JURNAL AKTA
Volume 13 No. 1, March 2025
SINTA 2,
Decree No. 164/E/KPT/2021

CRIMINAL LIABILITY FOR DEEPFAKE CYBERCRIME: EVIDENTIARY CHALLENGES AND REGULATORY GAPS IN INDONESIAN LAW

M. Rochman[1)]
[1)] Universitas Prof. Dr. Hazairin SH, Bengkulu, Indonesia, E-mail: drrochman18@gmail.com

Abstract. *The rapid advancement of artificial intelligence has enabled the emergence of deepfakes synthetic audio-visual content generated using machine-learning models such as Generative Adversarial Networks. While deepfake technology offers beneficial applications, its misuse in cybercrimes such as identity fraud, extortion, and others raises serious concerns in Indonesia due to the absence of explicit legal regulation. This study aims to analyze the criminal liability associated with deepfake-based cybercrime by examining evidentiary challenges and identifying regulatory gaps within the Indonesian Criminal Code and the Electronic Information and Transactions Law. Employing a qualitative library research methodology, this study synthesizes primary legal sources, academic literature, and authoritative institutional reports. Findings reveal that deepfake cybercrime presents substantial evidentiary difficulties, particularly in authenticating electronic evidence and maintaining chain-of-custody integrity, as deepfake content often lacks detectable traces of manipulation. Furthermore, the Criminal Code and Electronic Information and Transactions Law provide only general provisions for forgery, defamation, pornography, and electronic manipulation, without addressing AI-based content fabrication or synthetic impersonation. This regulatory vacuum restricts law-enforcement capability, creates inconsistent legal interpretations, and weakens victim protection. The study concludes that Indonesia requires explicit statutory definitions, specialized criminal provisions for deepfake offenses, and standardized forensic guidelines to ensure legal certainty.*

Keywords: *Artificial Intelligence, Criminal Liability, Deepfake Cybercrime, Electronic Evidence, Evidentiary Challenges.*

1. Introduction

The rapid advancement of Artificial Intelligence (AI) technology has enabled the creation of highly realistic synthetic audio-visual content known as deepfakes a technology capable of producing fake videos, images, or voices that are nearly indistinguishable from authentic ones (Al Alif et al., 2025). Although initially developed for artistic and entertainment purposes, this technology has been increasingly misused in various forms of cybercrime, such as fraud, digital identity theft, extortion, non-consensual

pornography, and defamation. In Indonesia, the growing prevalence of deepfake misuse has sparked significant concern regarding privacy, reputation, and digital security (Verma, 2024).

Cybercrime refers to criminal activities conducted through digital networks and computer systems, including data theft, ransomware, phishing, and technology-enabled fraud. In the past five years, the rapid development of generative artificial intelligence has expanded the cyber-threat landscape, enabling the creation of highly realistic synthetic images, audio, and video known as deepfakes. According to Europol (2022), the increasing accessibility of face-swapping, voice-cloning, and lip-synchronization tools has introduced new challenges for law enforcement, as deepfakes can be deployed to deceive individuals and institutions.

Criminal misuse of deepfakes includes financial fraud such as impersonating executives in video or audio calls to authorize unauthorized transfers sexual extortion and the distribution of fabricated intimate content, as well as the manipulation of audio-visual evidence to damage reputations. Recent reports highlight large-scale incidents, including corporate losses resulting from deepfake video-call scams (Milmo, 2024). Interpol (2024) also warns of the rising use of AI-generated child sexual abuse material and deepfake-based extortion schemes, demonstrating how traditional cybercrime techniques are now reinforced by advanced AI-driven deception.

The challenges posed by deepfake-enabled cybercrime span technical, legal, and organizational domains. Detection technologies continue to lag behind the rapid evolution of synthetic-media generation, creating an ongoing arms race between offenders and digital forensics specialists. Policy gaps also persist, as legal systems struggle to address evidentiary integrity and criminal liability in cases involving manipulated content. Reports from the European Parliament (2025), Corder (2025) and Europol (2025) emphasize the need for multi-stakeholder cooperation, improved forensic capacity, and updated regulatory frameworks to mitigate these risks effectively.

However, Indonesia's current legal framework has yet to explicitly anticipate the deepfake phenomenon. Although laws such as the Information and Electronic Transactions Law (*Undang Undang Infromasi dan Transaksi Elektronik*/UU ITE), Law Number 19 of 2016 as its amendment, and the Indonesian Criminal Code (*Kitab Undang Undang Hukum Pidana*/KUHP) contain general provisions on cyber-related offenses, none of them specifically regulate AI-based content manipulation (Khan et al., 2025; Putri & Rusmini, 2025). As a result, law enforcement authorities face serious difficulties in identifying offenders, proving digital evidence, and fulfilling the specific elements of a crime (Afif, 2025).

The technical nature of deepfakes particularly their reliance on neural networks such as Generative Adversarial Networks (GANs) creates substantial challenges in the authentication and verification of digital evidence (Masood et al., 2023). Within the judicial process, reliable digital evidence is essential to ensure lawful and fair adjudication. However, the detection of deepfakes remains problematic due to technological limitations: existing detection models often fail against newer deepfake variants or exhibit reduced reliability in real-world scenarios (Wang et al., 2024).

Empirical and juridical analyses by several scholars indicate that reliance on general provisions such as those governing data falsification, pornography, or defamation under the KUHP and UU ITE remains insufficient to address the distinctive features of deepfake-related crimes, especially regarding AI content manipulation, identity falsification, and the wide scope of potential harm (Afif, 2025). Meanwhile, the phenomenon continues to evolve rapidly in both scale and complexity (Arvitto, 2025).

Given these circumstances, there is an urgent need to conduct a comprehensive study on criminal liability in AI-based deepfake cybercrimes, focusing on evidentiary challenges and regulatory gaps in the KUHP and UU ITE. The primary urgency lies in assessing whether Indonesia's existing criminal law is adequate to address deepfake crimes and how it might be reformed to ensure legal certainty, justice, and victim protection (Abidin, 2023).

Previous studies have highlighted the presence of a legal vacuum surrounding deepfake crimes in Indonesia and the inadequacy of current regulations in providing justice for victims or deterrence for perpetrators (Arvitto, 2025). However, most existing works remain general, focusing primarily on cases of pornographic deepfakes or identity fraud, without a holistic analysis that integrates digital forensic proof, criminal elements, and a normative examination of both the KUHP and UU ITE.

Therefore, this research aims to critically analyze evidentiary challenges in deepfake cybercrime cases, including issues of digital evidence validity and detection reliability. Futhermore, identify regulatory gaps in Indonesia's KUHP and UU ITE concerning deepfakes and their implications for law enforcement effectiveness, and formulate normative recommendations for developing specific criminal regulations to govern deepfakes in Indonesia, thereby enhancing both victim protection and judicial integrity. This study is expected to contribute theoretically and practically to the modernization of Indonesia's criminal law in the digital age.

Based on the background and objectives above, the problem formulation for this research is as follows:

1. How do evidentiary challenges arise in deepfake cybercrime cases, including issues of digital evidence validity and detection reliability?

2. What regulatory gaps exist in Indonesia's KUHP and UU ITE concerning deepfakes and their implications for law enforcement effectiveness?

3. How can normative recommendations for developing specific criminal regulations to govern deepfakes in Indonesia be formulated to enhance victim protection and judicial integrity?

2. Research Methods

This study employs a qualitative research approach using the library research (literature study) method. A qualitative approach is appropriate because it enables an in-depth exploration of complex legal and social phenomena particularly issues related to criminal liability in AI-based cybercrimes such as deepfake manipulation through interpretive

understanding of legal norms, doctrines, and enforcement practices (Creswell & Poth, 2016; Moleong, 2021). The library research method is intended to examine and analyze various primary, secondary, and tertiary legal sources relevant to the research topic (Zed, 2018). Through this method, the researcher aims to identify normative gaps and examine how Indonesia's positive law responds to the rapid development of deepfake technology.

The data sources in this study consist of three categories. First, primary legal materials include legislation such as the Indonesian Criminal Code, Law Number 19 of 2016 on Amendments to Law Number 11 of 2008 concerning Electronic Information and Transactions, as well as relevant international instruments, including the Budapest Convention on Cybercrime. Second, secondary legal materials consist of academic research, scientific journals, legal books, scholarly articles, and official institutional reports that discuss cybercrime and deepfake technology (Soerjono & Mamudji, 2010). Third, tertiary legal materials include legal dictionaries, legal encyclopedias, and other credible online resources that serve to clarify and complement the understanding of the primary and secondary legal sources. These layered data sources ensure that the research comprehensively captures both the theoretical framework and practical implementation of criminal law related to deepfake technology.

Data were collected using the documentation study technique, which involves the systematic identification, examination, and analysis of academic and legal documents (Bungin, 2020). The researcher conducted an extensive review of national and international academic journals addressing cybercrime, artificial intelligence, and criminal lawboth within Indonesian legal contexts and in comparative law frameworks. Data collection also included legal documents, court decisions, and expert opinions from the fields of criminal law and information technology. To maintain analytical rigor, only credible, peer-reviewed, and up-to-date sources were selected, ensuring that findings accurately represent the current normative and empirical conditions.

The study utilizes a qualitative descriptive analysis combined with a normative-juridical approach. Data analysis was carried out through three main stages. First, data reduction was conducted by selecting and categorizing relevant information according to key research themes, namely criminal liability, evidentiary challenges, and regulatory gaps in the KUHP and UU ITE. Second, data presentation involved organizing the findings into descriptive narratives that illustrate the relationship between existing legal norms and the phenomenon of deepfake-related crimes. Third, conclusion drawing was undertaken through a systematic interpretation of the analyzed data to address the research questions and to propose normative recommendations for legal reform (Miles et al., 2020). Through this method, the study seeks to provide both theoretical and practical contributions to the development of Indonesian criminal law in the context of AI-driven cybercrimes, while offering a normative foundation for future regulatory adaptation to digital technological advancements.

3. Results and Discussion

3.1. Evidentiary Challenges in Deepfake Cybercrime Cases

One of the main challenges in prosecuting deepfake-related cybercrimes in Indonesia lies in the difficulty of validating and authenticating digital evidence. Deepfake

technology, powered by Generative Adversarial Networks (GANs), can produce synthetic audio-visual content that is nearly indistinguishable from authentic materials, making it extremely difficult for investigators and courts to establish the origin, authenticity, and integrity of such content. Pursuant to Article 5 paragraphs (1) and (2) of the Electronic Information and Transactions Law (*Undang Undang Informasi dan Transaksi Elektronik*/UU ITE) Law Number 11 of 2008 as amended by Law Number 19 of 2016 electronic information and/or documents are deemed legally valid evidence, as long as the information is accurate and has not been altered (Indonesia, 2008, 2016a).

In deepfake cases, this standard of authenticity becomes problematic because manipulation at the pixel or waveform level is almost undetectable through conventional forensic methods. Current digital forensic tools often fail to differentiate between synthetic and original content, especially when perpetrators use advanced AI tools to mask metadata traces. This raises a burden-of-proof dilemma, where prosecutors must prove both falsification and intent, while defense attorneys can easily invoke reasonable doubt concerning authenticity.

Moreover, Article 184 of the Indonesian Criminal Procedure Code (*Kitab Undang Undang Hukum Pidana*/KUHAP) specifies the types of admissible evidence, including witness testimony, expert opinions, and documentary evidence. Yet, unlike physical evidence, deepfake content challenges the "chain of custody" principle because electronic data can be easily replicated or modified without trace. Without standardized forensic protocols and reliable detection systems, courts risk admitting fabricated evidence or rejecting legitimate ones, undermining procedural integrity.

The absence of national standards for digital forensic authentication particularly in identifying AI-generated content creates inconsistencies in evidentiary assessment across cases. Consequently, judicial authorities face technical barriers that hinder fair adjudication and reduce public confidence in the justice system.

Therefore, strengthening detection reliability and evidentiary verification mechanisms is essential to address deepfake-related cybercrime. The development of standardized digital forensic protocols, the adoption of AI-assisted detection systems, and the enhancement of technical expertise among law-enforcement officers and judicial actors could improve the credibility of electronic evidence presented before the court. Such measures would help ensure that the principles of evidentiary validity, procedural fairness, and legal certainty are maintained when dealing with increasingly sophisticated forms of synthetic digital manipulation.

In addition to technical verification difficulties, another evidentiary challenge arises from the limited reliability and generalizability of existing deepfake detection systems (Wang et al., 2024; Bendiab et al., 2025). Many current detection models are trained on specific datasets and perform well under controlled conditions; however, their effectiveness often declines significantly when applied to real-world scenarios involving diverse formats, compression levels, or adversarial manipulation techniques. This inconsistency weakens the probative value of digital evidence, as courts require a high degree of certainty regarding authenticity before admitting such evidence. The inability of forensic tools to consistently identify manipulation across varying contexts creates uncertainty not only for investigators but also for judges in assessing evidentiary weight (Sunil et al., 2025). Furthermore, as deepfake generation techniques evolve rapidly, detection mechanisms

tend to lag behind, resulting in a persistent technological asymmetry between perpetrators and law enforcement (Bendiab et al., 2025).

Another important dimension concerns the integrity of forensic processes in increasingly complex digital ecosystems. The proliferation of interconnected platforms, including social media and cloud-based systems, complicates the tracking and preservation of digital evidence. Deepfake content may originate from multiple sources, be redistributed across platforms, and undergo successive modifications, making it difficult to reconstruct its provenance (Patel et al., 2023). This fragmentation challenges not only attribution but also the maintenance of forensic integrity throughout the investigation lifecycle. In such contexts, even minor uncertainties in data handling procedures can be exploited to challenge the admissibility of evidence in court. Additionally, the difficulty in ensuring reliable documentation and verification mechanisms for digital content further exacerbates concerns regarding evidentiary credibility (Bendiab et al., 2025).

Beyond technical and procedural issues, evidentiary challenges in deepfake cases are also influenced by broader socio-legal considerations, particularly those related to victim impact and evidentiary sensitivity (Malik et al., 2024). Deepfake-related offenses frequently involve highly personal or reputational harm, such as fabricated explicit content or identity-based manipulation, which raises concerns about privacy, dignity, and secondary victimization during legal proceedings. These factors may affect how evidence is presented, examined, and contested in court, potentially limiting the availability or disclosure of crucial digital materials. Moreover, the persuasive realism of deepfakes can influence judicial perception and introduce cognitive bias, further complicating objective evidentiary evaluation. As a result, the intersection between technological deception and human perception adds an additional layer of complexity to the already challenging task of ensuring evidentiary reliability in deepfake cybercrime cases (Wang et al., 2024).

3.2. Regulatory Gaps in the KUHP and UU ITE Regarding Deepfakes

The Indonesian Penal Code (KUHP) and the Electronic Information and Transactions Law (UU ITE) do not yet contain explicit provisions that specifically regulate deepfake-related offenses (Indonesia, 1946; Indonesia, 2008). Under Article 27 paragraph (1) of the UU ITE, disseminating electronic information violating decency or morality is punishable, while Article 28 paragraph (1) criminalizes spreading false or misleading information. However, these provisions are content-oriented, focusing on the nature of information rather than the modus operandi or technology used to create it. Thus, deepfake acts involving synthetic impersonation or fraud without explicit obscene or defamatory intent may fall outside the law's scope.

In the KUHP, provisions such as Article 263 (forgery) and Article 310 (defamation) may theoretically apply. However, these articles presuppose tangible documents or verbal statements concepts inadequate to encompass AI-generated imagery or voice synthesis (Indonesia, 1946). For instance, the element of a "false document" in Article 263 does not explicitly include digital files, creating interpretive uncertainty when prosecutors attempt to link deepfake outputs to forgery offenses.

Moreover, Articles 30 and 32 of the UU ITE address unauthorized access and manipulation of electronic systems but are designed to criminalize system intrusion

rather than content creation (Indonesia, 2016b). A person fabricating a deepfake video using publicly available software without hacking a system technically does not violate these provisions. This regulatory vacuum allows many harmful deepfake acts such as identity theft, sexual exploitation, or political disinformation to go unpunished, particularly when they do not fit neatly into existing statutory categories.

This legal gap significantly impairs law enforcement effectiveness, forcing investigators and prosecutors to rely on analogical interpretation or indirect application of defamation or pornography articles, which leads to inconsistent jurisprudence. Without specific criminal norms for deepfakes, the law remains reactive rather than preventive, weakening deterrence and victim protection in Indonesia's digital landscape.

Another dimension of the regulatory gap lies in the absence of explicit legal recognition of AI-generated manipulation as a distinct form of digital falsification. Current Indonesian criminal provisions tend to regulate the content or the resulting harm rather than the technological process used to produce the content. Deepfakes, however, operate through complex machine-learning systems that fabricate realistic representations of individuals without their consent. Because the legal framework does not explicitly acknowledge this technological mechanism, law-enforcement authorities often struggle to categorize deepfake acts within existing legal definitions, leading to uncertainty in determining applicable charges and legal elements that must be proven in court.

Furthermore, the lack of specific procedural guidance for investigating AI-based offenses limits the operational capacity of investigators and prosecutors. Law-enforcement agencies may encounter difficulties in identifying perpetrators who utilize anonymous digital platforms, cloud-based AI tools, or cross-border servers to generate and distribute deepfake content. Without clear investigative authority, specialized training, or technological support embedded in the legal framework, the process of digital evidence collection, attribution of responsibility, and prosecution becomes significantly more complex. This situation can delay investigations and reduce the likelihood of successful prosecution.

Consequently, these regulatory deficiencies have broader implications for law enforcement effectiveness and legal certainty. The absence of precise statutory norms addressing deepfake production, distribution, and malicious use weakens deterrence and allows perpetrators to exploit legal loopholes. It also creates inconsistencies in judicial interpretation when courts attempt to apply traditional criminal provisions to emerging AI-driven misconduct. Strengthening the legal framework by explicitly regulating deepfakes would therefore be essential not only to close normative gaps but also to enhance the effectiveness, consistency, and credibility of law enforcement in Indonesia's rapidly evolving digital environment.

In addition to substantive normative gaps, regulatory challenges also stem from the general nature of cybercrime provisions within the KUHP and the UU ITE, which tend to be broadly formulated and open to multiple interpretations. Such ambiguity often leads to inconsistent enforcement in practice, particularly when authorities confront emerging technologies like deepfakes (Handayani et al., 2025). The lack of precise legal language creates uncertainty in determining the scope and limits of criminal liability, and in some instances, may even open the door to overbroad application of the law, potentially affecting fundamental rights such as freedom of expression in digital spaces. As

technological advancements in artificial intelligence continue to evolve rapidly, the disparity between static legal norms and dynamic digital threats becomes increasingly pronounced (Suseno et al., 2025).

Furthermore, the fragmented structure of existing regulations presents an additional obstacle in addressing deepfake-related risks comprehensively. Current legal provisions are dispersed across different statutory instruments and lack integration into a cohesive framework capable of responding to AI-driven misconduct. This fragmentation contributes to legal uncertainty, particularly in defining criminal responsibility for acts involving identity manipulation and digital misinformation (Khan et al., 2025). The widespread dissemination of deepfake content through social media platforms illustrates that its impact extends beyond individual harm, potentially undermining public trust in digital information ecosystems. In this context, the absence of adaptive and coordinated regulation weakens the state's capacity to effectively mitigate technology-driven disinformation threats (Ghariwala, 2025).

Moreover, the inability of legal frameworks to keep pace with increasingly sophisticated AI-based attack methods highlights the need for a more technologically informed regulatory approach. Advanced deepfake impersonation techniques, for instance, require legal recognition not only of their harmful consequences but also of the underlying technological processes that enable them (Samrouth et al., 2025). Without incorporating these technical dimensions, the law risks remaining reactive and insufficient in addressing new forms of cyber-enabled misconduct. Therefore, a more forward-looking regulatory strategy supported by cross-sector collaboration and the integration of technological expertise is essential to ensure that legal systems remain responsive and effective in the face of evolving digital threats (Setyaningrum et al., 2022).

3.3. Normative Recommendations for Legal Reform

To address these challenges, a comprehensive legal framework is required to govern deepfake technology and AI-generated content within Indonesia's criminal law system. First, it is necessary to introduce a statutory definition of deepfake, defined as "digitally manipulated visual, audio, or audiovisual content generated through artificial intelligence with the intent to deceive or cause harm." Such a definition should clearly distinguish malicious uses from legitimate creative, artistic, or educational applications.

Second, special criminal provisions should be established by amending the Electronic Information and Transactions Law (*Undang Undang Infromasi dan Transaksi Elektronik*/UU ITE) or by incorporating specific articles within the New Criminal Code (Law Number 1 of 2023) to explicitly criminalize the creation, distribution, and use of deepfakes for fraudulent, defamatory, or exploitative purposes (Indonesia, 2023). This would provide clearer legal grounds for prosecution and strengthen law-enforcement capacity in addressing AI-based cybercrime.

Third, the government should develop national standards for digital forensic examination and AI-content verification, including mechanisms such as blockchain-based authentication, digital watermarking, and AI-detection protocols that can be formally recognized by the Supreme Court. Standardized procedures would help ensure consistency in evidentiary evaluation and improve the reliability of digital evidence in court proceedings.

In addition, victim protection mechanisms must be strengthened through restorative and compensatory measures, particularly for victims of deepfake-based sexual exploitation. These protections should align with the constitutional guarantee of personal security and dignity as recognized in Article 28G of the Indonesian Constitution (Indonesia, 1945). Finally, effective regulation requires cross-sectoral collaboration among key institutions, including the National Police, the Ministry of Communication and Information Technology (*Kementerian Komunikasi dan Informasi*/Kominfo), and specialized cybercrime units, to synchronize technological capabilities with legislative enforcement and enhance the overall effectiveness of cybercrime prevention. By integrating these reforms, Indonesia can move toward a proactive and adaptive criminal justice system capable of addressing deepfake-related offenses effectively. A harmonized legal structure acknowledging both technological realities and human rights principles will enhance judicial integrity, ensure fair adjudication, and restore public trust in digital law enforcement.

Beyond institutional and doctrinal adjustments, effective legal reform concerning deepfake regulation must be grounded in a robust normative framework that balances competing legal values and societal interests (Burton, 2012). Normative legal scholarship emphasizes that reform should not merely respond to technological developments but must also reflect fundamental principles such as justice, human dignity, and the rule of law. In this regard, a pluralistic approach one that accommodates multiple values rather than prioritizing a single objective becomes essential in addressing the multifaceted harms caused by deepfakes, including fraud, reputational damage, and social manipulation. Such an approach allows lawmakers to reconcile tensions between innovation, freedom of expression, and protection against harm, thereby producing more balanced and context-sensitive regulations (Dagan, 2014).

The formulation of legal reforms should integrate both empirical findings and normative reasoning to ensure that proposed regulations are not only theoretically sound but also practically feasible. The method of reflective equilibrium, for instance, enables policymakers to align legal principles with real-world constraints, such as technological limitations, enforcement capacity, and potential unintended consequences (Van der Burg, 2019). This is particularly relevant in the context of deepfake-related cybercrime, where rapid technological evolution often outpaces institutional readiness. Empirical insights into how deepfakes are used in organized fraud or digital misinformation ecosystems can inform more targeted and adaptive legal responses, while normative analysis ensures that such measures remain consistent with constitutional and ethical standards (Lin, 2025).

In addition, meaningful legal reform requires sensitivity to sociocultural contexts and a long-term commitment to institutional development. Deepfake harms are not experienced uniformly; they often disproportionately affect vulnerable individuals, particularly in cases involving identity exploitation or non-consensual content (Burton, 2012; Ali et al., 2025). As such, reform efforts should incorporate victimological perspectives and broader access-to-justice considerations, ensuring that legal mechanisms are responsive to the needs of affected communities. This includes not only formal legislative changes but also the development of complementary guidelines, institutional practices, and cross-sector collaboration frameworks. Without sustained engagement and adaptive strategies, legal reform risks becoming symbolic rather than transformative, failing to address the evolving challenges posed by AI-driven digital manipulation (Ghariwala, 2025).

4. Conclusion

This study concludes that deepfake-based cybercrime presents two critical problems in Indonesian criminal law: significant evidentiary challenges and clear regulatory gaps. First, the authentication of deepfake electronic evidence is highly problematic due to the sophistication of AI-generated manipulation, which often cannot be detected using existing forensic tools. This undermines the reliability of evidence under both the Information and Electronic Transactions Law and Criminal Procedure Code. Second, neither the Criminal Procedure Code nor the Information and Electronic Transactions Law contain explicit provisions addressing synthetic impersonation, AI-generated falsification, or malicious deepfake distribution. As a result, law-enforcement authorities must rely on general articles such as forgery, defamation, or morality-based offenses that do not adequately capture the unique nature and harm potential of deepfake technology. Therefore, the current legal framework is insufficient to ensure legal certainty, deterrence, and victim protection in deepfake-related cases.

To respond effectively, Indonesia must urgently strengthen its legal and institutional framework. Legislators should introduce clear legal definitions of deepfakes and formulate specific criminal provisions regulating their creation, dissemination, and harmful use. Law-enforcement institutions require standardized digital forensic protocols and AI-content detection systems to ensure accurate authentication of electronic evidence. Inter-agency cooperation particularly between Ministry of Communication and Information Technology, law-enforcement cyber units, and forensic laboratories must be enhanced to improve investigative capacity. Victim-protection mechanisms should also be expanded, particularly in cases of deepfake-based sexual exploitation and identity abuse.

Future research should focus on empirical examination of deepfake cases handled by Indonesian courts to evaluate practical enforcement obstacles. Comparative legal studies examining regulatory approaches adopted by the EU, US, and other jurisdictions may offer valuable insights for shaping Indonesia's legislative reform. Technical-legal interdisciplinary research is also essential, particularly studies on deepfake detection tools, evidentiary reliability, and AI governance frameworks. Longitudinal research is recommended to assess how evolving AI technologies continue to influence cybercrime patterns and legal responses.

5. References

Journals:

Abidin, M. I. (2023). Legal review of liability from deepfake artificial intelligence that contains pornography. *MIMBAR: Jurnal Sosial Dan Pembangunan*, 344–352.

Afif, M. (2025). Tindak Pidana Deepfake Pornography di Indonesia: Analisis Yuridis terhadap Kekosongan Norma dalam KUHP dan UU ITE. *MERDEKA : Jurnal Ilmiah Multidisiplin*, *3*(2 SE-Articles), 27–35.

Al Alif, S. S., Rindiani, A., & Marhayati, C. (2025). Konstruksi Hukum Pidana Dalam Penanggulangan Kejahatan Siber Berbasis Teknologi Deepfake di Indonesia. *Legal Standing: Jurnal Ilmu Hukum*, *9*(5 SE-Articles), 1169–1183.

Ali, M., Fernando, Z. J., Huda, C., & Mahmutarom, M. (2025). Deepfakes and Victimology: Exploring the Impact of Digital Manipulation on Victims. *Substantive Justice International Journal of Law*, *8*(1).

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.

Bendiab, G., Haiouni, H., Moulas, I., & Shiaeles, S. (2025). Deepfakes in digital media forensics: Generation, AI-based detection and challenges. *Journal of Information Security and Applications*, *88*, 103935.

Burton, S. J. (2012). Normative legal theories: The case for pluralism and balancing. *Iowa L. Rev.*, *98*(8), 535-552.

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Dagan, H. (2014). Normative jurisprudence and legal realism. *University of Toronto Law Journal*, *64*(3), 442-457.

Ghariwala, L. (2025). Impact of Deepfake Technology on Social Media: Detection, Misinformation and Societal Implications. *INTERNATIONAL JOURNAL*, *13*(3), 2982-2986.

Handayani, E. P., Arifin, Z., & Fernando, Z. J. (2025). Criminal Penalties in Cyberspace: Between the Development of Digital Democracy and Authoritarianism. *Indonesian Journal of Criminal Law Studies*, *10*(1), 45-82.

Interpol. (2024). *Beyond Illusions: The misuse of deepfake technology*. file:///C:/Users/asus/Downloads/BEYOND ILLUSIONS_Report_2024.pdf

Khan, A. A., Laghari, A. A., Bacarra, R., Alroobaea, R., Algarni, S., Baqasah, A. M., & Alsayaydeh, J. A. J. (2025). Cybersecurity, digital forensics, and the IoT for deepfake investigation on social media platforms: A review. *Human-Centric Computing and Information Sciences*, *15*(3), 103632.

Lin, L. S. (2025). Examining the role of deepfake technology in organized fraud: Legal, security, and governance challenges. *Frontiers in Law*, *4*, 6-17.

Masood, M., Nawaz, M., Malik, K. M., Javed, A., Irtaza, A., & Malik, H. (2023). Deepfakes generation and detection: State-of-the-art, open challenges, countermeasures, and way forward. *Applied Intelligence*, *53*(4), 3974–4026.

Patel, Y., Tanwar, S., Gupta, R., Bhattacharya, P., Davidson, I. E., Nyameko, R., ... & Vimal, V. (2023). Deepfake generation and detection: Case study and challenges. *IEEE Access*, *11*, 143296-143323.

Putri, P. N. T., & Rusmini, A. A. A. N. T. (2025). Pengaturan Hukum dalam Penanggulangan Deepfake Artificial Intelligence (AI) terhadap Anak sebagai Korban Kejahatan Siber di Indonesia. *Jurnal Magister Hukum Udayana (Udayana Master Law Journal)*, *14*(3), 713–728.

Samrouth, K., El Housseini, P., & Deforges, O. (2025). Siamese network-based detection of deepfake impersonation attacks with a person of interest approach. *ACM*

Transactions on Multimedia Computing, Communications and Applications, *21*(3), 1-23.

Setyaningrum, W., Morana, A. C., Vaizi, K. N., Damarina, R., Akbar, S. A., & Oktasari, S. (2022). Anticipation of the ITE Law and Reconciliation of Its Forms Freedom of Expression through the E-Hights Website. *Jurnal Hukum Novelty (1412-6834)*, *13*(2), 93-112.

Soerjono, S., & Mamudji, S. (2010). *Penelitian Hukum Normatif Suatu Tinjauan Singkat*. Jakarta: RajaGrafindo Persada.

Sunil, R., Mer, P., Diwan, A., Mahadeva, R., & Sharma, A. (2025). Exploring autonomous methods for deepfake detection: A detailed survey on techniques and evaluation. *Heliyon*, *11*(3).

Suseno, S., Ramli, A. M., Mayana, R. F., Safiranita, T., & Aurellia Nathania Tiarma, B. (2025). Cybercrime in the new criminal code in Indonesia. *Cogent Social Sciences*, *11*(1), 2439543.

Van der Burg, W. (2019). The merits of law: An argumentative framework for evaluative judgements and normative recommendations in legal research. *ARSP: Archiv für Rechts-und Sozialphilosophie/Archives for Philosophy of Law and Social Philosophy*, *3*(2), 11-43.

Verma, K. (2024). Digital Deception: The Impact of Deepfakes on Privacy Rights. *Lex Scientia Law Review*, *8*(2), 859–896.

Wang, T., Liao, X., Chow, K. P., Lin, X., & Wang, Y. (2024). Deepfake detection: A comprehensive survey from the reliability perspective. *ACM Computing Surveys*, *57*(3), 1–35.

Zed, M. (2018). *Metode penelitian kepustakaan*. Yogyakarta: Yayasan Pustaka Obor Indonesia.

Books:

Handoko, B.L., Lindawati, A.S.L., Budiarto, A.Y. 2020. *Using Mixed Reality to Provide* Bungin, B. (2020). *Metodologi penelitian kualitatif*. Jakarta: Kencana.

Miles, H., Huberman, A. M., & Saldana. (2020). *Qualitative data analysis: A methods sourcebook*. New York: Sage Publications, Inc.

Moleong, L. J. (2021). *Metodologi penelitian kualitatif*. Bandung: PT. Remaja Rosdakarya.

Conference:

Malik, S., Surbhi, A., & Roy, D. (2024, October). Blurring boundaries between truth and illusion: analysis of human rights and regulatory concerns arising from abuse of deepfake technology. In *AIP Conference Proceedings* (Vol. 3220, No. 1, p. 050016). AIP Publishing LLC.

Regulation:

Indonesia. (1945). *Undang-Undang Dasar Negara Republik Indonesia Tahun 1945*. Jakarta: Sekretariat Negara.

Indonesia. (1946). *Kitab Undang-Undang Hukum Pidana (KUHP)*. Jakarta: Pemerintah Republik Indonesia.

Indonesia. (2008). *Law No. 11 of 2008 on Electronic Information and Transactions (UU ITE)*. Jakarta: Government of the Republic of Indonesia.

Indonesia. (2016a). *Law No. 11 of 2008 on Electronic Information and Transactions (as amended by Law No. 19 of 2016)*.

Indonesia. (2016b). *Law No. 19 of 2016 on Amendments to Law No. 11 of 2008 on Electronic Information and Transactions*. Jakarta: Government of the Republic of Indonesia.

Indonesia. (2023). *Law No. 1 of 2023 on the New Criminal Code*. Jakarta: Government of the Republic of Indonesia.

Report:

Corder, M. (2025). *AI is turbocharging organized crime, EU police agency warns*. https://apnews.com/article/europe-crime-europol-ai-security-cyber-attack-846847536f6feb2bbb423943fd96e1f1

European Parliament. (2025). *Children and Deepfakes*. https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/775855/EPRS_BRI%282025%29775855_EN.pdf?

Europol. (2022). *Facing reality? Law enforcement and the challenge of deepfakes*. https://www.europol.europa.eu/cms/sites/default/files/documents/Europol_Innovation_Lab_Facing_Reality_Law_Enforcement_And_The_Challenge_Of_Deepfakes.pdf?

Milmo, D. (2024). Company worker in Hong Kong pays out £20m in deepfake video call scam. The Guardian. Retrieved in May 23, 2025 from https://www.theguardian.com/world/2024/feb/05/hong-kong-company-deepfake-video-conference-call-scam?utm_source=chatgpt.com.