Deepfakes and its social impacts: Proposed modalities of governance from the approach of Lessig's Cyberspace Regulatory Theory

Keywords: *Deepfakes, Lessig's Theory, social impact, regulatory governance, disinformation*

Abstract:
As Elvis Presley once sang:

"Just by seeing is believing

I don't need to question why"[1]

For some people, even if something is strange, if you witness it for yourself, you will think it is true.[2] Nina Iacono Brown – an Associate Professor at Syracuse University specializing in the intersection of media law and technology,[3] pointed out a reason for such a belief that before the Internet era, people gathered information from traditional broadcast and print media, which was normally verified in advance by fact-checking procedures.[4] However, in the age of non-conventional news consumption (e.g., via online content sharing platforms) where the information sharing process often lacks verification steps, along with the emergence of new technology, "deepfake", it seems that "seeing is believing" became obsolete. Deepfake technology can "replace faces, manipulate facial expressions, synthesize faces, and synthesize speech" in media products, which can "make people appear to do and say things that they never did or said".[5]

Utilizing that function, malicious actors have negatively used deepfake to produce disinformation.[6] For instance, to interfere in the US 2016 presidential election, Russia injected numerous deepfake videos "spreading divisive and politically inflammatory

---

[1] Presley, Elvis. *Seeing Is Believing (Official Audio)*, 2021. https://www.youtube.com/watch?v=MMfhqgKXSDY [Accessed 30 May 2025].
[2] "Seeing Is Believing." In *Cambridge Dictionary*, n.d. https://dictionary.cambridge.org/vi/dictionary/english/seeing-is-believing [Accessed 30 May 2025].
[3] "Nina Iacono Brown." Syracuse University, n.d. https://newhouse.syr.edu/people/nina-brown [Accessed 30 May 2025]
[4] Brown, Nina I. "Deepfakes and the Weaponization of Disinformation." Virginia Journal of Law & Technology 23, no. 1 (2020): 16.
[5] Donald, Bernice B. and Hedges, Ronald J. "Deepfakes Bring New Privacy and Cybersecurity Concerns," Corporate Counsel Business Journal (2020). https://ccbjournal.com/articles/deepfakes-bring-new-privacy-and-cybersecurity-concerns [Accessed 30 May 2025].
[6] See definitions of disinformation in House of Commons Digital, Culture, Media and Sport Committee (DCMSC). "Disinformation and 'Fake News': Interim Report." United Kingdom, 2018: 10; Independent High-Level Expert Group on Fake News and Online Disinformation (HLEG). "A Multi-Dimensional Approach to Disinformation." European Commission, 2018: 10.

messages" into social media,[7] which attracted millions of viewers to share and comment. Another warning situation is that with the advent of deepfakes, identifying such "digital manipulations of audio or video" by our natural eyes has become increasingly challenging. Thus, if people still embrace the belief that "seeing is believing", they may endanger themselves and society.

Notwithstanding, from a positive perspective, there are considerable advantages of deepfakes that cannot be denied. Some reputable examples can be seen in a deepfake experimental project of the Korean television channel MBN to produce television programs more vividly without the participation of human newsreaders[8] or the application of deepfake in modifying videos of historical figures to educate history livelier for children. These benefits make matters more complicated and pose another challenge for national authorities: any solutions against deepfake should not be too restrictive but must strike a proper balance between protecting the community and enjoying technological advancements.

Internationally, concerning negative uses of deepfakes, several countries have conducted legal reforms to counter the chilling effects of this technology. However, it seems that the authorities still "hesitate to legislate".[9] It can be evidenced that, for example, though the UK government has prepared for the criminalisation of the deepfakes use for sexual abuse, the measure in the proposed Crime and Policing Bill was assessed as "not fully protect victims from the broader harm posed by technology-facilitated abuse".[10] In Southeast Asia, for instance, in Thailand and Indonesia, specific regulations tackling deepfakes are also missing.[11] A justification for such acts is that, due to the "grey" characteristic of deepfakes, legislation may cause a "disaster for online expression".[12] Moreover, given the ambiguity

---

[7] Zorn, Eric. "Polls Reveal Sobering Extent of Nation's Fact Crisis." *Chicago Tribune*, 2017. https://www.chicagotribune.com/columns/eric-zorn/ct-polling-ignorance-facts-trump-zorn-perspec-0106-md-20170105-column.html [Accessed 8 June 2022]; Chesney, Robert and Citron, Danielle. "Deepfakes and the New Disinformation War The Coming Age of Post Truth Geopolitics." *Foreign Affairs*, 2019. https://search.ebscohost.com/login.aspx?direct=true&AuthType=sso&db=edswss&AN=000452813200013&site=edslive&scope=site [Accessed 7 June 2022]

[8] PR Newswire. "Moneybrain to Supply Kim Joo-Ha AI Anchor Solution to MBN," 2020. https://www.prnewswire.com/news-releases/moneybrain-to-supply-kim-joo-ha-ai-anchor-solution-to-mbn-301169202.html [Accessed 10 June 2022].

[9] Anderson, Martin. "European and UK Deepfake Regulation Proposals Are Surprisingly Limited." *Unite AI*, 2022. https://www.unite.ai/european-and-uk-deepfake-regulation-proposals-are-surprisingly-limited/ [Accessed 10 July 2022].

[10] Queen Mary University of London, 'Deepfakes and the Law: Why Britain needs stronger protections against technology-facilitated abuse', 23 January 2025, https://www.qmul.ac.uk/media/news/2025/humanities-and-social-sciences/hss/deepfakes-and-the-law-why-britain-needs-stronger-protections-against-technology-facilitated-abuse.html, accessed 30 May 2025.

[11] Rouse, 'AI-generated deepfakes: what does the law says?', 4 September 2024, https://rouse.com/insights/news/2024/ai-generated-deepfakes-what-does-the-law-say, accessed 30 May 2025.

[12] Centre for Data Ethics and Innovation (CDEI). "Snapshot Paper - Deepfakes and Audiovisual Disinformation." Ethical Issues in AI. Department for Digital, Culture, Media & Sport, 2019. https://www.gov.uk/government/publications/cdei-publishes-its-first-series-of-three-snapshot-papers-ethical-issues-in-ai/snapshot-paper-deepfakes-and-audiovisual-disinformation#what-should-we-do-about- deepfakes [Accessed 10 July 2022].

surrounding different interventions, the governments should be careful not to "squeeze out beneficial use" of deepfakes.

In Vietnam, deepfakes are currently governed only from the perspective of criminal offenses. However, the use of deepfakes to produce disinformation remains unregulated. For example, using AI-powered deepfake tools to generate a video simulating a human face with realistic visuals and voice is not yet subject to any legal sanctions.

Regarding the criminal aspects, deepfakes deepfakes have emerged as a tool for bypassing authentication systems and facilitating illegal transactions by criminal groups engaged in fraud and extortion.[13] Additionally, deepfakes have been misused to superimpose victims' faces onto sensitive or explicit videos, creating highly realistic fabricated content that is then used to blackmail victims—threatening to release such materials publicly unless they remain silent or pay a ransom.

According to a 2024 report by the Global Initiative, the Asia-Pacific region witnessed a 1.53% increase in deepfake-related cases between 2022 and 2023. Vietnam recorded the highest rate of deepfake fraud in the region, accounting for 25.3% of cases.[14] Despite this alarming trend, Vietnam currently lacks a specific legal instrument that directly governs deepfakes. Presently, deepfakes' victims are indirectly regulated through provisions related to the protection of individual rights and the prohibition of spreading false information, as set forth in the Law on Cybersecurity 2018 and Decree No. 13/2023/ND-CP on the Protection of Personal Data dated 17 April 2023. Moreover, the use of deepfakes in cybercrimes such as phishing or malware attacks is criminalized under Vietnam's Criminal Code 2015.

Clearly, deepfakes are a double-edged sword, generating significant social implications. In light of the current context in the UK and the US, this study raises two essential research questions:

(i) What is the nature of deepfake and how does it impact our human life?
(ii) Should deepfakes be governed and to what extent?
(iii) If so, which modalities should Vietnam consider addressing disinformation problems arising from deepfakes?

This paper concludes that regulating deepfakes is necessary due to their complex nature. However, any regulatory framework should carefully balance the risks and the potential benefits that deepfakes may offer, in order not to hinder technological innovation and community advantages.

Accordingly, the paper proposes a two-pronged approach for governing deepfakes in

---

[13] 'Vietnamese authorities warn of rising AI-driven deepfake extortion scams' (Việt Nam News, 1 April 2025) https://vietnamnews.vn/society/1694933/vietnamese-authorities-warn-of-rising-ai-driven-deepfake-extortion-scams.html accessed 30 May 2025.

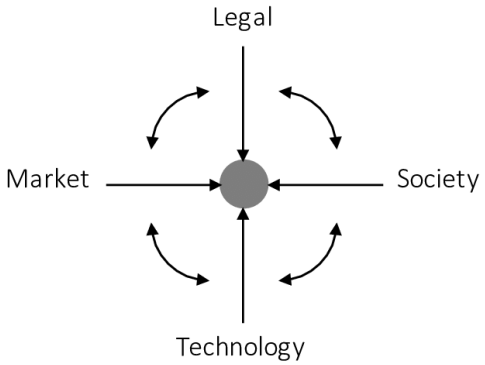
[14] Natnicha Surasit, 'Rogue Replicants: Criminal Exploitation of Deepfakes in South East Asia' (Global Initiative Against Transnational Organized Crime, 29 February 2024) https://globalinitiative.net/analysis/deepfakes-ai-cyber-scam-south-east-asia-organized-crime/ accessed 30 May 2025.

Vietnam:

(i) **Top-down approach**: Update existing legal instruments to specifically address disinformation involving the malicious use of deepfakes that negatively impact society.

(ii) **Bottom-up approach**: Promote the development of soft law instruments and guidelines to encourage responsible practices among businesses in Vietnam—particularly those engaged in data collection and processing. These entities should establish internal codes of conduct to safeguard user data and maximize the positive applications of deepfakes. At the same time, a mandatory disclosure mechanism should be introduced, requiring such organizations to report suspected abuse of deepfake technology to competent state authorities.

Regarding the methodology, the paper will use doctrinal and analytical legal research, together with different arguments of scholars and policymakers for and against the regulations of deepfakes. Significantly, the paper consolidates the doctrinal legal research and philosophical method to illustrate the operation of Lessig's theory, particularly to highlight the importance of the interaction of four modalities (law, social norms, market and technology) in addressing problems of cyberspace.



Picture 1: Lessig's four modalities of regulation

Research findings

Deepfakes, powered by increasingly sophisticated technology, have heightened public anxiety by blurring the line between truth and falsehood. This uncertainty threatens individual reputations, public order, and even democratic processes. While academics and regulators are in the race to find appropriate solutions, it should be remembered that this innovation is not only used to pose hazards to people and society, but it can also improve our lives (e.g., by the contribution to media industry and children's education). Hence, this magnitude of potential negative effects of malicious use makes it more difficult to develop regulations because any remedies against deepfakes may be considered as restrictive to freedom of expression or unnecessary as the self-regulation mechanism and current laws can sufficiently handle arising issues of deepfakes.

(i) Limitations of the status quo

Under the research, the potential risks associated with the use of deepfake technology include: violations of privacy rights, damage to reputation and honor, risks of fraud and deception, and infringements of intellectual property rights.

Currently in Vietnam, the impacts of deepfake technology are governed through a number of key legal instruments, particularly those relating to personal data protection and the prohibition of unlawful personal data processing. Specifically:

- The Law on Cybersecurity 2018 prohibits acts of misusing cyberspace to falsify information or infringe upon the privacy, reputation, and honor of individuals and organizations. Accordingly, the use of deepfake technology to generate fake videos or images that attack the reputation or dignity of others falls under the prohibited acts set out in Article 8 of this law.
- The Law on Information Technology 2006 (as amended in 2018) includes provisions on the protection of personal data against unauthorized collection, use, and dissemination, as stipulated in Article 12.
- In addition, the rights of data subjects have been explicitly recognized in Decree No. 13/2023/ND-CP dated April 17, 2023, on Personal Data Protection, alongside the Law on Electronic Transactions 2023, marking significant progress towards aligning Vietnam's data protection framework with international standards such as the GDPR.
- With respect to sanctions for the misuse of deepfakes, Decree No. 15/2020/ND-CP provides administrative penalties in areas such as postal services, telecommunications, and radio frequency use. Moreover, violations related to personal data may also constitute criminal offenses under Articles 159 and 288 of the 2015 Penal Code, covering crimes such as "Infringement upon the secrecy or safety of mail, telephone or other forms of private communication" and "Illegal provision or use of information on computer networks or telecommunications networks."

However, from the authors' perspective, current regulations remain incomplete, as Vietnamese law does not yet contain any specific provisions dedicated to deepfakes. In addition, platforms such as YouTube, Facebook, and TikTok have not yet assumed clear responsibility for identifying and managing deepfake videos. For instance, they have not implemented measures such as labeling content as AI-generated or limiting the circulation of deepfakes that convey misleading or false information.

Furthermore, while freedom of expression is a fundamental right, it is not absolute and may be lawfully restricted where such limitations satisfy the proportionality test. Therefore, there is no reason why deepfakes should not be regulated. The real challenge lies in determining the appropriate scope and extent of such regulation.

(ii) A multidimensional regulatory strategy

To address the legal, technical, market, and social dimensions of deepfakes, the paper adopts Lawrence Lessig's theory of "modalities of regulation" in cyberspace.

1. Legal constraints	Lawmakers should prioritise to regulate pornographic deepfakes. Their prevalence provides firmer empirical grounds for regulation that can withstand free-speech scrutiny. Statutes should categorically prohibit the creation and distribution of non-consensual pornographic deepfakes, irrespective of motive
2. Technological and market constraints	Regarding the constraint of technology and market, to avoid the risks of violating net neutrality rules and freedom of expression, rather than trying to slow down the speed of circulation of deepfake contents or applying the ex-ante detection techniques, this paper suggests that the regulators should compel the intermediaries to apply ex-post detection technologies in censoring their users' uploaded contents. For example, online intermediaries would be required to deploy post-upload detection tools, label verified deepfakes, and promptly remove unlabeled content when flagged through notice-and-takedown procedures. Penalties for non-compliance must be calibrated: severe enough to deter misconduct, yet not so harsh as to cripple platforms.
3. Social-norm constraints	Legislation should influence both internal dispositions and external behaviours. - *Intrinsic measures*: invest in media-literacy programmes. Taiwan's "nerd immunity" model—countering disinformation without censorship—offers a useful template. - *Extrinsic measures*: foster a pluralistic media environment so *that* citizens can access diverse viewpoints. Achieving genuine media pluralism is complex and warrants further study; the challenge is to broaden diversity without chilling economic activity or competition.