CADERNOS DE DEREITO ACTUAL
www.cadernosdedereitoactual.es

© ***Cadernos de Dereito Actual*** Nº 32. Núm. Ordinario (2026), pp. 177-200
·ISSN 2340-860X - ·ISSNe 2386-5229

Criminal liability for criminal use of deepfake technology that causes serious consequences

Viktoriia Shcherba[1]
National Academy of Internal Affairs
Mykhailo Akimov[2,*]
National Academy Internal Affairs
Inna Vartyletska[3]
National Academy of Internal Affairs
Olha Kryshevych[4]
National Academy of Internal Affairs
Mariia Diakur[5]
Yuriy Fedkovych Chernivtsi National University

Abstract: Generative AI has accelerated deepfake crime beyond the limits of traditional victim-centered criminal qualification. This study develops a dual-scale

[1] PhD in Law Sciences, Associate Professor of the Department of Criminal Law and Criminology, Educational and Scientific Institute of Law and Psychology, National Academy of Internal Affairs, Kyiv, Ukraine. ORCID: 0000-0002-8529-7979; E-mail: viktoriia.shcherba9@mail.com.
[2] PhD in Legal Sciences, Associate Professor of the Criminal Law Department, National Academy Internal Affairs, Kyiv, Ukraine. ORCID: 0000-0001-7715-0259; E-mail: mykhailo144medical@gmail.com (corresponding author).
[3] PhD in Legal Sciences, Professor of Criminal Law, National Academy of Internal Affairs, Kyiv, Ukraine. ORCID: 0000-0002-3447-0567; E-mail: inna.vartyletskaa2@gmail.com.
[4] Doctor of Juridical Sciences, Professor of Criminal Law, National Academy of Internal Affairs, Kyiv, Ukraine. ORCID: 0000-0001-6136-8106; E-mail: olha.kryshevych11@gmai.com.
[5] PhD, Associate Professor of the Department of Criminal Law, Faculty of Law, Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine. ORCID: 0000-0003-3265-6584; E-mail: mariia.dyakur21@chnu.edu.ua.

framework that links personal and systemic harm to criminal-judicial decision-making. Using classification, topology building, cross-impact analysis, judicial content analysis, regression, scenario modeling, and UML formalization, the study shows that automation, diffusion, and institutional targeting drive the most severe outcomes.

Keywords: Deepfake Crime Severity, Systemic Digital Harm, Synthetic Media Forensics, Algorithmic Legal Qualification, Criminal Justice Adaptation, Evidentiary Destabilization, Cybercrime Escalation

Resumo: A inteligência artificial generativa acelerou os crimes com deepfakes, ultrapassando os limites da qualificação criminal tradicional centrada na vítima. Este estudo desenvolve uma estrutura de dupla escala que liga o dano pessoal e sistémico à tomada de decisões criminais e judiciais. Utilizando a classificação, a construção de topologia, a análise de impacto cruzado, a análise de conteúdo judicial, a regressão, a modelação de cenários e a formalização UML, o estudo demonstra que a automatização, a difusão e o targeting institucional impulsionam os resultados mais graves.

Palavras-chave: Gravidade do Crime de Deepfake, Danos Digitais Sistémicos, Análise Forense de Meios Sintéticos, Qualificação Jurídica Algorítmica, Adaptação da Justiça Criminal, Desestabilização Probatória, Escalada do Cybercrime

1. Introduction

The study addresses the widening gap between the growth of digital threats and the ability of criminal justice systems to respond under conditions of societal informatization and hybrid security risk[6,7,8,9].There is also a growing need for technologically sensitive and procedurally robust mechanisms for qualifying harm from non-traditional harm vectors documented in current standards for investigating complex crimes[10]Deepfakes sit at the intersection of cybercrime, informational warfare, and evidentiary destabilization. For that reason, analytically grounded models for qualifying harm severity are both timely and operationally important.

Prior scholarship addressed governance, privacy, detection, and offense-specific doctrine, but it did not provide an integrated model for measuring systemic harm caused by deepfake crime. In judicial practice, the links between automation, diffusion, institutional targeting, and criminal severity remained weakly specified. Large-scale synthetic harm therefore stayed analytically visible but procedurally under-addressed.

---

[6] KOPOTUN, I., et al. "Expanding the potential of the preventive and law enforcement function of the security police in combating cybercrime in Ukraine and the EU", *TEM Journal*, v. 9, n. 2, 2020, p. 460–468. https://doi.org/10.18421/tem92-06

[7] BONDARENKO, S., et al. "Improving the state system of strategic planning of national security in the context of informatization of society", *Journal of Information Technology Management*, v. 14, 2022, p. 1-24. https://doi.org/10.22059/jitm.2022.88861

[8] KIRCHENGAST, T. "Deepfakes and image manipulation: Criminalisation and control", *Information & Communications Technology Law*, v. 29, n. 3, 2020, p. 308–323. https://doi.org/10.1080/13600834.2020.1794615

[9] RENAUD, L. "Will you believe it when you see it? How and why the press should prepare for deepfakes", *Georgetown Law Technology Review*, v. 4, 2019, 241. Available at: https://georgetownlawtechreview.org/will-you-believe-it-when-you-see-it-how-and-why-the-press-should-prepare-for-deepfakes/GLTR-01-2020/4 (accessed on 2 May 2025).

[10] LOSHYTSKYI, M., et al. "International legal standards for documentation and investigation of war crimes", *Clio Revista De Historia Ciencias Humanas Y Pensamiento Critico*, v. 5, n. 10, 2025, p. 1818–1855. https://doi.org/10.5281/zenodo.15598037

Criminal justice systems still assess deepfake offenses mainly through individualized harm and proximate causation. That approach underestimates cumulative cross-platform and institutional damage. The result is inconsistent qualification and fragmented responses to serious deepfake harm.

Research questions. Which technological and contextual factors influence deepfake harm severity beyond individual impact? To what extent does judicial reasoning recognize the systemic impact of deepfakes? Can personal and systemic harm be integrated within an algorithmic criminal qualification framework?

Research hypothesis. Deepfake crime severity can be modeled as a dual-scale structure of personal and systemic harm, with systemic variables explaining a larger share of severe outcomes than current judicial qualification models.

Research aim. The study aims to develop a dual-scale model to assess harm from deepfake-enabled crimes and align it with criminal justice decision-making.

Research objectives: (1) To classify deepfake-related criminal conduct through a coherent technology-centered scheme. (2) To structure the consequences of deepfake crimes by severity, with attention to individual and systemic impact. (3) To examine how technological modality interacts with harm outcomes and escalation patterns. (4) To analyze judicial reasoning in deepfake cases, especially harm, causation, and evidence. (5) To estimate the contribution of technological and contextual factors to severe harm. (6) To distinguish and integrate personal and systemic harm predictors. (7) To model harm escalation under different technological and contextual conditions. (8) To develop an algorithmic framework for proportional qualification of deepfake crimes. (9) The article proceeds from the literature review to the methodological design, the empirical results, the discussion, and the concluding implications. Each section builds on the previous one and preserves the same analytical sequence.

2. Literature review

The use of deepfake technology has introduced new pressure into an already changing legal environment. Scale, anonymity, and automation now interact in ways that existing rules did not anticipate. This makes the legal response uneven and often unstable.

Early debates treated deepfakes as pressure on existing rights rather than as a rupture requiring a new regulatory model. One line of scholarship tried to stretch established intellectual-property concepts[11] Another accepted deepfakes as embedded in media production and argued for sector-specific governance. The tension between doctrinal preservation and institutional redesign remained unresolved[12].

A second line of scholarship shifted attention from rights to harm. These studies showed that deepfakes weaken relational proximity and enlarge the field of abuse[13]. They also moved responsibility outward to creators, distributors, and platforms[14]. In

[11] PAVIS, M. "Rebalancing our regulatory response to Deepfakes with performers' rights", *Convergence: The International Journal of Research Into New Media Technologies*, v. 27, n. 4, 2021, p. 974–998. https://doi.org/10.1177/13548565211033418

[12] YADAV, H.; OZA, J. "The deepfake dilemma: Balancing innovation, ethics, and accountability through law", *Authorea Preprints,* 2025. https://doi.org/10.36227/techrxiv.175744943.37338204/v1

[13] FLYNN, A.; CLOUGH, J.; COOKE, T. "Disrupting and preventing deepfake abuse: Exploring criminal law responses to ai-facilitated abuse", in *The Palgrave handbook of gendered violence and technology*. Cham: Springer International Publishing, 2022, p. 583–603. https://doi.org/10.1007/978-3-030-83734-1_29

[14] JASSERAND, C. "Deceptive deepfakes: Is the law coping with ai-altered representations of ourselves?" in *2024 International conference of the biometrics special interest group*

that sense, accountability became distributed rather than singular.

The sharpest disruption appeared in the law of proof. Research showed that algorithmic fabrication complicates collection, authentication, and admissibility[15]. Once deepfakes started to compromise evidentiary integrity and distort judicial processes, the issue extended beyond synthetic media detection and became a systemic challenge for criminal justice, legal safeguards, and democratic security[16].

Civil and tort responses proved similarly strained. Patchwork adaptations could absorb some cases, but they did not capture anonymous circulation, virality, or diffuse social injury[17]. Low-risk liability classifications also understated the damage caused by malicious deepfake use[18].

Criminal-law analyses reached a more severe diagnosis. They pointed to enforcement delay, weak victim redress, and reactive doctrinal tools[19]. The shared conclusion was not that technology itself should be criminalized, but that legal systems require a more coherent framework for synthetic harms[20].

Taken together, the literature showed fragmented regulation, thinner evidentiary certainty, and dispersed liability. Yet the criminal-law implications of severe psychological, economic, democratic, and systemic harms remained underdeveloped. That gap justified a focused inquiry into criminal responsibility for deepfake-enabled conduct.

3. Methods and materials

3.1. Research procedure

The research procedure followed a staged and integrative design. It moved from technological classification and harm topology to cross-impact analysis and judicial content analysis, and only then to regression, dual-scale indexing, scenario modeling, and algorithmic formalization. This sequence allowed empirical observation, doctrinal interpretation, and procedural modeling to be connected without collapsing them into one method.

3.2. Methods

The methodological design used a sequential, multi-layered analytical pipeline to study deepfake crime as a socio-technical and legal phenomenon. It combined classificatory, inferential, and modeling techniques. This combination made it possible to cross-validate empirical findings, doctrinal interpretation, and algorithmic formalization.

(BIOSIG). Darmstadt, Germany: IEEE, 2024, p. 1-4. https://doi.org/10.1109/biosig61931.2024.10786729

[15] MEKKAWI, M. H. "The challenges of digital evidence usage in deepfake crimes era", *Journal of Law and Emerging Technologies*, v. 3, n. 2, 2023, p. 176–232. https://doi.org/10.54873/jolets.v3i2.123

[16] SANDOVAL, M.-P.; DE ALMEIDA VAU, M.; SOLAAS, J.; et al. "Threat of deepfakes to the criminal justice system: A systematic review", *Crime Science*, v. 13, 2024, Article 41. https://doi.org/10.1186/s40163-024-00239-1

[17] KADRI, T. E.; WEST, S. R. "Deepfake torts: Emerging tort frameworks in U.S. deepfake regulation", *Journal of Tort Law,* v. 18, n. 2, 2025, p. 515-552. https://doi.org/10.1515/jtl-2025-0032

[18] HRISTOV, G. "Genuine harms behind artificial content: How EU regulation can combat malicious use of deep fake technology", *SSRN Electronic Journal,* 2025. https://doi.org/10.2139/ssrn.5634715

[19] DARMA, M., et al. "Legal implications of deepfake technology misuse in digital content on social media", *Science of Law*, v. 3, 2025, p. 98-103. https://doi.org/10.55284/eqazc148

[20] MARQUES MOREIRA, J. "The use of deepfakes in the criminal justice domain: An emerging reality?" *SSRN Electronic Journal*, 2025. https://doi.org/10.2139/ssrn.5329688

(1) *Technological typology of deepfake crimes (2019–2025)*. A structured classification method was applied to group deepfake incidents by underlying generation, transformation, and dissemination technologies. The method served to isolate technological drivers of harm and to establish a comparable analytical baseline across cases.

(2) *Consequence-based topology by severity of harm (2019–2025)*. Topology construction was used to map harm outcomes along an ordinal severity gradient. This method enabled differentiation between localized injury and large-scale systemic disruption, forming the basis for later severity calibration.

(3) *Cross-impact mapping of modality and consequences*. A matrix-based cross-impact analysis linked technological modalities to observed harm intensities. It was employed to detect non-linear amplification effects and interaction patterns between automation, diffusion, and institutional targeting.

(4) *Judicial content analysis of deepfake verdicts (2019–2025)*. Qualitative-quantitative content analysis was conducted on judicial decisions to examine harm operationalization, causal attribution, and evidentiary thresholds. The method exposed structural biases toward individualized harm in judicial reasoning.

(5) *Regression-based contribution analysis*. Multivariate regression modeling quantified the marginal contribution of technological and contextual predictors to severe harm. This method identified systemic variables as dominant explanatory factors relative to victim-centered indicators.

(6) *Dual-scale predictor structuring*. Index construction techniques were used to separate and weight personal and systemic harm predictors. The objective was to reveal divergence between judicial recognition and latent harm potential.

(7) *Scenario-based consequence modeling*. Counterfactual and scenario modeling simulated harm escalation under varying diffusion and automation conditions. This method validated the disproportionate growth of systemic harm beyond individual impact thresholds.

(8) *Dual-scale algorithmic framework modeling*. UML-based algorithmic modeling formalized the empirical findings into a procedural qualification framework. The method translated analytical results into an operational structure suitable for criminal-justice implementation.

3.3. Sample

The empirical sample was designed to capture the transition from implicit tolerance of synthetic media to explicit regulatory and criminal responses. The 2019–2025 period marked the point at which deepfake technologies moved from experimental misuse to socially consequential harm. The sample therefore linked AI governance, platform accountability, and criminal liability within one comparative frame.

The dataset for Table 1 was built entirely from publicly accessible materials. These included judicial decisions, legislative and regulatory documents, enforcement reports, and peer-reviewed legal analysis. The sample is therefore transparent, verifiable, and replicable. Table 1 compares jurisdictions across three dimensions: regulatory level, technological modality, and legally recognized consequences. The aim was not to rank jurisdictions. The aim was to show how similar synthetic practices trigger different legal responses under different evidentiary and regulatory conditions.

Table 1. Comparative analysis of deepfake-related criminal law and cases (2019–2025).

Country + Instrument Regulating Deep Fake Offending	**Explicit Provisions (Section, Article or Paragraph)**	**Precedent(s) Relating to Deep Fakes (Outcome)**	**Deep Fake Technology Used**	**Academic Research**
EU – AI Act	Art 50(4) obliges providers of "deep fakes" (synthetic/manipulated images/audio/videos) to providesufficient transparency/labelling.	There is no supranational EU "criminal deep fake verdict", as the enforcement route is via Member State Criminal Law, and the EU Layer is largely compliance obligations.	Synthetic media generation; Disclosure/Provenance Bypass Patterns.	Buckingham; de Souza[21, 22]
EU – Digital Services Act (DSA)	Art 16(1) provides for a notice-and-action procedure for illegal content (e.g. deep fake abuse, where it constitutes an illegal act under national law).	There is no DSA "deep fake crime" judgement found that relates directly to this provision; instead, the DSA focuses on procedural/ platform governance.	Platform mediated dissemination; Rapid Re-Upload/Virality.	Yavuz; Schmitz-Berndt et al.[23,24]
UK - Prosecution by CPS (CSAM + AI Image Generation).	CPS did not provide any detail on its charging structure in the CPS Note; CPS considered the AI generated images as "indecent images" within the current Sexual Offences framework.	R v Hugh Nelson (Bolton Crown Court, 28 October 2024): 18 years imprisonment plus six years extended license for creating and distributing multiple AI generated child sexual abuse images.	AI Image Generation / Transformation of Real Photos into CSAM-like Imagery	Montasari; Gaitis et al.[25,26]

[21] BUCKINGHAM, E. "The challenges of deepfake technology on the decision-making processes within law enforcement: A study within the EU landscape" (Master's thesis). University of Malta, 2025. Available at: https://www.um.edu.mt/library/oar/handle/123456789/141835 (accessed on 18 January 2026).

[22] DE SOUZA, R. R. M. "Legal remedies and regulatory frameworks to combat ai-driven deepfakes", In *Mitigating the risks of AI deepfakes*, Boca Raton: CRC Press, 2026, p. 94–116. https://doi.org/10.1201/9781003539032-5

[23] YAVUZ, C. "Criminalization of the dissemination of nonconsensual deepfake pornography in the European Union. A comparative legal analysis", *Research Day Law and Criminology,* 2024, 2024. Available at: https://biblio.ugent.be/publication/01JMFTQENS8CCYFVRZGNVRE56V (accessed on 18 January 2026).

[24] SCHMITZ-BERNDT, S. et al. "Non-consensual deepnudes: Responses under EU law to a novel form of sexual abuse", *International Review of Law, Computers & Technology*, 2026, p. 1–29. https://doi.org/10.1080/13600869.2026.2654235

[25] MONTASARI, R. "Responding to deepfake challenges in the United Kingdom: Legal and technical insights with recommendations", in *Advanced sciences and technologies for security applications*. Cham: Springer International Publishing, 2024, p. 241–258. https://doi.org/10.1007/978-3-031-50454-9_12

[26] GAITIS, K. K., et al. "Legal challenges in tackling AI-generated CSAM across the UK, USA, Canada, Australia and New Zealand: Who is accountable according to the law?", *Searchlight 2025–Who Benefits? Shining a Light on the Business of Child Sexual Exploitation and Abuse*, 2025, p. 50-59. Available at: https://www.research.ed.ac.uk/en/publications/legal-challenges-in-tackling-ai-generated-csam-across-the-uk-usa-/ (accessed on 18 January 2026).

Country + Instrument Regulating Deep Fake Offending	**Explicit Provisions (Section, Article or Paragraph)**	**Precedent(s) Relating to Deep Fakes (Outcome)**	**Deep Fake Technology Used**	**Academic Research**
USA (Federal) - CSAM Enforcement	DOJ Case Applies Federal CSAM Offenses to "Deep Fake CSAM" (Possession/Access with Intent)	U.S. v Smelko (Pennsylvania, Sentencing announced May-June 2024): 14 Years, 7 Months imprisonment for possessing "Deep Fake" CSAM.	Deep Fake CSAM (AI Based Face Synthesis / Compositing).	Vyas; Lin[27,28]
The Commonwealth of Australia – Criminal Code Act 1995 (Cth)	at S 474.17(1) defines "Carriage Service Use to Menace/Harass/Offend", which are a number of elements that comprise an offence.	Hayler (District Court of NSW, 21 Jun 2024) sentenced a defendant to nine years of imprisonment for a non-parole term of five-and-one-half years after he was convicted of twenty-eight counts of using a carriage service	Cause offence by posting/uploading and distributing digitally altered intimate images ('deep fakes')	Berry; Celli[29,30]
Australian Online Safety Act - Civil Penalty for Image-Based Abuse	Australia has enacted legislation which permits for civil penalties to be levied against individuals who post non-consensual pornography and/or synthetic images.	The federal court imposed a civil penalty of $343,500 plus legal fees upon a person for posting a deep-fake pornographic image of another individual without their consent and failing to remove it once informed to do so.	Synthetic images are those that have been artificially produced using artificial intelligence, and represent a new form of non-consensual pornography.	Sheehy; Martin[31,32]

[27] VYAS, A. D. "United States of Deepfake", *Tennessee Law Review*, v. 92, 2024, 307. Avaialble at: https://ssrn.com/abstract=4910852 (accessed on 18 January 2026).
[28] LIN, L. S. F. "Organisational challenges in US law enforcement's response to ai-driven cybercrime and deepfake fraud", *Laws*, v. 14, n. 4, 2025, 46. https://doi.org/10.3390/laws14040046
[29] BERRY, E. "Is misuse of deepfake technology adequately addressed by Australian intellectual property law, or does Australia need to introduce a right of publicity tort?" *Intellectual Property Forum: Journal of the Intellectual and Industrial Property Society of Australia and New Zealand*, v. 140, 2025, p. 7-17. Available at: https://search.informit.org/doi/abs/10.3316/informit.T202506150000039149828 1295 (accessed on 18 January 2026).
[30] CELLI, F. "Deepfakes are coming: Does Australia come prepared?" Canberra Law Review, v. 17, 2020, 193. Available at: https://search.informit.org/doi/10.3316/agis.20211118057010 (accessed on 18 January 2026).
[31] SHEEHY, S. "The effects and influence of artificial intelligence and likelihood of impacts to Australia", *Journal of the Australian Institute of Professional Intelligence Officers*, v. 29, n. 2, 2021, p. 3-9. https://doi.org/10.3316/informit.387429703808070
[32] MARTIN, N. "Online safety regulation of deepfake abuse: A case study on Australia's eSafety Commissioner", *Griffith Law Review*, 2025, p. 1–24. https://doi.org/10.1080/10383441.2025.2504791

Country + Instrument Regulating Deep Fake Offending	**Explicit Provisions (Section, Article or Paragraph)**	**Precedent(s) Relating to Deep Fakes (Outcome)**	**Deep Fake Technology Used**	**Academic Research**
South Korean Digital/Sexual Offenses	Criminal Prosecution through Judicial Decision-Making - South Korean criminal law has established a statutory provision that provides for a judicial determination regarding whether an individual's creation and distribution of AI generated, sexually explicit, images is a digital sexual crime. In a recent case,	In a recent case, the Seoul Central District Court (as reported on 30th October 2024), sentenced a principal defendant to ten (10) years imprisonment and co-defendant to four (4) years imprisonment for distributing over 2,000 AI-generated images.	The images were distributed through messaging and/or social media platforms. The defendants were convicted of using facial recognition software to create the AI-generated images, and then distribute them through various platforms.	Kim; Ji[33,34]
Chinese Deep Synthesis Legislation	Translation of Article 6 - Chinese criminal law provides a general prohibition on the use of "deep synthesis" to produce, publish or disseminate unlawful material.	This includes prohibitions on spreading false information, producing and publishing pornographic materials and violating individuals' rights.	Although no cited criminal judgment was found within the retrieved articles (the article mainly establishes a compliance or prohibition baseline; prosecutions will generally occur pursuant to general provisions of criminal law), this regulation defines what constitutes a "deep synthesis" pipeline - namely the creation of face-swap, voice-cloning and synthetic avatars.	Zheng et al.; Peng and Lee[35,36]

[33] KIM, K. "Deepfakes: Challenges to intellectual property rights in South Korea", *GRUR International*, v. 74, n. 6, 2025, p. 532-542. https://doi.org/10.1093/grurint/ikaf044

[34] JI, S. "#MeToo in an AI-generated deepfake sexual violence era in South Korea", *Women's Studies International Forum*, v. 112, 2024, 103146. https://doi.org/10.1016/j.wsif.2025.103146

[35] ZHENG, G.; SHU, J.; LI, K. "Regulating deepfakes between lex lata and lex ferenda – A comparative analysis of regulatory approaches in the U.S., the EU and China", *Crime, Law and Social Change*, v. 83, n. 1, 2025. https://doi.org/10.1007/s10611-024-10197-z

[36] PENG, H.; LEE, P.-W. "Reimagining U.S. tort law for deepfake harms: Comparative insights from China and Singapore", *Journal of Tort Law,* v. 18, n. 2, 2025, p. 579-607. https://doi.org/10.1515/jtl-2025-0028

Country + Instrument Regulating Deep Fake Offending	**Explicit Provisions (Section, Article or Paragraph)**	**Precedent(s) Relating to Deep Fakes (Outcome)**	**Deep Fake Technology Used**	**Academic Research**
Texas, USA - Election Code - Political Deception	Texas Election Code, Title 15, Chapter 255, Section 255.004(A), prohibits a political advertisement purporting to come from a false source.	Although no convictions have been found in the retrieved literature for DeepFakes related to elections (the statute does exist, but sample cases could not be found), this statute would likely prohibit the use of political Deepfakes to impersonate candidates and/or other politicians.	Source spoofing is a common tactic used in Deepfakes where a political candidate's voice is manipulated to make it appear as though they said something they did not actually say.	Rodriguez; Goldberg[37,38]
California, USA - Elections Code - Deceptive Media in Elections	California Elections Code, § 20010, prohibits the dissemination of "materially deceptive" audio/video near the time of an election.	The code also provides relief and restrictions for candidates whose campaigns have been impacted by such media.	Like Texas, California has no reported Deepfake convictions under this statute, however, this statute would likely apply to manipulated election videos ("materially deceptive" media).	Farish; Ugwuoke and Sanfilippo[39,40]

Source: developed by the authors.

3.4. Research tools

This section translates observed harm patterns into explicit mathematical form. The regression and dual-scale formulations were used because deepfake harm grows non-linearly with automation, diffusion, and institutional reach. The resulting model is transparent, reproducible, and compatible with evidentiary and sentencing logic.

(1) Regression-Based Contribution of Technological and Contextual Factors to Severe Deepfake Harm. The severity of deepfake harm was modeled as a function of technological and contextual predictors using a generalized linear specification:

[37] RODRIGUEZ, X. "Artificial intelligence (AI) and the practice of law in Texas", *South Texas Law Review*, v. 63, 2023, 1. Available at: https://texasbarsections.com/wp-content/uploads/2023/11/Rodriguez-Paper.pdf (accessed on 18 January 2026).

[38] GOLDBERG, H. "States legislating against digital deception: A comparative study of laws to mitigate deepfake risks in American political advertisements", *Notre Dame Journal on Emerging Techologies,* v. 6, 1, 2025. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4912795 (accessed on 18 January 2026).

[39] FARISH, K. "Do deepfakes pose a golden opportunity? Considering whether English law should adopt California's publicity right in the age of the deepfake", *Journal of Intellectual Property Law & Practice*, v. 15, n. 1, 2019, p. 40–48. https://doi.org/10.1093/jiplp/jpz139

[40] UGWUOKE, V.; SANFILIPPO, M. R. "The current landscape of deepfake legislation in the United States", *Journal of Information Policy*, v. 15, 2025, p. 97-129. https://doi.org/10.5325/jinfopoli.15.2025.0004

$$H_i = \beta_0 + \sum_{k=1}^{K} \beta_k T_{ik} + \sum_{m=1}^{M} \gamma_m C_{im} + \varepsilon_i \qquad (1)$$

where H_i – observed harm severity score for case i (ordinal–continuous, normalized to [0,1]); β_0 – intercept term; T_{ik} – technological factors (automation depth, batch synthesis, diffusion velocity, cross-platform propagation); C_{im} – contextual factors (victim vulnerability, institutional targeting, duration of exposure, jurisdictional reach); β_k , γ_m – partial effect coefficients estimating marginal contribution to harm severity; ε_i – stochastic error term capturing unobserved variance.

Coefficient magnitudes $(|\beta_k|)$ demonstrated that systemic technological factors dominated individual-contextual predictors in explaining high-severity outcomes.

(2) Dual-Scale Predictor Structure of Deepfake Crime Severity (Judicial Recognition vs. Systemic Harm Potential). To capture the divergence between legally recognized harm and latent systemic impact, severity was decomposed into two orthogonal indices:

$$P_i = \sum_{j=1}^{J} \omega_j^{(P)} \times p_{ij}, \; S_i = \sum_{l=1}^{L} \omega_l^{(S)} \times S_{il} \qquad (2)$$

where P_i – personal harm index (judicially salient); S_i – systemic harm index (often judicially latent); p_{ij} – personal-level indicators (psychological trauma, reputational loss, coercion, economic injury); s_{il} – systemic-level indicators (scale of diffusion, automation intensity, institutional distortion, persistence); $\omega_j^{(P)}$, $\omega_l^{(S)}$ – normalized weights derived from regression loadings and scenario modeling.

Overall severity qualification followed a dominance-sensitive fusion rule:

$$Q_i = \max(\mathrm{P}_i, \alpha \times \mathrm{S}_i), \; \alpha > 1 \qquad (3)$$

where Q_i – final severity class used for procedural routing; α – amplification coefficient reflecting the empirically higher social cost of systemic harm.

This formulation formally encoded the empirical finding that systemic harm escalated severity even when personal injury remained moderate, resolving the structural underestimation observed in judicial practice. The variables in the regression and dual-scale models were coded from the judicial cases included in the empirical dataset. Information came from public court decisions, enforcement reports, and regulatory case descriptions listed in Table 1. Each case was coded through a predefined protocol that converted qualitative findings into standardized indicators.

Contextual variables were extracted from the factual record of each case. Victim vulnerability reflected legally recognized vulnerability factors. Institutional targeting captured cases directed at elections, public institutions, financial systems, or judicial evidence. Duration and reach were coded from documented persistence and dissemination scope.

All variables were normalized as ordinal or binary indicators before entering the regression model and dual-scale indices. This ensured comparability across jurisdictions and improved reproducibility.

Quantitative analysis was conducted in a Python environment using NumPy, pandas, statsmodels, scikit-learn, and SciPy. Matplotlib was used for analytical visualization. The dual-scale qualification framework was then formalized through UML modeling to represent decision points, parallel harm pathways, and procedural transitions.

To ensure transparency, all variables were coded under a predefined rule-based protocol derived from the case materials. Each case was annotated with technological indicators, contextual indicators, and harm indicators. Severity levels

were assigned through the consequence-based topology introduced earlier in the study.

4. Results

The first analytical stage adopted a technology-centered perspective rather than a purely legal one. Deepfake incidents were grouped by synthetic media type and by the mechanisms used to generate, transform, or distribute them. This reduced legal noise and made technological drivers of harm easier to compare across jurisdictions.

The categories in Table 2 were derived through a structured empirical classification of judicial cases and regulatory materials from 2019 to 2025. Recurrent technological mechanisms were first extracted from the materials. They were then grouped according to their functional role in production, transformation, and dissemination. The resulting typology is therefore technology-centered rather than doctrine-centered.

Table 2. The technological typology of deepfake crimes identified in this empirical sample (2019–2025).

Type of deep fake	**What type of technology is being used to make a deep fake?**	**Documented cases of each type**	**Judicial outcome/enforcement level**
Synthetic sexual imagery (CSAM/non-consensual porn)	GAN- and diffusion-based image generation; Face synthesis; Image-to-image transformation	2024, UK – R v Hugh Nelson (18 years); 2024, USA – U.S. v Smelko (14y7m); 2024, South Korea – Seoul Central District Court (10 + 4 years); 2024–2025, Australia – NSW District Court (9 years), Federal Court (AUD 343,500)	High: custodial sentences (9–18 years); significant financial penalties
Transformative sexual deep fakes (Real image -> synthetic abuse)	Face swap; Morphing of a real photograph; Dataset driven composition	2024, UK – R v Hugh Nelson; 2024, USA – U.S. v Smelko; 2024, South Korea – Mass production of deep fakes	High: custodial sentencing (linked to CSAM qualification)
Mass produced sexual humiliation deep fakes	Automated batch processing; Face swap pipelines; Encrypted redistribution	2024, South Korea – Seoul Central District Court (Approximately 2,000 images; Custody Sentences)	High: custodial sentences; severity driven by scale
Harassment and coercive deep fakes	Synthetic intimate imagery; Platform based distribution; Reposting automation	2024, Australia – NSW District Court (Hayler); 2025, Australia – Federal Court (Rotondo)	Moderate–High: imprisonment or substantial civil penalties
Political deep fakes (Electoral manipulation)	Audio visual impersonation; Source spoofing; Edited video that induces misperception	2019–2025, USA (Texas §255.004; California §20010) – No Criminal Convictions Identified	Low: no criminal enforcement; regulatory provisions only

Type of deep fake	What type of technology is being used to make a deep fake?	Documented cases of each type	Judicial outcome/enforcement level
Regulated synthetic media (Compliance layer)	Synthetic image/audio/video generation; Provenance suppression	2022–2025, EU – AI Act Article 50(4); China – Deep Synthesis Provisions Article 6 (No Direct Criminal Verdicts)	Low: regulatory compliance; absence of criminal sanctions
Platform amplified deep fakes	Algorithmic recommendation; Rapid re upload; Virality Loops	2022–2025, EU – DSA Article 16; Australia – Criminal Code Section 474.17(1)	Moderate: indirect enforcement via platforms; limited criminal liability

Source: developed by the authors.

Table 2 shows a clear asymmetry in enforcement. Sexualized deepfakes based on high-fidelity face synthesis are regularly associated with custodial sanctions or significant penalties. Political and compliance-oriented deepfakes show much weaker enforcement despite comparable dissemination capacity. The most severe outcomes arise where generative synthesis intersects with batch production and platform amplification.

Table 3. Consequence-based topology of deepfake crimes by harm severity (2019–2025).

Types of Consequences	Deepfake Crime Types Which Result in Such Consequences	Documented Cases
The severe sexual exploitation and long-lasting psychological damage	The creation of AI-created CSAM; and the non-consensual use of sexual deep fakes to create large amounts of sexually humiliating content.	UK 2024 – R v Hugh Nelson (18 years); 2024 USA — U.S. v Smelko (14 y 7 m); 2024 S. Korea — Seoul Central District Court (10+4 yrs.)
The long-term harm to an individual's reputation and the coercion to be a victim	The use of synthetic sexualized deep fake harassment; synthetic sexual intimate images for extortion purposes; the coerced dissemination of such images.	Australia 2024 — NSW District Court (Hayler, 9 years); Australia 2025 — Federal Court (Rotondo, AUD 343,500)
The high level of psychological trauma that can result with the use of technology without being sexual in nature.	Synthetic impersonation; and the humiliation and embarrassment of victims as a result of use of face swap technology; the coordinated campaign of harassment using technology.	Australia 2024-2025 — image-based abuse cases prosecuted under Criminal Code and Online Safety Act.
Large-scale financial loss due to fraudulent activity as a result of use of deepfakes.	Identity theft and voice-cloning scams by means of impersonating individuals through the use of deepfakes.	China - 2019 to 2025 - Enforcement of laws on the large scale of fraud and deception using the general criminal code of China and reported cases of prosecution for fraud in comparison with other countries regulatory responses
Disruption of democracy and elections	The creation of false political audio and video materials that create a false appearance of an individual participating in a campaign or election.	USA - 2019 to 2025 - Section 255.004 of the Texas Election Code and Section 20010 of the California Elections Code (No convictions found).

Types of Consequences	**Deepfake Crime Types Which Result in Such Consequences**	**Documented Cases**
The undermining of the overall system of trust for evidence and the integrity of all procedures in litigation.	Creation of unreliable evidence by the creation of false synthetic audio-visual materials; and contamination of evidence by the presence of deepfakes.	Multi-jurisdictional - 2023-2024 - Findings from both the legal and forensic communities (No singular decision or verdict).
Societal harm and destabilization at the platform level	Widespread dissemination of deepfakes through social media platforms and repetitive posting of deepfakes after they have been removed from those same platforms.	EU - 2022 to 2025 - Governance framework for Digital Services Act (DSA) and Australia - Ongoing enforcement actions against social media companies.

Source: developed by the authors.

Table 3 shifts the analysis from technological modality to harm topology. It orders deepfake outcomes from individual psychological injury to broad societal disruption. The table shows that criminal liability tracks severity, irreversibility, and social externality of harm more closely than technological novelty alone.

The topology also reveals a stratified legal response. Custodial sentencing clusters around sexual exploitation and lasting psychological injury, whereas economic, electoral, and epistemic harms are still handled indirectly or fragmentarily. This pattern justified the next step: identifying composite high-severity risk profiles through cross-impact analysis.

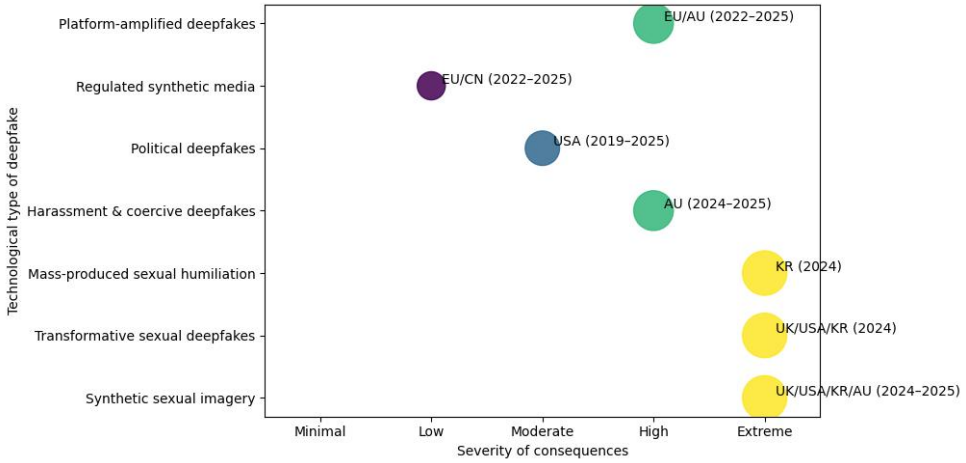

Figure 1. Cross-impact mapping of deepfake crimes, technological modality, and consequences. Source: developed by the authors.

Figure 1 shows a stable connection between technological configuration and sanctionable harm. Face-synthesis and automated image transformation cluster at the highest severity level and correlate with custodial sentences of 9 to 18 years. Political and compliance-layer deepfakes occupy lower sanction zones despite similar dissemination capacity. Platform-amplified deepfakes form a transitional category in which persistence and propagation drive harm escalation. Figure 1 does not imply that sexualized and political deepfakes are treated as equivalent categories in law. It shows instead how current criminal justice practice operationalizes harm within existing doctrinal frameworks. Sexualized synthetic imagery fits established offenses with identifiable victims and direct injury. Political deepfakes more often remain within regulatory or electoral frameworks, where systemic informational harm is harder to criminalize directly.

Table 4 shows a common judicial pattern across jurisdictions. Courts establish causation most readily when deepfakes produce direct, tangible, and lasting personal harm, especially sexual exploitation, reputational destruction, and severe psychological distress. By contrast, societal harms such as electoral distortion or institutional distrust remain weakly specified and often appear as regulatory rather than criminal concerns.

Table 4. Judicial content analysis of deepfake verdicts (2019–2025).

The known case	**What type of a deepfake crime?**	**What consequences did the deepfake crime have?**	**What was the sentence/outcome?**	**What operationalization and causal linkages were made by the court?**
R v Hugh Nelson (UK, 2024)	Transformative sexual deep fake; AI-generated CSAM	Long term psychological damage to children; long term victimization of the child; severe sexual exploitation	18 years in prison + 6 year extended license	Court clearly connected the use of AI based image manipulation to the autonomous generation of CSAM; causality was established by court that synthetic manipulation caused permanent harm and increased culpability beyond physical contact.
U.S. v Smelko (USA, 2024)	Possession of deepfake CSAM	Risk of further distribution of CSAM; Sexual exploitation of children; Normalizing the behavior of abusive adults	14 years and 7 months in prison	Judge reasoned that CSAM created using algorithms is functionally identical to authentic CSAM; Court extended causal chain of creation, dissemination and re-victimization of the child via the use of AI generated images.
Seoul Central District Court (South Korea, 2024)	Severe mass production of sexual humiliation deepfakes.	Victims experienced extreme psychological distress; victims were socially excluded; victims had their reputations completely destroyed.	Principal offender sentenced to 10 years in prison; co-offender sentenced to 4 years in prison.	The court stressed that both the large-scale synthesis of images and the scale at which they were produced were causative factors that amplified harm. The court did not measure harm on an individual basis using the severity of individual images, but by measuring harm based on volume, duration, and victim identifiability of the images.

The known case	**What type of a deepfake crime?**	**What consequences did the deepfake crime have?**	**What was the sentence/outcome?**	**What operationalization and causal linkages were made by the court?**
New South Wales District Court (Hayler) (Australia, 2024)	Deepfake harassment with a sexualized theme.	Victim experienced long-term damage to reputation; victim was coerced into certain behaviors as a result of the harassment.	Offender sentenced to 9 years imprisonment; non-parole period is 5.5 years.	Harm was caused by the defendant disseminating images of the victim through a platform. The court also determined that the algorithmic reposting of the images by the platform was a factor that increased the harm and the intent of the defendant.
The Rotondo case in Australia's Federal Court (Civil enforcement) from 2025.	A case about non-consensual deep fake pornography	In the case there were ongoing psychological harms and the victim had lost control over her own digital life.	She was ordered to pay AUD 343,500 in addition to court costs.	Judge Rotondo found that the ongoing psychological harm and loss of control was a direct result of the fact that she did not take action to have the fake images taken down, and that the continued use of the fake images by others was an independent cause for the harm
Cases relating to election activity in the U.S. (Texas / California) (2019 – 2025)	Audio Visual Deep Fakes Distortion of political information;	Potential deceptions of voters	No Criminal Charges	While the Courts recognized that the audio-visual deep fakes could be used to mislead people, they failed to establish a specific amount of damage that would need to occur before legal action could be taken against someone using them; the courts also said that it would be difficult to show that the deep fakes caused some form of measurable harm to voters.

Source: developed by the authors.

Courts primarily focus on individualized harm in deepfake cases. This matches the cross-impact mapping, where custodial sentences cluster around sexualized deepfakes and political or infrastructural manipulation remains weakly sanctioned. The reason is practical as well as doctrinal: courts have stronger evidence for direct victim injury than for algorithmic diffusion or cumulative social damage.

The next analytical step therefore moved from qualitative judicial reasoning to formalized inference. Regression analysis was used to quantify how technological features, dissemination scale, and victim vulnerability relate to severe harm. This made diffuse threats more measurable and more useful for criminal-law threshold setting.

The regression in Figure 2 was estimated from the structured case dataset compiled for this study. Each case was coded for technological characteristics and contextual factors. The dependent variable captured normalized harm severity derived from the consequence-based topology, and the independent variables were entered as ordinal or binary indicators in a generalized linear model.

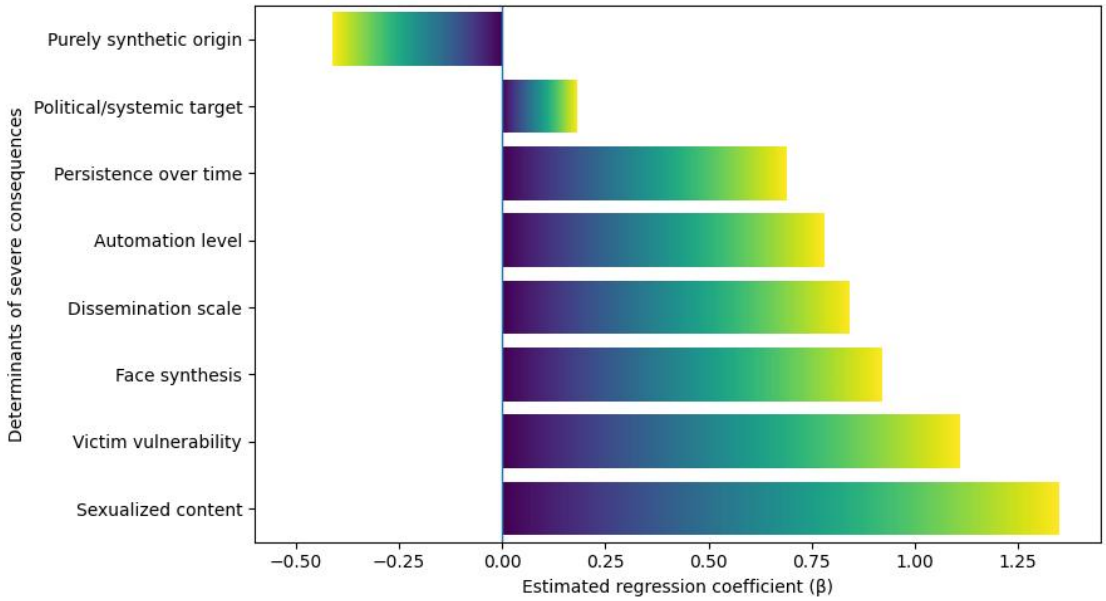

Figure 2. Regression analysis of variables contributing to severe deepfake harm
Source: developed by the authors.

Figure 2 makes the hierarchy of harm drivers explicit. Sexualized content, victim vulnerability, and face synthesis produced the strongest marginal effects. Dissemination scale, automation, and temporal persistence acted as second-order amplifiers. Political or systemic targeting remained weakly weighted within judicially legible harm, revealing a structural blind spot.

The variables in Figure 2 were quantified through a standardized coding procedure applied to the empirical dataset. Technological features were coded as binary or ordinal indicators, while harm severity was normalized from the consequence-based classification of outcomes. This allowed the model to estimate the relative contribution of technological and contextual factors to severe harm.

This gap motivated the next analytical step: identifying robust predictors of deepfake crime severity capable of capturing both individualized and systemic harm, thereby enabling a more coherent framework for criminal responsibility in algorithmically mediated offenses (Figure 3).

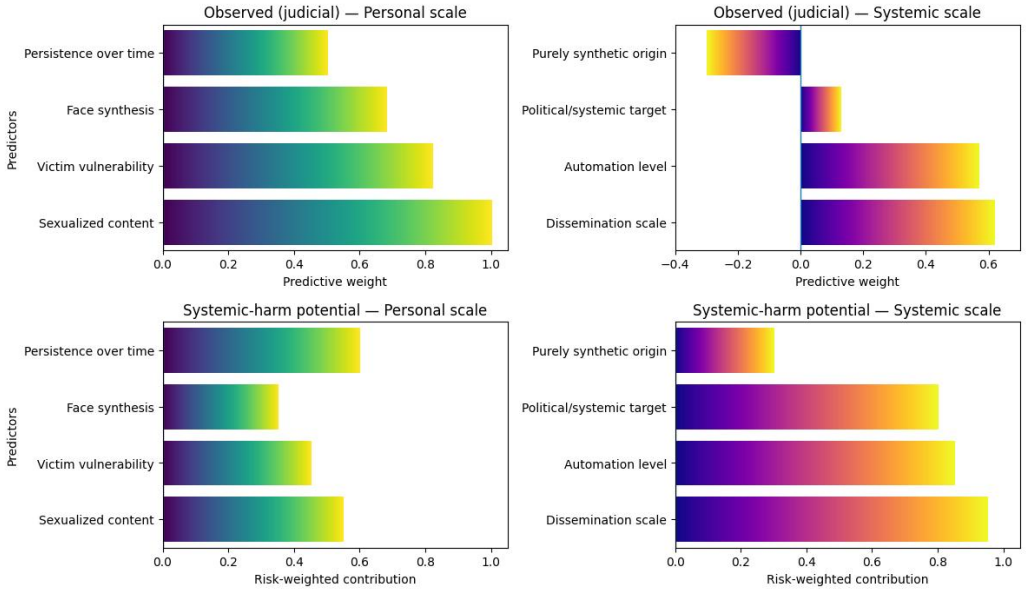

Figure 3. Dual-scale predictors of deepfake crime severity: judicial recognition vs systemic harm potential. Source: developed by the authors.

Figure 3 contrasts recognized and latent predictors of deepfake harm. On the judicial scale, most predictive weight lies in personal variables such as sexualized content, victim vulnerability, and face synthesis. On the systemic scale, dissemination, automation, and institutional targeting become the key determinants of large-scale harm, even where no direct victim is easily identified.

The figure therefore shows that harm is shaped less by content alone than by the way content is propagated through algorithmic systems. Courts respond most strongly where high-fidelity synthetic representation is tied to identifiable victim injury. At the same time, cumulative and societal harm remains underweighted in legal practice.

Table 5. Scenario modeling of personal and systemic impacts of deepfakes.

Scenario	**Deepfake configuration**	**Main harm source**	**Modelled personal harm**	**Modelled systemic harm**	**Interpretation of ANALYSIS**
S1. Sexual Deepfake targeting an individual;	ace-synthesized; distributed to a few victims.	Individual.	Severely High (Psychological Trauma; Annihilation of Reputation)	Very Low	This is the classical case that receives most judicial attention, as it has severe localized harm.
S2. A series of harassments using Deepfakes.	Reposting of synthetic images in an intimate setting, repeatedly posted.	Individual → Network.	High (Chronic Distress; Coercion)	Moderate.	Harm is increased by persistence but the harm is still located at the victim.
S3. Mass produced humiliations using deep fakes.	Batch made, automated facial swaps with many output possibilities.	Large numbers of people sharing information on the internet.	Moderate - High (Victimization can be diffuse).	High (Normalization of abuse and social chilling effects)	From an individual to a group of people, the impact of a deep fake is now a collective harm dynamic.

Scenario	Deepfake configuration	Main harm source	Modelled personal harm	Modelled systemic harm	Interpretation of ANALYSIS
S4. Deep Fake Video Amplification through Social Media Platforms.	Deep fake video created with high fidelity; amplified through algorithms.	A society.	None, because no one can be identified as the victim of this type of crime.	Extremely high (Trust in institutions has been eroded by large-scale behavioral manipulation.)	Legal Blind Spot: Harm caused by deep fakes are systemic, but very poorly defined within the criminal justice system.
S5. Deep fake in elections.	Timed impersonation of politicians for political purposes.	Democracy is being threatened.	Low.	Extremely dangerous to all levels of government as an institution, and extremely damaging to legitimacy.	Even though there is no victim in terms of an individual who has been directly harmed by the impersonation, the amount of harm caused exceeds that of S1-S2.
S6. Financial impersonation through use of voice cloning.	Automated voice synthesis; Mass impersonation.	The population at large.	Moderate; Distributed economic loss.	Very High; Losses due to market instability and fraud ability.	Scale converts the economic losses that are moderate for each individual to be very high for the system as a whole.
S7. Synthetic Evidence Pollution.	Synthetic "Evidence" in Form of Audio or Video.	Justice System.	Low.	Extremely high risk of distrust of evidence, and procedural failures.	Damage to an individual's systemic integrity is secondary to overall systemic damage to society as a whole.
S8. Large-scale campaign through multiple social media channels to create and distribute DeepFakes.	Through automation, viral spread, and persistence across social media channels.	All sections of society.	Moderate.	A high-risk configuration was identified (information system instability).	We have identified Peak Risk Configuration from our Cross-Impact Analysis.

Source: developed by the authors.

Table 5 makes the asymmetry between personal and systemic harm visible. Personalized deepfake abuse can produce very intense injury, but that harm saturates quickly at the victim level. Systemic harms—such as platform disinformation, electoral manipulation, financial impersonation, and evidentiary

contamination—generate much greater aggregate damage because they scale through automation, speed, and institutional reach.

These results reinforce the earlier typologies and the regression findings. Automation and dissemination operate as harm multipliers, while courts still find personal harm easier to prove than systemic harm. This justified the final step of the analysis: developing an algorithmic framework that converts latent large-scale harm into legally usable categories.

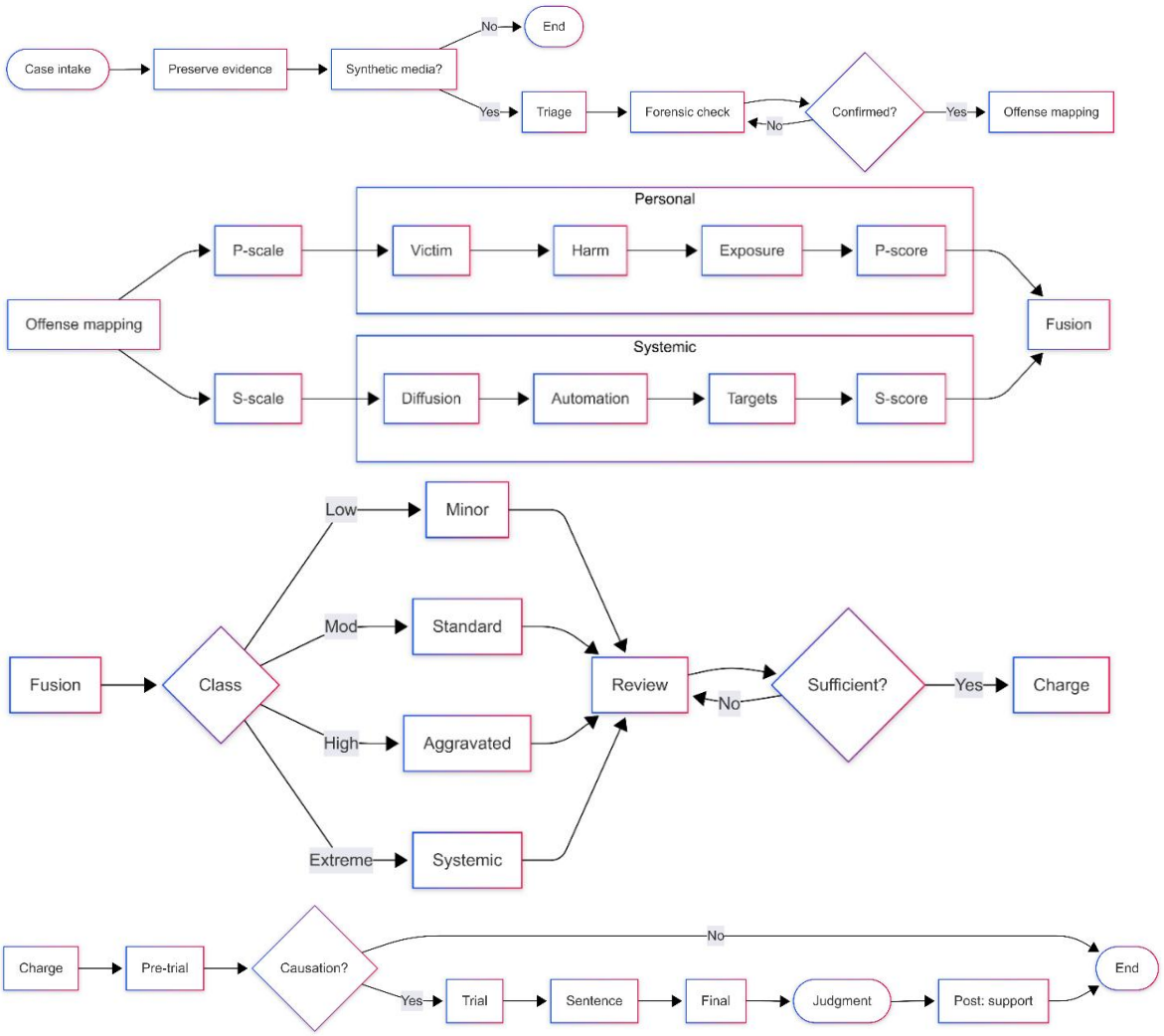

Figure 5. Dual-scale framework for qualifying harm severity in deepfake crimes
Source: developed by the authors.

Figure 5 integrates the empirical findings into one operational decision architecture. It separates personal harm from systemic harm and then recombines them through transparent qualification gates. In this way, it resolves the central contradiction of the study: the legal over-visibility of individualized injury and the under-representation of structural damage.

The framework aligns with the technological typology, the harm topology, the regression results, and the scenario simulations. Its escalation rules reflect the observed pattern in which platform amplification, batch generation, and cross-domain diffusion sharply increase aggregate harm. The fusion logic therefore translates latent systemic risk into legally legible severity classes without abandoning individualized causation.

Implemented as a decision-support layer rather than as an automated adjudicator, the framework can route cases through differentiated judicial pathways according to impact scale. It preserves a continuous assessment of harm across personal and systemic dimensions. This improves proportionality, consistency, and foresight in the qualification of deepfake-enabled offenses.

5. Discussion

The discussion relates the empirical findings of the study to the broader and often fragmented literature on deepfake harm. The key question is whether existing theories of governance, detection, and legislative reform adequately explain harm under conditions of automation, virality, and institutional penetration.

The empirical findings were then compared with the main lines of prior scholarship. This comparison was necessary because the literature often describes harm normatively, while the present study models it empirically and procedurally.

Some scholarship focused on platform governance and democratic epistemics. Those studies clarified why deepfakes destabilize public trust, but they remained largely ex ante. The present study shifted the focus ex post and showed that systemic harm dominates aggregate severity even when judicial practice barely recognizes it.

Another strand emphasized detection, provenance, and technical trust infrastructures. That work was valuable, but it did not resolve the criminal-law problem of qualifying severe harm once diffusion and institutional targeting are activated.

Doctrinal studies improved the treatment of sexual offenses, consent, and privacy. Yet those frameworks mainly explain baseline individualized injury. The present results show that escalation depends on system-level multipliers that offense-specific reforms only partly capture.

Comparative criminal-law research clarified legal fragmentation in fraud and cybercrime responses. It also exposed evidentiary and enforcement difficulties. What it did not provide was an operational mechanism for translating dispersed harm into proportionate criminal pathways.

Taken together, the literature clarified ethical risk, privacy erosion, evidentiary instability, and enforcement gaps. It did not, however, conceptualize systemic harm as a central criminal dimension. The present study addressed that gap by demonstrating the dominance of systemic harm and by proposing a dual-scale algorithmic model for criminal qualification.

5.1. Limitation

A cross-jurisdictional empirical sample of recorded cases (from 2019 to 2025) is likely to be an incomplete representation of all systemic harm that may have occurred but has no judicial evidence to prove it. Regression modeling and the modeling of scenarios used proxy measures for the diffusion, automation and institutional targeting of AI tools in legal processes; this resulted in increased levels of measurement error. Judicial decision making was based solely on the published decisions of courts and as such did not provide insight into how prosecutors make discretionary decisions prior to trial or the effect these decisions have on what is ultimately prosecuted and therefore adjudicated.

5.2. Recommendations

Further longitudinal studies are needed that link procedural decisions by the legal system to their effects on society. This will help establish a clearer causal relationship as it relates to how severe a punishment is. As institutional courts begin using these tools, they should do so through pilot projects in conjunction with

judicial education programs to ensure that the judiciary maintains its ability to interpret the law as well as maintain its own legitimacy.

6. Conclusions

The authors' study of deep fakes demonstrated that there is no way to determine the frequency of serious crimes associated with the misuse of deep fakes by considering only who was harmed by the crime and how the crime affected the victimized person. This is due to the fact that the data collected for the period of 2019 – 2025 indicated that the quality of deep fakes produced using automatic generation technology greatly increased as the production technology for producing deep fakes improved; further, the damage and consequences of deep fakes increased as the rate at which deep fakes could be produced, the amount of time required for the generated deep fakes to reach their intended audience, and the number of organizations involved in the creation, distribution, and creation of deep fakes increased. Further, the damage and consequences of deep fakes did not vary based upon the realism of the video depicting the subject of the deep fake or the degree of deception contained in the video.

As such, according to the results of the regression analysis and the results of the analyses of the scenario models, the degree of systemic harm represented the greatest predictor of the variance in the severity of the crimes involving deep fakes as compared to the degree of individual harm. Individual harm was substantially ignored in the court processes. Therefore, the majority of the social and economic effects of deep fake-related cases were felt in terms of societal procedures and economic impacts rather than the treatment of the cases in the courts where the identical "who hurt whom" narrative paradigm was applied.

In terms of doctrine, the authors conducted a content analysis of court decisions and found that there is a growing tendency towards focusing upon the individual (i.e., the person whose likeness was used to create the deep fake) who suffered harm directly from the crime and the timing of when the harm was suffered by the individual; therefore, the courts ignore the systemic harm resulting from the dissemination of the video. The authors developed a dual scale algorithmic framework for bridging the gap between the two scales of harm by translating harms of the systemic scale into illegible procedural measures of severity that do not exceed the existing boundaries of criminal law. The authors also combined the personal and systemic predictors into one qualification engine in order to develop a new operational model that will assure that the prosecution of deep fake crimes will be proportionate, consistent, and foreseeable. In addition, the authors have identified deep fake crimes as both technology-based anomalies and as scalable/offense categories that require the criminal justice system to respond to these crimes with a data-driven and adaptable approach.

6.1. Academic relevance

The authors' most important academic contribution has been to provide a two-tiered understanding of the severity of crimes committed utilizing deepfakes and to separate individual and system harm within one analytical framework. Additionally, the authors determined through empirical research that systemic factors such as the speed of dissemination, the degree of automation utilized in generating the deepfakes and the degree to which institutions were involved in the creation and dissemination of the deepfakes were the primary determinants of the degree of severity of crimes committed utilizing deepfakes; however, these factors are typically missing from decisions regarding whether to prosecute. Further, the authors provided an algorithm to convert cumulative harms into procedural degrees

of severity; similar to how courts reason to determine the degree of severity for a crime.

6.2. Practical implications

The authors created a proposed operational decision-support framework for prosecutors and courts to utilize when determining the severity of crimes committed utilizing deepfakes. The authors' framework creates multiple pathways for processing evidence, charging and developing sentencing strategies depending on the scale of the harm caused by the crime as opposed to the type of content used to commit the crime. The implementation of the authors' framework should result in greater proportionality, consistency and predictability in the criminal justice response to synthetic media crimes.

7. References

BERRY, E. "Is misuse of deepfake technology adequately addressed by Australian intellectual property law, or does Australia need to introduce a right of publicity tort?" Intellectual Property Forum: Journal of the Intellectual and Industrial Property Society of Australia and New Zealand, v. 140, 2025, p. 7-17. Available at: https://search.informit.org/doi/abs/10.3316/informit.T2025061500000391498281295 (accessed on 18 January 2026).

BONDARENKO, S., et al. "Improving the state system of strategic planning of national security in the context of informatization of society", Journal of Information Technology Management, v. 14, 2022, p. 1-24. https://doi.org/10.22059/jitm.2022.88861

BUCKINGHAM, E. "The challenges of deepfake technology on the decision-making processes within law enforcement: A study within the EU landscape" (Master's thesis). University of Malta, 2025. Available at: https://www.um.edu.mt/library/oar/handle/123456789/141835 (accessed on 18 January 2026).

CELLI, F. "Deepfakes are coming: Does Australia come prepared?" Canberra Law Review, v. 17, 2020, 193. Available at: https://search.informit.org/doi/10.3316/agis.20211118057010 (accessed on 18 January 2026).

DARMA, M., et al. "Legal implications of deepfake technology misuse in digital content on social media", Science of Law, v. 3, 2025, p. 98-103. https://doi.org/10.55284/eqazc148

DE SOUZA, R. R. M. "Legal remedies and regulatory frameworks to combat ai-driven deepfakes", In Mitigating the risks of AI deepfakes, Boca Raton: CRC Press, 2026, p. 94–116. https://doi.org/10.1201/9781003539032-5

FARISH, K. "Do deepfakes pose a golden opportunity? Considering whether English law should adopt California's publicity right in the age of the deepfake", Journal of Intellectual Property Law & Practice, v. 15, n. 1, 2019, p. 40–48. https://doi.org/10.1093/jiplp/jpz139

FLYNN, A.; CLOUGH, J.; COOKE, T. "Disrupting and preventing deepfake abuse: Exploring criminal law responses to ai-facilitated abuse", in The Palgrave handbook of gendered violence and technology. Cham: Springer International Publishing, 2022, p. 583–603. https://doi.org/10.1007/978-3-030-83734-1_29

GAITIS, K. K., et al. "Legal challenges in tackling AI-generated CSAM across the UK, USA, Canada, Australia and New Zealand: Who is accountable according to the law?", Searchlight 2025–Who Benefits? Shining a Light on the Business of Child Sexual Exploitation and Abuse, 2025, p. 50-59. Available at: https://www.research.ed.ac.uk/en/publications/legal-challenges-in-tackling-ai-generated-csam-across-the-uk-usa-/ (accessed on 18 January 2026).

GOLDBERG, H. "States legislating against digital deception: A comparative study of laws to mitigate deepfake risks in American political advertisements", Notre Dame Journal on Emerging Techologies, v. 6, 1, 2025. Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4912795 (accessed on 18 January 2026).

HRISTOV, G. "Genuine harms behind artificial content: How EU regulation can combat malicious use of deep fake technology", SSRN Electronic Journal, 2025. https://doi.org/10.2139/ssrn.5634715

JASSERAND, C. "Deceptive deepfakes: Is the law coping with ai-altered representations of ourselves?" in 2024 International conference of the biometrics special interest group (BIOSIG). Darmstadt, Germany: IEEE, 2024, p. 1-4. https://doi.org/10.1109/biosig61931.2024.10786729

JI, S. "#MeToo in an AI-generated deepfake sexual violence era in South Korea", Women's Studies International Forum, v. 112, 2024, 103146. https://doi.org/10.1016/j.wsif.2025.103146

KADRI, T. E.; WEST, S. R. "Deepfake torts: Emerging tort frameworks in U.S. deepfake regulation", Journal of Tort Law, v. 18, n. 2, 2025, p. 515-552. https://doi.org/10.1515/jtl-2025-0032

KIM, K. "Deepfakes: Challenges to intellectual property rights in South Korea", GRUR International, v. 74, n. 6, 2025, p. 532-542. https://doi.org/10.1093/grurint/ikaf044

KIRCHENGAST, T. "Deepfakes and image manipulation: Criminalisation and control", Information & Communications Technology Law, v. 29, n. 3, 2020, p. 308–323. https://doi.org/10.1080/13600834.2020.1794615

KOPOTUN, I., et al. "Expanding the potential of the preventive and law enforcement function of the security police in combating cybercrime in Ukraine and the EU", TEM Journal, v. 9, n. 2, 2020, p. 460–468. https://doi.org/10.18421/tem92-06

LIN, L. S. F. "Organisational challenges in US law enforcement's response to ai-driven cybercrime and deepfake fraud", Laws, v. 14, n. 4, 2025, 46. https://doi.org/10.3390/laws14040046

LOSHYTSKYI, M., et al. "International legal standards for documentation and investigation of war crimes", Clio Revista De Historia Ciencias Humanas Y Pensamiento Critico, v. 5, n. 10, 2025, p. 1818–1855. https://doi.org/10.5281/zenodo.15598037

MARQUES MOREIRA, J. "The use of deepfakes in the criminal justice domain: An emerging reality?" SSRN Electronic Journal, 2025. https://doi.org/10.2139/ssrn.5329688

MARTIN, N. "Online safety regulation of deepfake abuse: A case study on Australia's eSafety Commissioner", Griffith Law Review, 2025, p. 1–24. https://doi.org/10.1080/10383441.2025.2504791

MEKKAWI, M. H. "The challenges of digital evidence usage in deepfake crimes era", Journal of Law and Emerging Technologies, v. 3, n. 2, 2023, p. 176–232. https://doi.org/10.54873/jolets.v3i2.123

MONTASARI, R. "Responding to deepfake challenges in the United Kingdom: Legal and technical insights with recommendations", in Advanced sciences and technologies for security applications. Cham: Springer International Publishing, 2024, p. 241–258. https://doi.org/10.1007/978-3-031-50454-9_12

PAVIS, M. "Rebalancing our regulatory response to Deepfakes with performers' rights", Convergence: The International Journal of Research Into New Media Technologies, v. 27, n. 4, 2021, p. 974–998. https://doi.org/10.1177/13548565211033418

PENG, H.; LEE, P.-W. "Reimagining U.S. tort law for deepfake harms: Comparative insights from China and Singapore", Journal of Tort Law, v. 18, n. 2, 2025, p. 579-607. https://doi.org/10.1515/jtl-2025-0028

RENAUD, L. "Will you believe it when you see it? How and why the press should prepare for deepfakes", Georgetown Law Technology Review, v. 4, 2019, 241. Available at: https://georgetownlawtechreview.org/will-you-believe-it-when-you-see-it-how-and-why-the-press-should-prepare-for-deepfakes/GLTR-01-2020/4 (accessed on 2 May 2025).

RODRIGUEZ, X. "Artificial intelligence (AI) and the practice of law in Texas", South Texas Law Review, v. 63, 2023, 1. Available at: https://texasbarsections.com/wp-content/uploads/2023/11/Rodriguez-Paper.pdf (accessed on 18 January 2026).

SANDOVAL, M.-P.; DE ALMEIDA VAU, M.; SOLAAS, J.; et al. "Threat of deepfakes to the criminal justice system: A systematic review", Crime Science, v. 13, 2024, Article 41. https://doi.org/10.1186/s40163-024-00239-1

SCHMITZ-BERNDT, S. et al. "Non-consensual deepnudes: Responses under EU law to a novel form of sexual abuse", International Review of Law, Computers & Technology, 2026, p. 1–29. https://doi.org/10.1080/13600869.2026.2654235

SHEEHY, S. "The effects and influence of artificial intelligence and likelihood of impacts to Australia", Journal of the Australian Institute of Professional Intelligence Officers, v. 29, n. 2, 2021, p. 3-9. https://doi.org/10.3316/informit.387429703808070

UGWUOKE, V.; SANFILIPPO, M. R. "The current landscape of deepfake legislation in the United States", Journal of Information Policy, v. 15, 2025, p. 97-129. https://doi.org/10.5325/jinfopoli.15.2025.0004

VYAS, A. D. "United States of Deepfake", Tennessee Law Review, v. 92, 2024, 307. Avaialble at: https://ssrn.com/abstract=4910852 (accessed on 18 January 2026).

YADAV, H.; OZA, J. "The deepfake dilemma: Balancing innovation, ethics, and accountability through law", Authorea Preprints, 2025. https://doi.org/10.36227/techrxiv.175744943.37338204/v1

YAVUZ, C. "Criminalization of the dissemination of nonconsensual deepfake pornography in the European Union. A comparative legal analysis", Research Day Law and Criminology, 2024, 2024. Available at: https://biblio.ugent.be/publication/01JMFTQENS8CCYFVRZGNVRE56V (accessed on 18 January 2026).

ZHENG, G.; SHU, J.; LI, K. "Regulating deepfakes between lex lata and lex ferenda – A comparative analysis of regulatory approaches in the U.S., the EU and China", Crime, Law and Social Change, v. 83, n. 1, 2025. https://doi.org/10.1007/s10611-024-10197-z