International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

AI-Generated Deepfake Fraud and Digital Identity Manipulation: Emerging Legal Challenges, Regulatory Gaps, and Governance Strategies

Rajat Umrey[1], Dr. Richa Singh[2]

[1]*Assistant Professor, Department of Management, Lakshmi Narain College of Technology University, Bhopal (M.P.), India*
[2]*Assistant Professor, Department of Legal Studies, Lakshmi Narain College of Technology University, Bhopal (M.P.), India*

***Abstract*—The rapid development of generative artificial intelligence has transformed deepfakes from experimental synthetic media into sophisticated instruments of impersonation, identity manipulation, and fraud. This study critically examines the legal and governance challenges arising from AI-generated deepfake fraud, with particular emphasis on digital identity, consent, attribution, financial deception, platform accountability, evidentiary authenticity, and cross-border enforcement. A qualitative doctrinal and comparative legal methodology is employed to analyse relevant legislation, regulatory instruments, policy documents, and interdisciplinary scholarship across selected jurisdictions. The findings reveal substantial regulatory fragmentation, particularly in defining unlawful synthetic impersonation, establishing intent and attribution, protecting biometric identity, allocating platform responsibility, and providing timely remedies to victims. Existing fraud, cybercrime, privacy, and data-protection frameworks address several consequences of deepfake misuse but do not consistently respond to its technological characteristics and transnational scale. The study proposes a layered, risk-based governance framework integrating content provenance, identity authentication, platform safeguards, proportionate liability, victim redress, and international regulatory cooperation to strengthen digital trust.**

***Keywords*-- Artificial Intelligence; Deepfake Fraud; Digital Identity; Generative AI; Synthetic Media; Cybercrime; AI Governance; Platform Accountability; Digital Trust**

I. INTRODUCTION

Generative artificial intelligence (AI) has fundamentally transformed the creation, manipulation, and dissemination of digital content. Contemporary generative models can produce highly realistic synthetic images, videos, voices, and audiovisual representations that closely resemble identifiable individuals. Commonly described as deepfakes, these outputs have legitimate applications in entertainment, education, accessibility, creative industries, and digital communication.

However, the same capabilities are increasingly associated with impersonation, identity manipulation, financial deception, social engineering, misinformation, reputational harm, and other forms of technology-enabled fraud.

The principal concern is not synthetic media itself, but its deceptive and unauthorized use. AI-generated voice cloning, facial manipulation, and fabricated audiovisual communication can enable offenders to impersonate executives, relatives, public officials, customers, or other trusted persons. Such techniques may be used to induce financial transfers, obtain confidential information, circumvent verification procedures, or manipulate individuals into taking actions they would otherwise reject. Consequently, digital identity can no longer be understood solely through passwords or conventional credentials; facial appearance, voice, behavioural characteristics, and other biometric attributes have also become vulnerable to synthetic replication.

These developments create significant legal challenges involving consent, attribution, fraudulent intent, privacy, biometric-data protection, electronic evidence, intermediary responsibility, jurisdiction, and victim remedies. Existing legal frameworks often regulate the consequences of fraud or identity misuse without specifically addressing the technological processes through which synthetic impersonation is generated and disseminated.

Against this background, the present study examines the adequacy of existing legal and regulatory approaches to AI-generated deepfake fraud. It identifies major regulatory gaps and comparatively evaluates emerging governance responses. The study further proposes a layered, risk-based governance framework combining legal accountability, technological safeguards, platform responsibility, identity authentication, victim protection, and cross-border cooperation while preserving legitimate uses of generative AI.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

II. Literature Review

Deepfake scholarship has developed from an early concentration on synthetic-media generation and forensic detection toward a broader examination of cybersecurity, identity, privacy, fraud, and governance. Technical studies have investigated spatial artefacts, temporal inconsistencies, physiological signals, audio anomalies, facial features, and multimodal approaches for distinguishing manipulated from authentic content. Nevertheless, continuous improvements in generative models have created an adversarial cycle in which detection techniques must repeatedly adapt to increasingly realistic synthetic outputs.

Recent literature consequently conceptualizes deepfakes as a socio-technical governance problem rather than exclusively a detection problem. AI-generated voice and video impersonation can strengthen social-engineering attacks by exploiting interpersonal trust and perceived audiovisual authenticity. The resulting risks extend to financial fraud, biometric spoofing, reputational damage, misinformation, and identity-related harms.

Legal scholarship further identifies difficulties involving consent, privacy, defamation, data protection, criminal attribution, evidentiary integrity, and intermediary liability. Conventional fraud and cybercrime provisions may apply where deception produces financial loss or unauthorized access, but they may provide less direct protection where harmful synthetic impersonation does not fit established offence categories.

The literature also reveals fragmentation between technical and legal responses. Detection, provenance, platform governance, criminal liability, and victim remedies are frequently examined independently. A significant research gap therefore concerns the development of an integrated governance model connecting these mechanisms across the deepfake lifecycle. The present study addresses this gap through comparative legal analysis and a multidimensional governance framework.

III. Materials And Methods

3.1 Research Design

The study adopts a qualitative doctrinal and comparative legal research design. This approach was selected because the research objective is not to measure the prevalence of deepfake fraud experimentally, but to assess the adequacy, scope, and limitations of existing legal and governance mechanisms applicable to AI-generated impersonation and digital identity manipulation.

The doctrinal component examines legal principles relating to fraud, cybercrime, privacy, data protection, biometric information, digital identity, electronic evidence, online intermediaries, and artificial intelligence governance. The comparative component evaluates how selected jurisdictions respond to these issues through both deepfake-specific measures and broader technology-neutral legislation.

3.2 Sources and Scope of Analysis

Primary sources include statutes, regulations, regulatory guidance, policy documents, and official AI-governance instruments. Secondary sources comprise peer-reviewed journal articles, academic literature, institutional reports, and interdisciplinary scholarship concerning generative AI, deepfakes, cybersecurity, synthetic identity, and digital fraud.

The comparative analysis focuses on the European Union, United States, United Kingdom, and India. These jurisdictions were selected because they illustrate different regulatory models, including comprehensive AI regulation, sectoral or state-level intervention, online-safety governance, and broader cybercrime and digital-data frameworks.

Table 1. Analytical Framework for AI-Generated Deepfake Fraud

Analytical Dimension	Key Issues Examined	Evaluation Criterion
Fraud and deception	Financial loss, impersonation, social engineering	Adequacy of legal coverage
Digital identity	Face, voice and biometric replication	Identity protection
Consent and privacy	Unauthorized synthetic representation	Rights protection
Attribution	Creator, distributor and beneficiary identification	Accountability
Platform governance	Detection, labelling, reporting and response	Intermediary responsibility
Electronic evidence	Authenticity, integrity and provenance	Evidentiary reliability
Jurisdiction	Cross-border generation and dissemination	Enforcement capacity
Victim remedies	Removal, reporting, compensation and redress	Remedial effectiveness

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

3.3 Analytical Procedure

The research proceeded through four stages. First, relevant legal and scholarly materials were classified according to their relationship with deepfake creation, dissemination, fraudulent use, and resulting harm. Second, thematic analysis was conducted to identify recurring legal problems. Six principal themes emerged: attribution and intent; consent and identity rights; platform accountability; evidentiary authenticity; jurisdiction and enforcement; and victim protection.

Third, comparative analysis examined the extent to which selected jurisdictions address these themes through existing or emerging regulatory mechanisms. The evaluation applied four criteria: preventive capacity, technological adaptability, enforceability, and availability of victim remedies.

Finally, the identified gaps were synthesized into a proposed risk-based governance model. Rather than treating every form of synthetic media as inherently unlawful, the framework distinguishes legitimate applications from deceptive or harmful uses according to factors such as absence of consent, fraudulent intent, impersonation, foreseeable harm, and scale of dissemination.

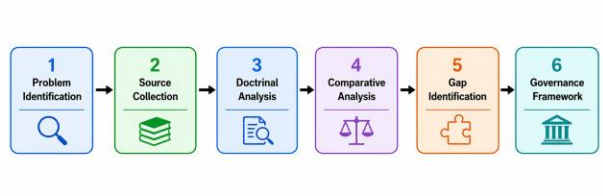

Figure 1. Research Methodology

This methodology enables the study to move beyond a purely descriptive account of deepfake regulation and critically evaluate how legal, technological, institutional, and remedial mechanisms can operate collectively.

IV. Results And Discussion

4.1 Fragmented Legal Protection

The analysis indicates that AI-generated deepfake fraud is not wholly outside existing law. Fraud, identity-related offences, privacy law, data protection, cybercrime provisions, and other established legal doctrines may apply depending on the conduct and jurisdiction. The central problem is therefore fragmentation rather than complete absence of regulation. Legal protection frequently depends on the consequences of a deepfake rather than the act of malicious synthetic impersonation itself.

This creates uncertainty where fabricated identity is used to manipulate or deceive but the conduct does not neatly satisfy traditional requirements for financial loss, unauthorized system access, defamation, or another established cause of action.

4.2 Attribution, Intent, and Digital Identity

Attribution emerged as a major enforcement challenge. Synthetic content can be generated anonymously, altered by multiple actors, and distributed rapidly across platforms and jurisdictions. Establishing the original creator, subsequent distributor, fraudulent beneficiary, and requisite state of mind may therefore require complex technical and evidentiary investigation.

Digital identity protection presents an equally significant concern. Facial appearance and voice increasingly function as authentication signals while simultaneously becoming reproducible through generative AI. Legal frameworks must consequently consider not only the collection and processing of biometric information but also its unauthorized synthetic reconstruction and deceptive use.

Table 2. Key Regulatory Gaps and Recommended Governance Responses

Regulatory Gap	Significance	Recommended Response
Ambiguous liability for malicious synthetic impersonation	High	Harm- and intent-based legal provisions
Inadequate protection against voice/facial replication	High	Stronger biometric and identity safeguards
Difficulty determining content origin	High	Provenance and content-credential standards
Inconsistent platform responsibilities	High	Proportionate risk-based platform duties
Cross-border enforcement difficulties	High	Regulatory and law-enforcement cooperation
Delayed victim remedies	High	Expedited reporting, removal and redress
Limitations of automated detection	Moderate–High	Multimodal detection plus human verification
Low public awareness	Moderate–High	AI and digital-media literacy programmes

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

Table 3. Comparative Regulatory Framework for AI-Generated Deepfake Fraud

Jurisdiction	Key Regulations	Main Strength	Key Gap
European Union	AI Act; DSA; GDPR	Deepfake transparency, privacy and platform accountability	Fragmented liability for deepfake fraud
United States	TAKE IT DOWN Act; FTC/FCC rules; state laws	Targeted protection against impersonation and harmful deepfakes	No comprehensive federal framework
United Kingdom	Online Safety Act; Fraud Act; data-protection laws	Online safety and fraud protection	Limited regulation of non-intimate synthetic impersonation
India	IT Act; DPDP Act; intermediary rules	Covers identity theft and digital personation	Limited deepfake-specific provisions

4.3 Platform Accountability and Authenticity

Online platforms occupy a critical position in the deepfake ecosystem because they facilitate dissemination at scale. A proportionate governance framework should therefore include accessible reporting mechanisms, preservation of relevant evidence, responses to verified fraudulent impersonation, and appropriate transparency concerning manipulated content.

However, platform obligations require careful calibration. Automatic removal of all synthetic content could interfere with legitimate satire, artistic expression, education, journalism, and consensual AI-generated media. Governance should therefore prioritize deception, absence of consent, fraudulent purpose, impersonation, and demonstrable or foreseeable harm rather than synthetic creation alone.

Another significant finding is that forensic detection should not operate as the sole defensive mechanism. As generation techniques become more sophisticated, detection performance may vary across models and contexts.

Content provenance, cryptographically supported credentials, watermarking, secure metadata, and stronger identity-verification procedures can provide complementary layers of authenticity assurance.

4.4 Proposed Integrated Governance Framework

The findings support a lifecycle-based approach in which responsibility is distributed across developers, platforms, financial institutions, regulators, law-enforcement bodies, and users.

Figure 2. Integrated Governance Framework for Deepfake Fraud

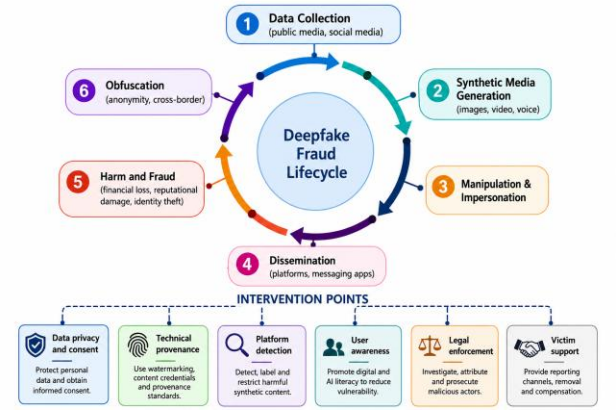

Figure 3. Deepfake Fraud Lifecycle and Intervention Points

This framework demonstrates that effective deepfake governance requires more than criminal prohibition. Preventive technical architecture can reduce opportunities for manipulation, while platform controls can limit dissemination. Legal rules can allocate responsibility according to intent and harm, and rapid-response mechanisms can reduce continuing victimization. International cooperation is particularly important because synthetic-media services, platforms, offenders, and victims may operate in different jurisdictions.

Accordingly, the preferred regulatory approach is layered, technology-adaptive, and risk-based. Such an approach avoids treating all synthetic media as harmful while imposing stronger obligations where AI is intentionally used for deceptive impersonation, identity manipulation, financial fraud, or comparable harms.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 15, Issue 08, August 2026)

V. Conclusion

AI-generated deepfake fraud presents significant challenges to digital identity, privacy, financial security, evidentiary trust, and regulatory enforcement. Existing laws provide partial protection but remain fragmented across legal domains and jurisdictions. An effective response requires a technology-adaptive and risk-based framework integrating provenance, identity authentication, platform accountability, proportionate liability, rapid victim redress, and cross-border cooperation. Such an approach can target malicious synthetic impersonation without unnecessarily restricting legitimate uses of generative AI, thereby strengthening digital trust while supporting responsible technological innovation.

REFERENCES

[1] Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. California Law Review, 107(6), 1753–1820.

[2] Coalition for Content Provenance and Authenticity. (2024). C2PA technical specification (Version 2.0).

[3] European Parliament and Council of the European Union. (2016). Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). Official Journal of the European Union, L 119, 1–88.

[4] European Parliament and Council of the European Union. (2022). Regulation (EU) 2022/2065 on a single market for digital services and amending Directive 2000/31/EC (Digital Services Act). Official Journal of the European Union, L 277, 1–102.

[5] European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union.

[6] Fallis, D. (2021). The epistemic threat of deepfakes. Philosophy & Technology, 34(4), 623–643. https://doi.org/10.1007/s13347-020-00419-2

[7] Government of India. (2000). The Information Technology Act, 2000 (Act No. 21 of 2000). Ministry of Law and Justice.

[8] Government of India. (2023). The Digital Personal Data Protection Act, 2023 (Act No. 22 of 2023). Ministry of Law and Justice.

[9] Kaur, A., Noori Hoshyar, A., Saikrishna, V., Firmin, S., & Xia, F. (2024). Deepfake video detection: Challenges and opportunities. Artificial Intelligence Review, 57, Article 159. https://doi.org/10.1007/s10462-024-10810-6

[10] Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? Business Horizons, 63(2), 135–146. https://doi.org/10.1016/j.bushor.2019.11.006

[11] Killion, V. L. (2025). The TAKE IT DOWN Act: A federal law prohibiting the nonconsensual publication of intimate images (Legal Sidebar LSB11314). Congressional Research Service.

[12] Mirsky, Y., & Lee, W. (2021). The creation and detection of deepfakes: A survey. ACM Computing Surveys, 54(1), Article 7. https://doi.org/10.1145/3425780

[13] Mubarak, R., Alsboui, T., Alshaikh, O., Inuwa-Dute, I., Khan, S., & Parkinson, S. (2023). A survey on the detection and impacts of deepfakes in visual, audio, and textual formats. IEEE Access, 11, 144497–144529. https://doi.org/10.1109/ACCESS.2023.3344653

[14] Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). Deepfakes and beyond: A survey of face manipulation and fake detection. Information Fusion, 64, 131–148. https://doi.org/10.1016/j.inffus.2020.06.014

[15] United Kingdom. (2006). Fraud Act 2006. The National Archives.

[16] United Kingdom. (2023). Online Safety Act 2023. The National Archives.

[17] United States Congress. (2025). Tools to Address Known Exploitation by Immobilizing Technological Deepfakes on Websites and Networks Act (TAKE IT DOWN Act), Pub. L. No. 119-12.

[18] Verdoliva, L. (2020). Media forensics and deepfakes: An overview. IEEE Journal of Selected Topics in Signal Processing, 14(5), 910–932. https://doi.org/10.1109/JSTSP.2020.3002101

[19] Viola, M., & Voto, C. (2023). Designed to abuse? Deepfakes and the non-consensual diffusion of intimate images. Synthese, 201, Article 30. https://doi.org/10.1007/s11229-022-04012-2

[20] Westerlund, M. (2019). The emergence of deepfake technology: A review. Technology Innovation Management Review, 9(11), 39–52.