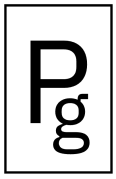
Deepfake Risks for Law Firms

A Whitepaper for Partners and CIOs

by Polyguard, Inc. and Breacher.ai,

May, 2025



Executive Summary

Law firms face rising threats from deepfake attacks that compromise client confidentiality, enable financial fraud, falsify evidence, damage reputations, and trigger malpractice claims. These threats are no longer hypothetical—threat actors are already using deepfakes to impersonate clients, partners, and judges, infiltrate meetings, and reroute funds.

The American Bar Association (ABA) has warned attorneys to "stay informed about deepfake generation as it increases in sophistication or else risk allowing manipulated materials to improperly influence their advice and advocacy."[1] In response, this whitepaper outlines the legal sector's deepfake risk landscape—beginning with impersonation and wire fraud—and offers actionable, prevention-focused recommendations.

Polyguard protects Zoom meetings and other sensitive communications using real-time, cryptographically secure identity verification. Unlike phone or email based verifications, Polyguard verifies every participant continuously, blocking unauthorized access in real time.

Deepfake-enabled fraud grew by over 1300% in 2024[2], with law firms among the top targets due to their control over high-value data and transactions. Polyguard mitigates these risks by:

- Verifying identities using government-issued ID (e.g., driver's license, passport) and live biometrics
- Requiring bi-directional verification for sensitive instructions
    - Training legal teams to detect and respond to deepfake threats

[1] American Bar Association, "Deepfakes and Malpractice Risk: Lawyers Beware" (Spring 2024)
[2] Cyber Daily, "Fake views: Deepfake fraud surged by 1,300% in 2024" (Summer 2025)

Embedding continuous identity attestation into legal workflows enables firms to prevent attacks before they happen—and meet growing regulatory and professional standards.

Introduction

Deepfake technology—leveraging generative AI to synthesize realistic audio, video, and images—has matured rapidly. Cyber adversaries can now impersonate clients, partners, or even judges with startling fidelity, creating new attack vectors that bypass traditional authentication controls. For law firms, which routinely handle privileged communications and large-value transactions, the stakes are exceptionally high. If unaddressed, deepfake-enabled fraud can lead to direct financial losses, breach of attorney-client privilege, erosion of client trust, and potential malpractice liability.

This whitepaper focuses on the following core threat categories:

1. **Client Impersonation & Confidentiality Breaches**
2. **Financial Fraud & Wire Transfer Exploits**
3. **Evidentiary Integrity & Litigation Risks**
4. **Reputational Damage & Client Trust Erosion**
5. **Professional Liability & Malpractice Exposure**
6. **Emerging Threats in Vendor & Third-Party Interactions**

Each section details how deepfakes manifest in a legal context, illustrates potential impact, and provides prevention-focused guidance tailored for the Chief Information Officer (CIO) or Chief Information Security Officer (CISO).

1. Client Impersonation & Confidentiality Breaches

1.1 Threat Overview

Deepfake-enabled client impersonation can occur via synthesized audio in calls, doctored video in virtual meetings, or falsified documents bearing forged signatures. Attackers impersonate legitimate clients to request privileged information—case strategies, settlement terms, or personal data—thereby compromising attorney-client privilege and exposing sensitive data to unauthorized parties. Such breaches undermine confidentiality obligations under ABA Model Rule 1.6 and risk violation of data protection laws (e.g., GDPR, CCPA) when personal client data is exposed.

1.2 Impact

- **Confidentiality Violation**: Unauthorized access to privileged communications can lead to reputational damage and loss of client confidence.
- **Regulatory Penalties**: Exposure of personal data may trigger fines under GDPR or CCPA, especially if firms fail to report breaches promptly.
- **Case Strategy Disruption**: Opposing counsel could exploit leaked information, potentially altering the outcome of litigation or negotiations.

1.3 Risk Factors

- **High-Value Matters**: Corporate M&A, high-stakes litigation, and white-collar investigations involve sensitive, high-value data. Adversaries targeting these matters stand to gain significantly.
- **Remote/Urgent Communications**: Lawyers often conduct interviews or strategy sessions via video conference under tight deadlines, increasing the likelihood that deepfakes will bypass cursory verification.
- **Overreliance on Known Channels**: Trusting email and phone calls alone is insufficient; adversaries craft deepfakes that mimic a familiar voice or visual appearance, outwitting multi-factor authentication (MFA) that doesn't incorporate live biometric checks.

2. Financial Fraud & Wire Transfer Exploits

2.1 Threat Overview

Law firms frequently manage escrow accounts or facilitate wire transfers for real estate closings, M&A deals, and settlement distributions. Deepfake-generated voices or videos can impersonate corporate executives or clients to urgently instruct finance teams to reroute funds to attacker-controlled accounts. In one documented incident, fraudulent synthetic audio and video led to a multimillion-dollar transfer before the firm detected inconsistencies, underscoring the limitations of standard verification processes.[3]

2.2 Impact

- **Direct Financial Loss**: Firms are often responsible for reimbursing clients when escrow or settlement funds are misappropriated due to fraud.
- **Operational Disruption**: Investigations tied to large, suspicious transactions can divert IT and finance resources, delaying legitimate client work.
- **Insurance and Premium Increases**: Cyber-insurance carriers may raise premiums or impose stricter underwriting requirements if firms lack robust identity-verification controls.

2.3 Risk Factors

- **Time Sensitivity of Transactions**: Wire instructions often include tight deadlines—attackers exploit urgency to bypass even manual cross-checks.
- **Complex Transaction Chains**: Multi-party closings (e.g., involving lenders, title companies, escrow agents) create multiple verification handoffs where deepfakes can be inserted.
- **Insufficient Verification of In-Call Identity**: Traditional voice-only or email-based OTP methods cannot distinguish a live client from a synthesized audio stream.

[3] WEForum, "This engineering firm was hit by a deepfake fraud. Here's what it learned"

3. Evidentiary Integrity & Litigation Risks

3.1 Threat Overview

In litigation, audio and video recordings are powerful evidence. Deepfake technology enables adversaries to fabricate "evidence" depicting clients or witnesses making admissions, entering agreements, or committing actions that never occurred. Conversely, genuine recordings may be dismissed as "fake," creating a "liar's dividend" that muddies the truth. As one legal expert noted, "Deepfakes will impact evidence authenticity, witness credibility, and the integrity of the judicial process".[4]

3.2 Impact

- **Extended Litigation Timelines**: Forensic experts may be required to validate any audio/video recording, adding cost and delay to proceedings.

- **"False Flag" Tampering**: Construction of synthetic versions of legitimate evidence can be used to undermine the credibility of such evidence, a new form of AI-era evidence tampering.
- **Increased Defense Costs**: Even when deepfakes are debunked, the expense of forensic analysis and expert testimony can be substantial.
- **Judicial Skepticism**: Judges and juries may become more distrustful of digital evidence, hindering cases reliant on recorded testimony or surveillance footage.

3.3 Risk Factors

- **Rise of Generative AI**: Tools that produce high-fidelity deepfakes are increasingly accessible to non-technical threat actors.
- **Lack of Standardized Authentication**: Courts have yet to adopt consistent protocols for authenticating digital evidence, leaving individual firms to develop ad hoc processes.
- **High Stakes in White-Collar Defense**: Defendants in corporate fraud or criminal

[4] Illinois State Bar Association, "Deepfakes in the Courtroom: Problems and Solutions" (Mar 2025)

matters may invest heavily in synthetic evidence to fabricate alibis or mislead investigators.

4. Reputational Damage & Client Trust Erosion

4.1 Threat Overview

Deepfake attacks on a firm's public image can take the form of fabricated videos showing partners or associates making unethical comments or disclosing confidential client matters. Such content, once published on social media or disseminated to journalists, can spread rapidly. Even after forensic analysis proves the material is fake, the initial damage to the firm's reputation is often irreversible.

4.2 Impact

- **Client Attrition**: High-profile clients may lose confidence and seek new counsel if they believe their sensitive matters are not secure.
- **Recruiting Challenges**: Top legal talent may hesitate to join a firm perceived as vulnerable to sophisticated social engineering campaigns.
- **Media Scrutiny**: Negative press cycles can strain marketing and PR resources, requiring sustained mitigation efforts.

4.3 Risk Factors

- **Open-Source Footprint**: Law firms often share partner profiles, firm logos, and office imagery online, giving attackers raw material to train deepfake generators.
- **Speed of Social Media**: Viral spread of short-form videos or GIFs means a deepfake can reach thousands of stakeholders before containing the narrative.
- **Limited Crisis-Response Protocols**: Many firms lack pre-approved workflows for verifying and refuting deepfake content under tight time constraints.

5. Professional Liability & Malpractice Exposure

5.1 Threat Overview

If an attorney fails to implement reasonable measures to verify client identity—especially for high-value transactions—a deepfake impersonator could instruct wire transfers or request privileged information, triggering a malpractice claim. Failure to incorporate robust identity authentication protocols may constitute professional negligence if deepfakes lead to client harm.

5.2 Impact

- **Malpractice Lawsuits**: Clients who suffer financial loss or exposure of privileged data due to forged instructions may sue the firm for negligent verification practices.
- **Bar Disciplinary Actions**: State bar associations could censure attorneys for failing to safeguard client interests when deepfake risks are foreseeable.
- **Insurance Premium Adjustments**: Firms with claims for deepfake-enabled malpractice may see rapid increases in professional liability insurance costs.

5.3 Risk Factors

- **Ambiguous Standard of Care**: As deepfake risks are relatively new, courts may require firms to adopt industry-leading technologies to meet the “reasonable care” standard.
- **Complex Transaction Flows**: In multi-jurisdictional matters, attorneys rely on remote communications, where reliance on unverified video calls heightens exposure.
- **Increasing Regulatory Scrutiny**: Regulators may mandate stronger anti-fraud controls for regulated industries (e.g., financial services) that intersect with legal work.

6. Emerging Threats in Vendor & Third-Party Interactions

6.1 Threat Overview

Attackers can target outside counsel, escrow agents, title companies, and other third parties by creating deepfake videos or audio that mimic authorized representatives, providing falsified approvals or contract amendments. In cross-border transactions, adversaries may exploit language barriers and time-zone differences, interjecting themselves into critical communications where verification is more challenging. "Conveyancing and property transactions, where the verification of identities and documents is paramount, are particularly exposed to potential fraudulent activity"[5].

6.2 Impact

- **Erroneous Contractual Changes**: Malicious actors may insert or modify clauses, leading to unintended obligations or financial liabilities.
- **Unauthorized Data Sharing**: Deepfakes could be used to authorize disclosures of privileged information to external advisors.
- **Supply Chain Disruption**: Relying on compromised third parties can halt deal closings and extend project timelines, damaging client relationships.

6.3 Risk Factors

- **Distributed Ecosystems**: Modern legal matters often involve multiple stakeholders—external advisors, co-counsel, and vendors—each presenting their own authentication challenge.
- **Inconsistent Verification Practices**: Third parties may lack standardized identity protocols, making it difficult for firms to ensure end-to-end trust.
- **Global Transactions**: Cross-border matters introduce regulatory complexities and amplify the difficulty of rapid, in-person identity checks.

[5] Lockton, "Deepfake and the risk of vendor fraud: challenges and solutions for solicitors," Lockton

Key Recommendations (Prevention-Focused)

1. Implement Multi-Modal Verification

- Combine government-issued ID verification with live facial recognition in every high-risk video call or document exchange.
- Require bi-directional biometric checks for any instructions related to client funds or privileged data, ensuring the requestor is physically present.

2. Embed Continuous Identity Attestation

- Adopt a solution that issues cryptographically signed attestations at the start of each critical interaction and periodically re-validates participants throughout the session.
- Ensure that every attendee's identity is backed by a tamper-resistant record, so deepfake-generated intruders cannot join mid-stream.

3. Train Staff on Deepfake Risks

- Create a "Deepfake Response Playbook" with clear escalation paths when suspicious counterparties are identified.

4. Update Engagement Letters & Risk Disclosures

- Explicitly state in new client agreements that remote instructions for wire transfers or document access must pass live identity verifications.
- Clarify responsibilities for both firm and client in confirming authenticity, reducing ambiguity in potential malpractice claims.

5. Expand Vendor & Third-Party Protocols

- Require all outside counsel and third-party service providers to adhere to the firm's deepfake prevention standards, including multi-factor and biometric verification before releasing sensitive data or approving changes.
- Institute a centralized "Trusted Partner Registry" that catalogs verified points of contact for key vendors.

6. Develop Rapid Incident-Response Processes

- Establish a cross-functional task force (IT, security, legal, compliance, and PR) to coordinate immediate action when a suspected deepfake breach occurs.

- Pre-position forensic contacts—such as trusted AI-forensics firms—for expedited analysis if needed, ensuring swift rebuttal of false evidence.

Conclusion

Deepfakes represent a paradigm shift in how adversaries can compromise law firm operations—eroding confidentiality, enabling financial fraud, undermining evidentiary trust, and exposing firms to malpractice risks. Traditional controls like passwords, email OTPs, or static MFA are no longer sufficient. The ABA's guidance underscores the urgency for law firms to proactively adopt stronger, live identity verification protocols.

By embedding continuous, cryptographically verifiable identity attestations into every high-risk interaction—whether client interviews, document requests, or escrow instructions—firms can prevent deepfake-driven attacks at the outset. Implementing these prevention-focused measures will not only safeguard client trust and firm reputation but also align with evolving regulatory expectations and rising professional liability standards.

Polyguard, Inc.
505 W 43rd St, 3J,
New York, NY
(844) 671-0790
info@polyguard.ai

Breacher.ai
6100 Waterfield Way,
St Cloud, FL, 34771
(407) 900-0799
support@breacher.ai

"Breacher ran highly effective, targeted simulations for us using convincing deepfake audio and video of our CEO. This quickly pinpointed our vulnerabilities whilst educating the right employees."

- CISO, US Bank

"Polyguard is emerging as a critical shield against the rising tide of AI-powered fraud. Co-founders Joshua McKenty and Khadem Badiyan are at the forefront of combating sophisticated scams that leverage deepfakes, caller ID spoofing, and generative AI."

- David Marshall, VMBlog