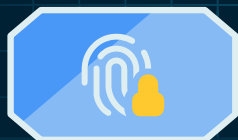
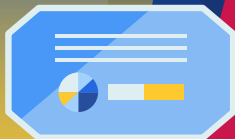


UNMASKING DECEPTIONS

Urging the Law to Take a Role
in Mitigating and Tackling
the Malicious Deepfake





Author

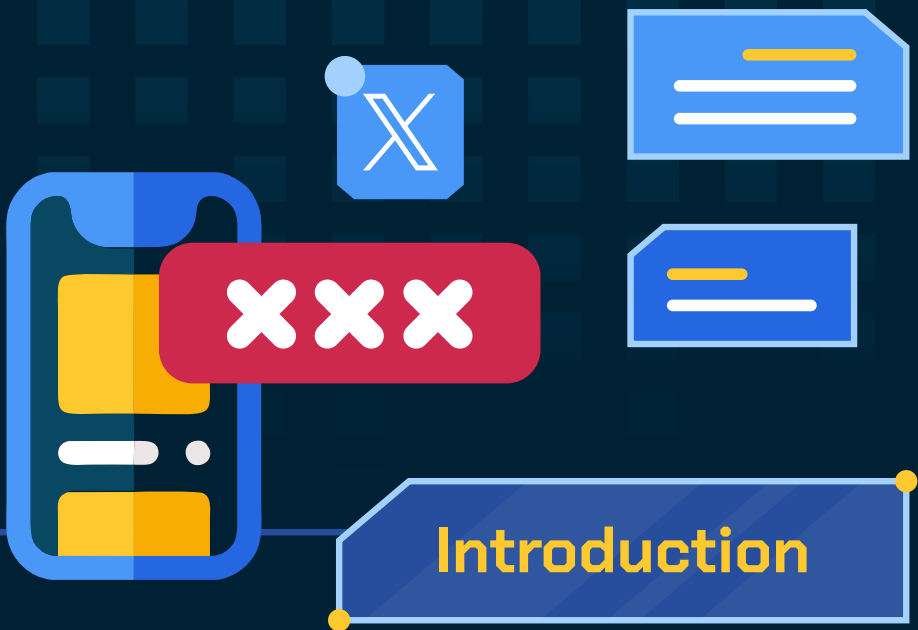
M. Irfan Dwi Putra

Editor

Rizka Khairunissa Herdiani

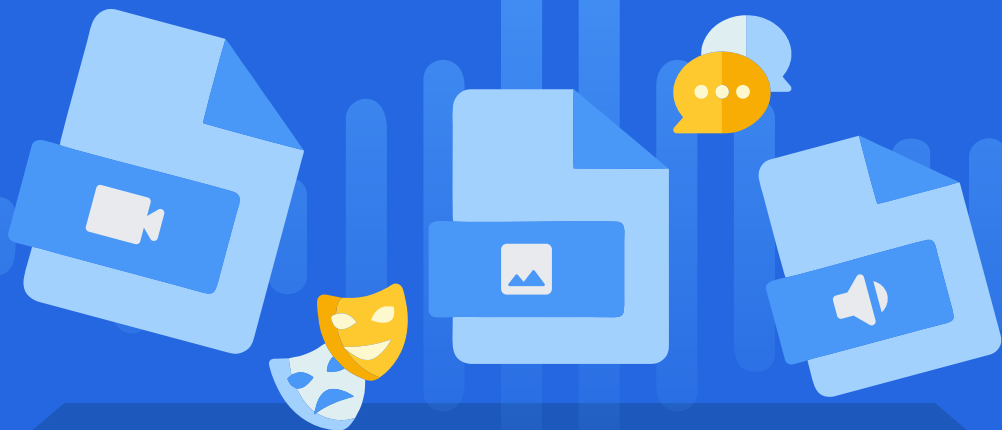
Designer & Layouter

Ahmad Nur Hasan



In late January 2024, the online community was horrified by the spread of pornographic content containing Taylor Swift's image on the X—previously known as Twitter. The content has drawn condemnations from Swifties—Taylor Swift's fans—around the world. The White House also responded to the content by calling it “alarming”.¹ After its horrendous release, however, X, the platform where the content was posted, was not responsive in taking down the content. The content circulated in the X for about a week and got over 47 million views. It was only after Taylor Swift's fans reported the content, that it was finally removed from X.²

Just a couple of months earlier, around October 2023, Indonesian cyberspace was similarly gripped by a video depicting President Joko "Jokowi" Widodo delivering a speech in fluent Mandarin. At a quick glance, the video seemed visually normal. However, the Ministry of Communication and Information Technology (Kemenkominfo) has confirmed that the video was artificially manipulated.³ Before it was confirmed as disinformation, many people believed that the video was real. In some cases, this also raised certain political sentiments against President Jokowi.⁴



Both cases involve the same problem: deepfake technology. Deepfake refers to synthetic media content comprising digitally manipulated images, videos, or audio, designed to emulate real individuals engaging in actions or utterances that never occurred in reality.⁵ The term deepfake derives from two words, "deep learning" and "fake". The term first used in 2017 when a *Reddit* user under the pseudonym Deepfakes uploaded a number of pornographic videos showing famous actresses such as Scarlett Johansson and Gal Gadot as porn stars. Since then, the term has evolved into a catch-all phrase denoting manipulated content generated by artificial intelligence (AI). As AI technology advances, deepfake technology increasingly blurs the line between authentic and manipulated content, rendering discernment more challenging.⁶

Undoubtedly, the proliferation of deepfake technology presents a new threat to society. In February 2024, a finance worker in Hong Kong fell victim to a scam orchestrated through deepfake manipulation. He sent US\$25 million to a scammer who used deepfake technology to impersonate the company's chief financial officer in a video conference.⁷ Such cases underscore the utilization of deepfake as a potent social engineering tool for perpetrating financial scams and fraud. Meanwhile, in the cases of non-consensual pornography spread against Taylor Swift and Scarlett Johansson, deepfake was used as a medium to destroy a person's reputation, image, and career.⁸ Other threats posed by deepfake include disinformation campaigns, data theft, the creation of negative sentiments, and threats to democracy.

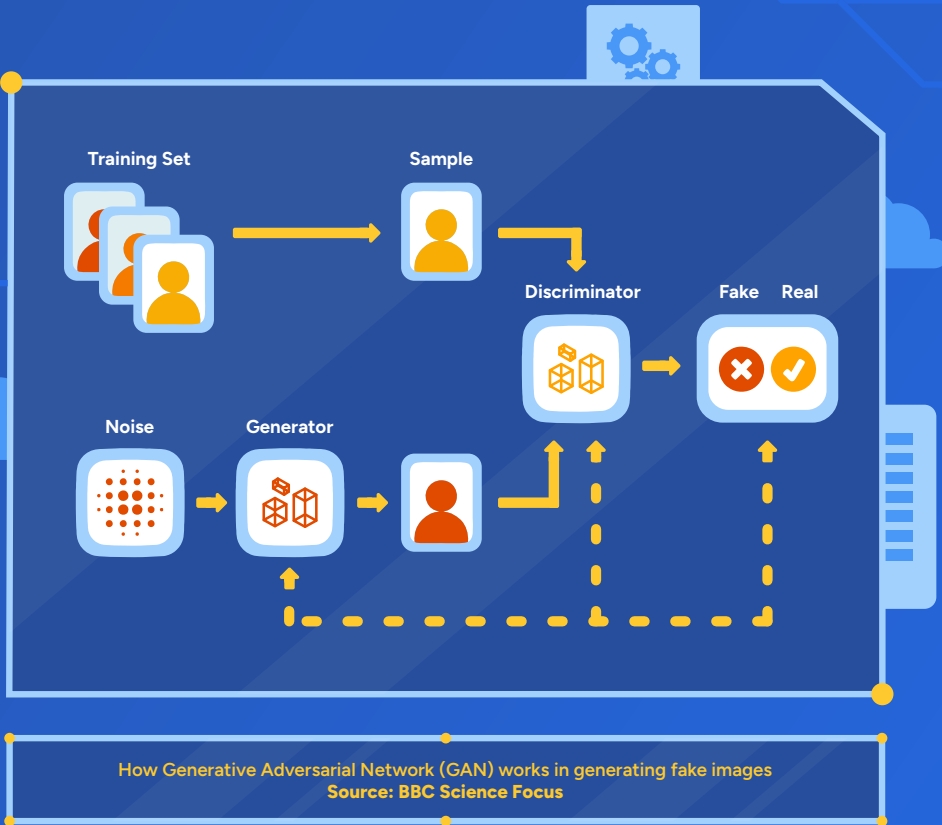
The complexities surrounding deepfake issues necessitate a legal response to mitigate their impacts. Law holds considerable potential in regulating the use of deepfake and addressing the multifaceted problems it engenders. This paper aims to discuss how deepfake distorts the socio-political realities in society and how the law should play a role in mitigating and tackling the harm deepfake poses. Comprising four sections, this paper first sets the stage with an introduction. After the introduction, the second section discusses how deepfake distorts the socio-political realities in society through disinformation and non-consensual pornography. In the third section, the paper delves into three key approaches to mitigating deepfake harm: technical, educational, and legal, with a particular emphasis on the urgency of legal intervention. The last section provides the conclusion of this paper.



Deepfake, disinformation, and non-consensual pornography: How AI distorts socio-political realities in society

Before discussing the correlation between deepfake, disinformation, and non-consensual pornography, this section will explain how deepfake works. Deepfake is generated by a machine learning (ML) algorithm called General Adversarial Networks (GANs). GANs operate through an adversarial process wherein two neural networks, namely the generator and the discriminator, collaborate to synthesize false data. The generator is a neural network that generates false data based on the input dataset. Meanwhile, the discriminator is a neural network that distinguishes between the false data produced by the generator and the original data.⁹

The generator operates by producing synthetic data closely resembling the original dataset provided as input. Subsequently, this falsified data undergoes scrutiny by the discriminator, which serves to detect discrepancies between original and fabricated content. Upon detecting inconsistencies in the generated data, the discriminator provides feedback to the generator, prompting refinement towards the creation of more authentic content. At this point, the generator and discriminator will compete with each other; this constitutes an adversarial framework. This happens repeatedly until the generator generates false data that looks so authentic and convincing that the discriminator is unable to distinguish it from authentic data. The false data that the discriminator cannot distinguish is called a deepfake.¹⁰



Depending on the intended content, creators utilize various types of deepfakes, which encompass photo, audio, and video formats. Firstly, photo deepfakes facilitate the substitution of an individual's facial features or other anatomical elements with those of another person. Secondly, audio deepfakes enable the alteration or emulation of a person's voice, often achieved through text-to-speech techniques. Thirdly, video deepfakes manipulate visual and auditory elements, permitting the replacement of a person's visage within a video with that of another individual, along with the synchronization of lip movements and dialogue alteration.¹¹ Taylor Swift's deepfake pornography falls under the first type of deepfake, while President Jokowi's deepfake speaking Mandarin is the third type.

In the age of AI, the rise of deepfake creates tremendous distortions to the socio-political realities in society. The utilization of AI algorithms to generate deepfake content paves the way for the new era of disinformation. This is chiefly attributable to one reason: the easier and wider access to AI. While in the past, the creation of manipulated content was only feasible for certain people who mastered image and video editing software, contemporary advancements afford virtually anyone the capability to manipulate images or videos through AI-based applications. As a result, manipulated content will be much easier to create and its number will increase compared to previous eras. Thus, a new era of disinformation begins where anyone can fake content and disseminate it to the public with a narrative as if it were authentic. Furthermore, the democratization of social media also opens up a wider path for deepfake-based disinformation to reach a wider and more diverse audience.¹²

Moreover, the heightened realism of deepfake-generated content exacerbates the distortion of societal perceptions of reality. Individuals encounter challenges in discerning between manipulated and authentic content as the boundaries are increasingly blurred. Consequently, this phenomenon fosters the propagation of erroneous beliefs, wherein falsehoods are embraced as truths.¹³ Such ramifications are particularly pronounced within communities characterized by low levels of digital literacy. In the absence of adept discernment skills to distinguish manipulated from original content, individuals are inclined to regard deepfake material as authentic.



Furthermore, a study has uncovered that even among individuals proficient at identifying deepfake content, its presence still yields adverse implications. These implications materialize as internal skepticism regarding the authenticity of such content, correlating with their trust in news disseminated via social media platforms. Subsequently, this doubt will lead to a decline in their trust towards news on social media, particularly given the prevalent dissemination of deepfake content within these platforms.¹⁴



In the political context, deepfake presents a formidable threat to electoral politics and democracy at large. Research findings underscore that deepfake video content featuring a political figure articulating specific statements has the potential to sway public perceptions and political inclinations towards said figure.¹⁵ Within this framework, deepfake serves as a tool for orchestrating the manipulation of public opinion, thereby instigating particular sentiments directed towards certain individuals. For instance, in the case of the aforementioned deepfake video involving President Jokowi, sentiments emerged suggesting a pro-China stance on the part of President Jokowi.



In addition to the societal and political realities distorted by the dissemination of disinformation, deepfake also poses a significant threat to individuals' personal and social life through the proliferation of manipulated pornography content or deepfake pornography. Statistics reveal that approximately 95 percent of all circulating deepfake video content comprises pornographic material. Deepfake pornography predominantly targets celebrities or public figures whose images or videos are publicly accessible.¹⁶ However, with the increasing prevalence of social media usage within society, ordinary individuals are increasingly susceptible to becoming victims. Perpetrators typically exploit publicly available images of victims, manipulating them into pornographic content.¹⁷

While male victims of deepfake pornography exist, the vast majority of those targeted are women. Through deepfake pornography, women are objectified and exploited to fulfill the creator's fantasies. In mid-2023, the issue of deepfake pornography garnered significant attention on social media

platform X, following a tweet by user @itsindahg regarding the circulation of online

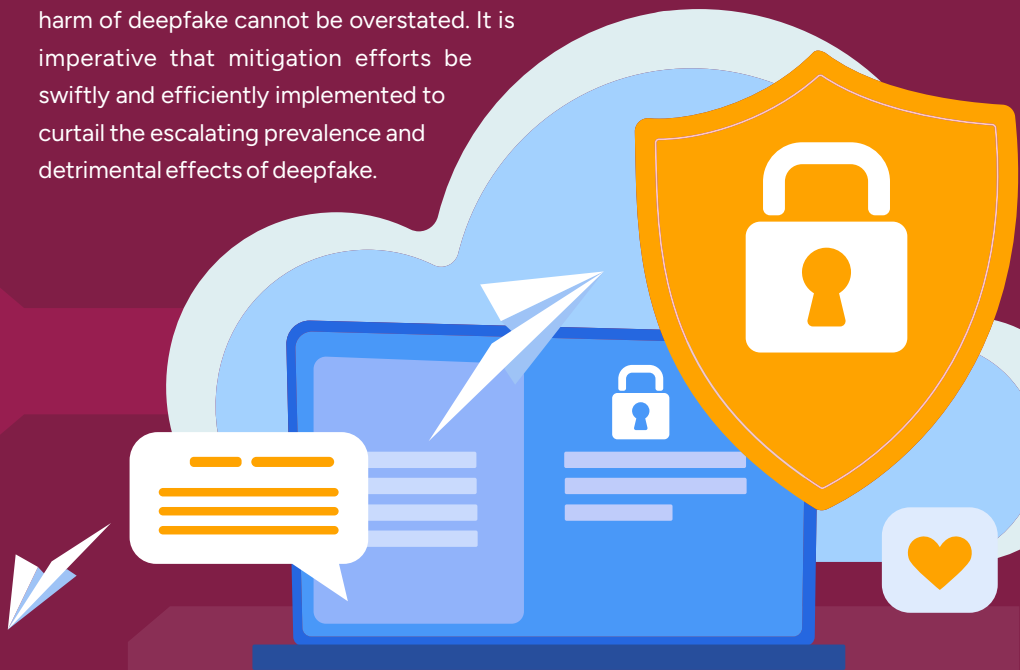
forms associated with deepfake pornography.

These forms originated from an online chat group on Telegram known for disseminating sexually abusive content, including deepfake pornography.¹⁸ Consequently, this development has sparked considerable apprehension, particularly among women, as the digital world is no longer perceived as a safe space, even in the absence of any proactive actions on their part.



The dissemination of deepfake pornography constitutes a grave intrusion upon an individual's privacy and sexuality, yielding profound psychological and social impacts for its victims. Psychologically, individuals subjected to deepfake pornography endure significant distress and social stigma as such illicit content circulates across social media platforms. Regrettably, victims often face societal blame (victim blaming) for the propagation of deepfake pornography content, despite their likely unawareness of its existence. This pervasive social pressure exacerbates the psychological strain on victims, precipitating symptoms of depression and, tragically, in some cases, culminating in suicide.¹⁹

Socially, the proliferation of deepfake pornography content causes a devastating toll on the image, reputation, integrity, and professional prospects of its victims. Moreover, individuals targeted by deepfake pornography encounter significant difficulty in securing employment opportunities, as prospective employers tend to conduct online assessments of their candidates. Upon encountering pornographic material associated with their candidates, employers are inclined to disqualify them from consideration, further exacerbating the adverse repercussions of deepfake victimization.²⁰ Consequently, the imperative for prompt and effective mitigation measures against the harm of deepfake cannot be overstated. It is imperative that mitigation efforts be swiftly and efficiently implemented to curtail the escalating prevalence and detrimental effects of deepfake.



Three key approaches in mitigating and tackling the harm of deepfake

Thus far, three approaches have emerged in addressing and mitigating the harm of deepfake technology: technical, educational, and legal approaches. The technical approach involves leveraging advancements in technology to detect and counter deepfake content. Computer experts have devised various techniques for identifying deepfake content. One such method, proposed by Li and Lyu in 2019, entails artifact-based detection. At the time, deepfake algorithms were limited in their ability to produce high-resolution images, necessitating further transformation to closely resemble the original face intended for replacement in the authentic video. This transformation process leaves traces within the resulting deepfake content, which can be effectively captured by Convolutional Neural Networks (CNNs). These traces serve as conclusive evidence of the presence of deepfake manipulation.²¹

Another approach, inconsistency-based detection techniques, has been developed to identify anomalies within deepfake content. Various studies have uncovered indicators of video deepfakes, such as unnatural blink patterns, facial distortion, and overall inconsistencies in the video imagery.²² Additionally, several other detection techniques have been devised, including forensic-based detection, Deep Neural Network (DNN)-based detection, GAN Redesign-Based Detection, and so on.²³ Despite its significant role in mitigating the threats posed by deepfake, the technical approach is inherently limited and may only offer temporary solutions. This limitation arises from the escalating sophistication of detection techniques, which often prompts corresponding advancements in deepfake technology, perpetuating a perpetual cycle of cat-and-mouse between deepfake creation and detection methods with no definitive resolution in sight.²⁴

The second approach entails an educational initiative aimed at enhancing media and digital literacy within society to enable the discernment of deepfake content from authentic one. This educational endeavor involves imparting knowledge to the public regarding the nature of deepfake content, its creation processes, associated risks, and techniques for distinguishing between deepfake and original content. Equipping individuals with the knowledge and skills to identify deepfake content diminishes the likelihood of its dissemination, thereby mitigating the adverse impacts of disinformation.²⁵

Nevertheless, akin to the technical approach, the educational strategy aimed at enhancing media literacy and digital proficiency also harbors limitations. Achieving genuine success and impact through this approach necessitates a protracted and resource-intensive process. The extended timeframe required for widespread adoption and proficiency in media literacy is incongruent with the rapid advancements in technology, which continually refine deepfake generation techniques to produce content that defies detection by human senses. Consequently, the efficacy of knowledge and skills in identifying deepfake content may diminish over time, rendering them increasingly irrelevant in the face of evolving deepfake capabilities.



The third is the legal approach, leveraging the law as a tool of social engineering to govern the utilization of deepfake. Typically, two legal approaches are employed in regulating technology: the principle-based approach, encompassing soft law, and the rule-based approach, involving hard law. In the context of AI usage, the former approach often manifests through the formulation of ethical guidelines governing AI deployment. This strategy prioritizes the establishment of standards to be upheld by regulated entities, emphasizing expected outcomes. However, the interpretation and execution of these standards are delegated to individual entities subject to regulation.²⁶

On one hand, the principle-based approach offers flexibility, ensuring continued relevance amidst the rapid evolution of technology. However, this approach is not without its limitations, rendering it unsuitable for effectively addressing the threat posed by deepfake technology. This kind of approach does not have a strong control mechanism to ensure that the regulated subjects will meet the established standards. When applied to the deepfake issue, principles such as "Do No Harm" may fall short in ensuring that individuals refrain from exploiting deepfake technology to harm others.



The second legal approach is the rule-based approach, implemented through hard law. This approach prioritizes the establishment of prescriptive rules, emphasizing legal certainty. Typically, this approach entails the enactment of regulations containing specific provisions that either mandate or prohibit actions by regulated entities. To enforce compliance among regulated entities, this approach often includes a range of sanctions—both administrative and criminal—for violations of these provisions.²⁷

This approach offers the advantage of a robust control mechanism to ensure compliance with regulatory provisions, particularly when accompanied by sanctions. However, its rigid nature renders it less adaptable to the rapid pace of technological advancements, presenting certain drawbacks.²⁸ Nonetheless, rule-based regulations can be formulated to effectively govern deepfake technology while ensuring its relevance for the future.



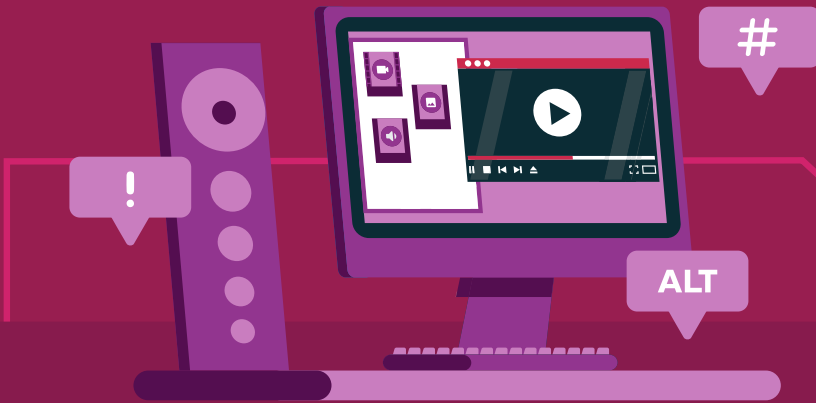
In the Indonesian context, several regulations are pertinent to the use of deepfake, as delineated in the preceding section. These include Law Number 27 of 2022 on Personal Data Protection (PDP Law), Law Number 12 of 2022 on Sexual Violence Crimes (TPKS Law), Law Number 11 of 2008 on Information and Electronic Transactions (EIT Law) and its amendments, Law Number 44 of 2008 on Pornography, and Law Number 1 of 2023 on the Criminal Code, which will come into effect in 2026.

However, the implementation of the aforementioned regulations primarily focuses on remedial measures when addressing the deepfake issue. In essence, these regulations are invoked only in response to specific instances of adverse deepfake occurrences. For instance, under the PDP Law, there exists a prohibition on the falsification of personal data leading to harm to others, carrying a maximum penalty of six years imprisonment and/or a fine of up to six billion rupiah.²⁹ Similarly, concerning instances where deepfake pornography intersects with sexual violence, the TPKS Law stipulates that perpetrators engaging in sexual violence via electronic means will face criminal sanctions, augmented by a third of the initial punishment.³⁰ Both of these regulations react to deepfake incidents only after harm has been inflicted upon victims. Notably, there is currently no legislation in Indonesia that prescribes mechanisms for preemptively addressing the widespread impact of deepfake and mitigating it from its inception.



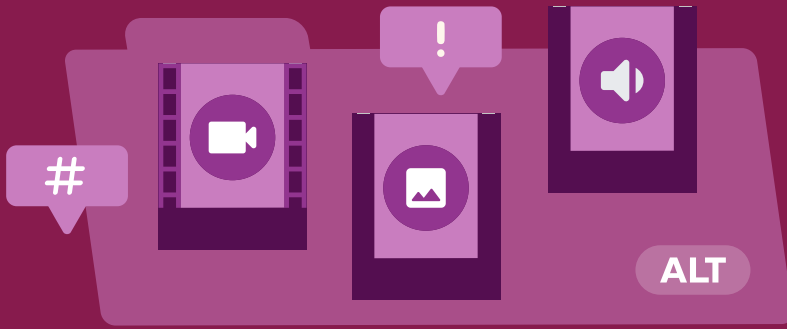
In several countries, a widely adopted legal mechanism for regulating deepfake content is the implementation of disclosure policy. This policy constitutes a facet of the transparency principles within AI ethics frameworks. Within the context of deepfake technology, numerous countries mandate the application of disclosure policy, requiring that deepfake content be explicitly labeled as artificially manipulated by AI. This regulatory approach was recently embraced by the European Union through the EU AI Act, which received approval from the European Union Parliament on March 13, 2024.³¹

Article 50, paragraph (4) of the EU AI Act outlines provisions mandating AI system deployers generating manipulated images, audio, or video to disclose the artificial nature of the content. This means that the AI system deployers have the obligation to disclose that the content has been artificially generated or manipulated.³² Furthermore, Recital 134 elucidates that this disclosure obligation entails labeling deepfake content and revealing the original content that underwent manipulation.³³ China, a pioneering nation in AI development, implemented a similar policy predating the EU AI Act. This policy prohibits the dissemination of deepfake content without proper disclosure of its manipulated nature.³⁴ Meanwhile, in the United States, draft legislation addressing deepfake was proposed to Congress in 2023, featuring provisions requiring disclosure in line with transparency principles.³⁵



The implementation of this disclosure policy serves as an effective control mechanism for proactively mitigating the impact of deepfake technology. By requiring applications generating deepfake content to provide clear labels on manipulated content, individuals encountering such content will be informed of its artificial nature, distinguishing it from original content. Consequently, this measure significantly reduces the potential ramifications associated with deepfake dissemination, particularly the propagation of widespread disinformation. Such a mechanism warrants consideration for adoption in Indonesia, necessitating the labeling of deepfake content to safeguard individuals against deception and discourage the sharing of falsified content, thereby minimizing its broader societal impact.

In tackling the issue of deepfake pornography, the legal framework must play a role in combating the proliferation of manipulative pornographic content disseminated via social media platforms. Drawing insights from the delayed response observed in the Taylor Swift incident on platform X, it is imperative for the law to enforce accountability measures to safeguard the online environment for all users. Granting platforms immunity from accountability only exacerbates the prevalence of deepfake pornography victimization. One example of these accountability measures is through mandating the social media platforms the prompt take down of deepfake pornography content upon detection. Additionally, the legal system should establish a clear and accessible mechanism for victims to seek legal redress against platforms that fail to expeditiously remove deepfake pornography content, which may cause harm to them.³⁶



Conclusion

Cases involving deepfake highlight the dual nature of technology, showcasing not only its benefits but also its potential harms. Deepfake represents a significant distortion of socio-political realities in the contemporary era, particularly through its utilization as a tool for spreading disinformation. Moreover, deepfake poses a threat to human safety and dignity when employed to disseminate non-consensual pornography. Given the multifaceted nature of the deepfake problem, it necessitates comprehensive problem-solving approaches to mitigate its impacts and prevent further detriment.

Among the three approaches available to mitigate and address the impact of deepfake—technical, educational, and legal—a synergistic combination of these approaches would undoubtedly present an ideal problem-solving strategy. However, considering the urgent nature of the issues at hand, the legal approach emerges as the most fitting solution, given its capacity as a tool of social engineering and its robust control mechanism. This legal approach can be implemented through the enactment of disclosure policy mandating AI system deployers to label deepfake content and imposing platform accountability, particularly in the context of deepfake pornography. By implementing disclosure policy and holding platforms accountable, the more detrimental impacts of deepfake can be mitigated and addressed.

References

- ¹ Witt, A. and Henry, N. (2024). 'Taylor Swift deepfakes: new technologies have long been weaponised against women. The solution involves us all' Available at: https://theconversation.com/taylor-swift-deepfakes-new-technologies-have-long-been-weaponised-against-women-the-solution-involves-us-all_222268 (Accessed: March ¹², 2024).
- ² Saner, E. (2024) 'Inside the Taylor Swift Deepfake Scandal: "It's Men Telling a Powerful Woman to Get Back in Her Box"' Available at: www.theguardian.com/technology/2024/jan/31/inside-the-taylor-swift-deepfake-scandal-its-men-telling-a-powerful-woman-to-get-back-in-her-box (Accessed: March ¹², 2024).
- ³ Afifa, L. (2023) 'Kominfo Ministry Asserts Jokowi's Speech in Chinese Produced by Deepfake' Available at: en.tempo.co/read/1789306/kominfo-ministry-asserts-jokowis-speech-in-chinese-produced-by-deepfake (Accessed: March ¹², 2024).
- ⁴ Dewi, I.R. (2023) 'Video Pidato Mandarin Jokowi Telan Korban, Kominfo Bongkar' Available at: www.cnbcindonesia.com/tech/20231208094512_37_495581/video-pidato-mandarin-jokowi-telan-korban-kominfo-bongkar (Accessed: March ¹², 2024).
- ⁵ Mustak, M. et al. (2023) 'Deepfakes: Deceptions, Mitigations, and Opportunities', *Journal of Business Research*, ¹⁵⁴, pp. ¹⁻¹⁵. doi:<https://doi.org/10.1016/j.jbusres.2022.113368>.
- ⁶ Dasilva, J.P., Ayerdi, K.M., and Galdospin, T.M. (2021) 'Deepfakes on Twitter: Which Actors Control Their Spread?', *Media and Communication*, ⁹(1), pp. ³⁰¹⁻³¹². <https://doi.org/10.17645/mac.v9i1.3433>.
- ⁷ Chen, H. and Magramo, K. (2024) 'Finance worker pays out \$²⁵ million after video call with deepfake 'chief financial officer' Available at: <https://www.google.com/amp/s/amp.cnn.com/cnn/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/index.html> (Accessed: March ¹², 2024).

- ⁸ Pascale, E. (2023) 'Deeply Dehumanizing, Degrading, and Violating: Deepfake Pornography and the Path to Legal Recourse', *Syracuse Law Review*, 73, pp. 335_366.
- ⁹ Yadav, D. and Salmani, S. (2019) 'Deepfake: A Survey on Facial Forgery Technique Using Generative Adversarial Network', *Proceedings of the International Conference on Intelligent Computing and Control Systems*, pp. 852_857.
- ¹⁰ Ibid.
- ¹¹ Masood, M. et al. (2023) 'Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward', *Applied Intelligence*, 53, pp. 3974_4026. <https://doi.org/10.1007/s10489-022-03766-z>.
- ¹² Gosse, C. and Burkell, J. (2020) 'Politics and porn: how news media characterizes problems presented by deepfakes', *Critical Studies in Media Communication*, 37(5), pp. 497_511. <https://doi.org/10.1080/15295036.2020.1832697>.
- ¹³ Ibid.
- ¹⁴ Vaccari, C. and Chadwick, A. (2020) 'Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News', *Social Media + Society*, 6(1), pp. 1_13. doi:10.1177/2056305120903408.
- ¹⁵ Ray, A. (2021) 'Disinformation, Deepfakes and Democracies: The Need for Legislative Reform', *UNSW Law Journal*, 44(3), pp. 983_1013. <https://doi.org/10.53637/DELS2700>.
- ¹⁶ Story, D. and Jenkins, R. (2023) 'Deepfake Pornography and the Ethics of Non-Veridical Representations', *Philosophy & Technology*, 56, pp. 36_56. <https://doi.org/10.1007/s13347-023-00657-0>.
- ¹⁷ Solichah, I.I., Sulistio, F., and Istiqomah, M. (2023) 'Protection of Victims of Deep Fake Pornography in a Legal Perspective in Indonesia', *International Journal of Multicultural and Multireligious Understanding*, 10(1), pp. 383_390. <http://dx.doi.org/10.18415/ijmmu.v10i1.4409>.
- ¹⁸ Pertiwi, S.P. (2023) 'Waspada KBGO, Ada Grup Obrolan Bikin Deepfake Porn via Formulir Daring' Available at: <https://www.konde.co/2023/07/darurat-kbgo-ada-temuan-group-chat-dan-permintaan-deepfake-porn-lewat-google-form/> (Accessed: March 21, 2024).

¹⁹ Pascale, E. (2023) 'Deeply Dehumanizing, Degrading, and Violating: Deepfake Pornography and the Path to Legal Recourse', *Syracuse Law Review*, ⁷³, pp. ^{335_366}.

²⁰ Ibid.

²¹ Li, Y. and Lyu, S. (2019) 'Exposing DeepFake Videos By Detecting Face Warping Artifacts', *CVPR Workshop*, pp. ^{46_52}.

²² Engler, A. (2019) 'Fighting deepfakes when detection fails' Available at: <https://www.brookings.edu/articles/fighting-deepfakes-when-detection-fails/> (Accessed: March ¹⁴, 2024).

²³ Zobaed, S. et al. (2021) 'DeepFakes: Detecting Forged and Synthetic Media Content Using Machine Learning', in Montasari, R. and Jahankhani, H. (eds.) *Artificial Intelligence in Cyber Security: Impact and Implications. Advanced Sciences and Technologies for Security Applications*. Cham: Springer. https://doi.org/10.1007/978_3_030_88040_8_7.

²⁴ Farid, H. (2022) 'Creating, Using, Misusing, and Detecting Deep Fakes', *Journal of Online Trust and Safety*, ¹(⁴), pp. ^{1_33}. <https://doi.org/10.54501/jots.v1i4.56>.

²⁵ Hwang Y., Ryu J.Y., and Jeong S.H. (2021) 'Effects of Disinformation Using Deepfake: The Protective Effect of Media Literacy Education', *Cyberpsychology, Behavior, and Social Networking*, ²⁴(³), pp. ^{188_193}. doi: ^{10.1089/cyber.2020.0174}.

²⁶ Black, J., Hopper, M., and Band, C. (2007) 'Making a success of Principle-based regulation', *Law and Financial Markets Review*, ¹(³), pp. ^{191_206}. <https://doi.org/10.1080/17521440.2007.11427879>.

²⁷ Decker, C. (2018) 'Goals-based and Rules-based Approaches to Regulation', *BEIS Research Paper*, ⁸, pp. ^{1_67}. <https://ssrn.com/abstract=3717739>.

²⁸ Ibid.

²⁹ Law Number ²⁷ of ²⁰²² on Personal Data Protection, Article ⁶⁶ jo. Article ⁶⁸.

³⁰ Law Number ¹² of ²⁰²² on Crimes of Sexual Violence, Article ¹⁵ paragraph (¹) point I.

³¹ Gilchrist, K. and Lordache, R. (2024) 'World's first major act to regulate AI passed by European lawmakers' Available at: <https://www.cnn.com/2024/03/13/european-lawmakers-endorse-worlds-first-major-act-to-regulate-ai.html> (Accessed: March 17, 2024).

³² EU Artificial Intelligence Act, Article ⁵⁰, paragraph (4).

³³ EU Artificial Intelligence Act, Recital ¹³⁴.

³⁴ Statt, N. (2019) 'China makes it a criminal offense to publish deepfakes or fake news without disclosure' Available at: <https://www.theverge.com/2019/11/29/20988363/china-deepfakes-ban-internet-rules-fake-news-disclosure-virtual-reality> (Accessed: March 17, 2024).

³⁵ Ramstack, T. (2023) 'Congress Wants Laws to Confront Deepfakes That Deceive Consumers' Available at: <https://www.thewellnews.com/technology/congress-wants-laws-to-confront-deepfakes-that-deceive-consumers/> (Accessed: March 17, 2024).

³⁶ Gieseke, A.P. (2020) "The New Weapon of Choice": Law's Current Inability to Properly Address Deepfake Pornography', *Vanderbilt Law Review*, ⁷³(5), pp. ^{1479_1516}.





Center for Digital Society

Fakultas Ilmu Sosial dan Ilmu Politik
Universitas Gadjah Mada
Gedung BC Lt. 2, Ruang BC 201-203
Jalan Socio Yustisia 1
Bulaksumur, Yogyakarta, 55281, Indonesia

Telepon: (0274) 563362, Ext. 116
Email: cfds.fisipol@ugm.ac.id
Situs Web: cfds.fisipol.ugm.ac.id



facebook.com/cfdsugm



Center for Digital Society (CfDS)



[cfds_ugm](https://www.instagram.com/cfds_ugm)



[@cfds_ugm](https://twitter.com/cfds_ugm)



[@cfds_ugm](https://x.com/cfds_ugm)



[CfDS UGM](https://www.youtube.com/CfDSUGM)