge of disseminating content through various social media accounts. This situation requires law enforcement to analyze the form of participation of each party. It is not appropriate for all parties to be automatically given the same qualification of accountability. It must be proven that each party's real contribution, intentionality, and relationship to the realization of a criminal act must be proven. This approach is in accordance with the principle of individualization of criminal liability. Everyone must be held accountable based on their own actions and mistakes.

The Position of AI Users as Responsible Subjects

AI is currently more accurately understood as a means or instrument used by humans. When a person enters instructions, data, images, or sounds with the aim of producing content that attacks someone's honor, the analysis of accountability is directed to the human who performs the act. The study of Muhammad Syafiq Wafi, Aloysius Wisnubroto, and Yudi Prayudi on AI-based deepfake crimes shows that there is a debate about the principle of error and the possibility of developing the concept of legal subjects in dealing with AI (Wafi et al., n.d.). However, based on Indonesia's positive criminal law framework, criminalization of perpetrators still requires a legal subject that can meet the elements of criminal responsibility.

Therefore, in a *deepfake* case, AI cannot simply replace the position of humans as the party who has the fault. Investigations must be directed to parties who operate, instruct, or use technology in a way that meets the elements of a criminal act.

Limits of Liability of Technology Developers

A more complex issue is the liability of the developer or *developer*. Not every developer of *deepfake* technology can be held accountable for all criminal acts committed by users. Criminal prosecution should not be carried out just because a technology has the potential to be abused. The developer's criminal liability must be based on proof of involvement and wrongdoing. If the developer only creates technology that has a legitimate use and is not aware of a specific criminal use, then the principle of error must be a barrier to criminalization. On the other hand, if in a situation it can be proven that a certain party intentionally designed, provided, or assisted the use of the system with the specific purpose of committing a criminal act, then there is a basis for analyzing the form of liability based on actual involvement and wrongdoing.

Thus, the accountability model for developers cannot be applied automatically. Research on deepfakes also shows that there are normative issues regarding the limits of responsibility of developers and technology providers that still require the development of criminal law policies (Satrul & Hasanah, 2026).

Proving Deepfake Content in the Criminal Justice Process

Proof is one of the important issues. In the case of defamation using *deepfakes*, it must be proven that the content circulating is indeed the result of manipulation or synthesis and not the original recording.

Proof can involve a digital forensic examination of:

1. original files of content;
2. metadata;
3. creation and upload history;
4. electronic devices;
5. digital accounts used;
6. the traces of communication between the parties;
7. patterns of manipulation of images or sounds; and
8. Testimony of digital forensic experts.

This proof is important to connect the perpetrator to the deed. It is not enough to just show that the victim was harmed by a video. Law enforcement must also try to prove who created, uploaded, or intentionally disseminated the content.

Research on the phenomenon of *deepfakes* in social media crimes also places the issue of electronic evidence and proof as one of the central issues in determining criminal liability (NurJannah et al., 2026).

Criminal Liability Analysis Based on Actus Reus and Mens Rea

To provide a clearer construction, the liability of the creator or spreader of deepfakes can be analyzed through two main aspects.

First, the actus reus or the element of the act. This element can be in the form of the act of creating, manipulating, uploading, transmitting, distributing, or making content available to the public, as long as the act meets the applicable criminal formula.

Second, mens rea or the element of error. This element is related to the knowledge and will of the perpetrator. It must be proven whether the perpetrator knows the false or manipulative nature of the content and has intentions or intentions that are relevant to the element of delicacy.

If a person without knowing the false nature of the content only forwards information that he believes to be true, then the analysis of the error is certainly different from the person who created fake content from the beginning. Therefore, proving the knowledge of the perpetrator is very important.

Legal Protection of Victims

Victims of deepfake abuse can suffer multiple layers of losses. In addition to good name, victims can experience privacy violations, loss of public trust, professional losses, and ongoing social impacts because content can be copied and redistributed.

Legal protection is not enough to be directed only at criminalization. Content handling also needs to pay attention to efforts to stop the spread, secure digital evidence, and restore the interests of victims in accordance with the available legal mechanisms.

Therefore, the law enforcement system must be able to respond to the special character of digital crimes. Speed is an important factor because the longer the content is in the digital space, the wider the possibility of losses experienced by the victim.

CONCLUSION

The regulation of criminal law concerning deepfake content that harms a person's reputation under Indonesia's positive law can generally be addressed through the application of existing criminal provisions, particularly provisions concerning attacks on honor or reputation in the Criminal Code and Article 27A of the ITE Law. Deepfakes do not constitute criminal offenses based solely on their technological characteristics; however, their use may constitute a criminal offense when they are employed to create or distribute false representations that fulfill the elements of an attack on a person's honor or reputation. Under certain circumstances, such acts may also be related to provisions concerning the manipulation of electronic information and personal data protection. The main challenge is the absence of specific regulations governing deepfakes, resulting in law enforcement relying on the application of several fragmented legal provisions.

Criminal liability for deepfake content is imposed on legal subjects who commit prohibited acts and possess the required element of fault. Content creators may be held liable if they intentionally use technology to produce manipulated content that attacks a victim's honor or reputation. Parties who distribute such content may also be held liable if they are aware of its manipulative or harmful nature and intentionally disseminate it. When multiple parties are involved, liability must be determined based on their respective roles, contributions, and forms of fault. Artificial intelligence (AI) as a technology does not automatically constitute a responsible legal subject. Similarly, technology developers cannot be held criminally liable merely because the technology they develop is misused, unless their involvement and fault fulfill the requirements for criminal liability.

REFERENCES

- Afanasyeva, T., & Yumasheva, I. (2022). Research on the effects of "DeepFake" technology for the modern digital space. In *Challenges and solutions in the digital economy and finance: Proceedings of the 5th International Scientific Conference on Digital Economy and Finances (DEFIN 2022), St. Petersburg 2022* (pp. 57–65).
- Alanazi, S., Asif, S., & Moulitsas, I. (2024). *Examining the societal impact and legislative requirements of deepfake technology: A comprehensive study*.
- Avalle, M., Di Marco, N., Etta, G., Sangiorgio, E., Alipour, S., Bonetti, A., Alvisi, L., Scala, A., Baronchelli, A., & Cinelli, M. (2024). Persistent interaction patterns across social media platforms and over time. *Nature*, 628(8008), 582–589.
- Doss, C., Mondschein, J., Shu, D., Wolfson, T., Kopecky, D., Fitton-Kane, V. A., Bush, L., & Tucker, C. (2023). Deepfakes and scientific knowledge dissemination. *Scientific Reports*, 13(1), 13429.
- Djuanda, H., Sihotang, E., & Saputra, K. E. D. (2026). Criminal liability for the misuse of deepfake videos in the crime of defamation under the Electronic Information and Transactions Law. *Journal of Law, Social Science and Humanities*.
- Du, X., Peng, J., Zhou, Y., Zhang, J., Chen, S., Jiang, G., Sun, X., & Ji, R. (2023). Pixelface+: Towards controllable face generation and manipulation with text descriptions and segmentation masks. In *Proceedings of the 31st ACM International Conference on Multimedia* (pp. 4666–4677).
- Gramigna, R. (2024). Preserving anonymity: Deep-fake as an identity-protection device and as a digital camouflage. *International Journal for the Semiotics of Law–Revue Internationale de Sémiotique Juridique*, 37(3), 729–751.
- Gunawan. (2026). Normative limitations on criminal liability for artificial intelligence-based

- deepfake abuse in Indonesia. *Journal of Law, Humanities and Politics*.
- Hameleers, M., van der Meer, T. G. L. A., & Dobber, T. (2024). They would never say anything like this! Reasons to doubt political deepfakes. *European Journal of Communication*, 39(1), 56–70.
- Hasbi, F. R., Utari, A. A. D., & Aringga, R. D. (2023). Criminal liability for perpetrators of negligent crimes resulting in the death of others in traffic accidents. *Sinergi International Journal of Law*, 1(3), 214–226.
- Ivanova, S. (2025). Legislative responses to deepfake technology and digital evidence manipulation. *Journal for Civil and Criminal Law for Legislative Studies*, 1(1), 21–25.
- Jafar, Z., Quick, J. D., Larson, H. J., Venegas-Vera, V., Napoli, P., Musuka, G., Dzinamarira, T., Meena, K. S., Kanmani, T. R., & Rimányi, E. (2023). Social media for public health: Reaping the benefits, mitigating the harms. *Health Promotion Perspectives*, 13(2), 105.
- Jaff, D. (2023). The surge of spreading harmful information through digital technologies: A distressing reality in complex humanitarian emergencies. *The Lancet Global Health*, 11(6), e821–e822.
- Moeljatno. (2008). *Principles of criminal law*. Rineka Cipta.
- Noggle, R. (2025). *Manipulation: Its nature, mechanisms, and moral status*. Oxford University Press.
- NurJannah, N., Sulatri, K., & Budiarti, D. (2026). The phenomenon of deep fake as a crime on social media related to defamation. *Causa: Journal of Law and Citizenship*.
- Park, T. J., & Rohatgi, A. (2024). Balancing the platform responsibility paradox: A case for amplification regulation to mitigate the spread of harmful but legal content online. *Computer Law & Security Review*, 52, 105960.
- Pasat, A. O. (2022). Criminal liability of a juvenile offender. *Acta Universitatis Danubius. Juridica*, 18(3), 71–90.
- Perriello, L. E. (2026). Blurred realities: Legal strategies for the deepfake era. *Maastricht Journal of European and Comparative Law*.
- Satrul, H. S., & Hasanah, N. H. (2026). Reconstruction of criminal liability for deepfake misuse in defamatory defenses in Indonesia. *Rechtsvinding*.
- Sijabat, S. A. U., & Lukitasari, D. (2024). Deepfake pornographic images and videos content as a form of defamation crime. *Recidive: Jurnal Hukum Pidana dan Penanggulangan Kejahatan*.
- Surjatmodjo, D., Unde, A. A., Cangara, H., & Sonni, A. F. (2024). Information pandemic: A critical review of disinformation spread on social media and its implications for state resilience. *Social Sciences*, 13(8), 418.
- Titahelu, J. A. S. (2022). Legal liability for crimes against humanity as a form of human rights violation: Criminal law perspective. *Law Reform*, 18(1), 28–42.
- Wafi, M. S., Wisnubroto, A., & Prayudi, Y. (n.d.). Artificial intelligence-based deepfake crimes: A conception of culpability principle as a criminal liability reform. *Legal Reform*, 29(2).
- Wang, L., Hu, Q., Lu, W., & Luo, X. (2025). Diffusion-based adversarial identity manipulation for facial privacy protection. In *Proceedings of the 33rd ACM International Conference on Multimedia* (pp. 11562–11571).
- Widyaningrum, T., Khoirunnisa, K., & Jubaidi, D. (2024). Corporate criminal liability: An analysis of corporate crime perpetrators under positive law in Indonesia. *JCH (Jurnal Cendekia Hukum)*, 9(2), 146–157.
- Law Number 1 of 2023 concerning the Criminal Code. (2023).
- Law Number 1 of 2024 concerning the Second Amendment to Law Number 11 of 2008 concerning Information and Electronic Transactions. (2024).
- Law Number 27 of 2022 concerning Personal Data Protection. (2022).