



T2026_01 Understanding the Dangers of Deepfakes (March 23rd, 2026)

Deepfakes are synthetic media created using artificial intelligence that can convincingly manipulate or fabricate audio, images, or videos to make individuals appear to say or do things they never actually did. While the technology has legitimate uses in entertainment and research, it also presents serious cybersecurity and information-integrity risks.

Malicious actors can use deepfakes to spread disinformation, conduct fraud, impersonate executives in social-engineering attacks, or damage reputations. For example, attackers have used AI-generated voice deepfakes to impersonate company executives and trick employees into transferring funds or revealing sensitive information. Deepfakes can also undermine trust in legitimate media, making it harder for the public and organizations to distinguish real evidence from fabricated content.

To reduce risk, organizations and individuals should verify unusual requests through secondary communication channels, educate staff about deepfake-enabled scams, implement strong identity verification procedures for sensitive transactions, and rely on trusted sources when consuming or sharing media. Developing awareness and verification practices is critical as AI-generated content becomes more realistic and widely accessible.

References

- National Institute of Standards and Technology. (2023). Evaluating Deepfake Detection Methods in Media Forensics. NIST. Retrieved from <https://www.nist.gov>
- Europol. (2022). Facing Reality? Law Enforcement and the Challenge of Deepfakes. Europol Innovation Lab. Retrieved from <https://www.europol.europa.eu>
- Federal Bureau of Investigation. (2023). Public Service Announcement: Criminals Use AI-Generated Content for Fraud and Social Engineering. FBI. Retrieved from <https://www.ic3.gov>
- Chesney, R., & Citron, D. (2019). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *California Law Review*, 107(6). Retrieved from <https://papers.ssrn.com>
- MIT Media Lab. (2020). The State of Deepfakes: Landscape, Threats, and Impact. MIT. Retrieved from <https://www.media.mit.edu>