Revista de Estudos Constitucionais, Hermenêutica e Teoria do Direito (RECHTD)
17(2):201-217, maio-agosto 2025
Unisinos - doi: 10.4013/rechtd.2025.172.04

Truth in the age of synthetic voices: constitutional and private-law remedies for deepfake financial fraud

A verdade na era das vozes sintéticas: soluções constitucionais e de direito privado para fraudes financeiras deepfake

Rizaldy Anggriawan[1]
University of Szeged (Hungary) / Universitas Muhammadiyah Yogyakarta (Indonesia)
rizaldyanggriawan@umy.ac.id

Abstract

This study analyzes the emergent threat of deepfake financial fraud, where generative AI technologies allow actors to impersonate others, fabricate identities, and interfere with financial transactions. The objective is to examine the constitutional, statutory, and private-law remedies that are available for this newly developed form of deception that is evolving rapidly. Utilizing recent case law, regulatory guidance, and comparative state and international responses, the analysis assesses whether our existing legal frameworks sufficiently incurred liability on platforms, made evidentiary standards flexible, and enforced anti-money laundering regulations and know your customer requirements. The analysis will also assess private-law remedies, such as fraud, negligence, and misappropriation claims. The analysis acknowledges that there are fragmented, albeit inadequate, responses from courts, regulators, and legislatures addressing the application of deepfakes and its threats, but there remains no comprehensive system to addresses the dimensions of deepfake threats. This study recommends that a collaborative system combining regulatory oversight, private-law accountability and technological infrastructure is necessary to protect private individuals and the entire financial markets.

Keywords: Accountability; Deepfakes; Financial Fraud; Platform Liability; Regulation.

Resumo

Este estudo analisa a ameaça emergente de fraude financeira deepfake, na

[1] PhD in Law at the Faculty of Law and Political Sciences, University of Szeged, Hungary. Lecturer in the Faculty of Law, Universitas Muhammadiyah Yogyakarta, Jalan Lingkar Selatan, Tamantirto – Kecamatan Kasihan, Bantul, 55183, Yogyakarta, Indonesia.

qual tecnologias generativas de IA permitem que atores se façam passar por outros, fabriquem identidades e interfiram em transações financeiras. O objetivo é examinar os recursos constitucionais, estatutários e de direito privado disponíveis para essa forma de fraude recém-desenvolvida e em rápida evolução. Utilizando jurisprudência recente, orientações regulatórias e respostas comparativas estaduais e internacionais, a análise avalia se nossos arcabouços jurídicos existentes incorreram suficientemente em responsabilidade nas plataformas, flexibilizaram os padrões probatórios e aplicaram regulamentações antilavagem de dinheiro e os requisitos do "conheça seu cliente". A análise também avaliará recursos de direito privado, como alegações de fraude, negligência e apropriação indébita. A análise reconhece que há respostas fragmentadas, embora inadequadas, de tribunais, reguladores e legisladores abordando a aplicação de deepfakes e suas ameaças, mas ainda não há um sistema abrangente para abordar as dimensões das ameaças de deepfakes. Este estudo recomenda que um sistema colaborativo que combine supervisão regulatória, responsabilização do direito privado e infraestrutura tecnológica seja necessário para proteger indivíduos privados e todo o mercado financeiro

Palavras-chave: Responsabilidade; Deepfakes; Fraude Financeira; Responsabilidade de Plataformas; Regulamentação.

Introduction

Artificial intelligence has initiated a new era of synthetic media with voice and facial deepfakes so realistic that truth and fabrication are increasingly indistinguishable. In this new era, technology has very quickly become a tool for financial crime: fraudsters impersonate a CEO, relative, or public official to trick a victim into wiring funds or disclosing sensitive information. The number of reported instances has risen sharply. In 2024 alone, FinCEN reported a sharp rise in identity fraud cases that involved AI-generated photos and identification documents (Fincen, 2024). Analysts predict staggering losses: FS-ISAC (2024) estimates that fraud using deepfakes will cost U.S. institutions more than $40 billion a year, by 2027. Criminals are also creating wholly synthetic identities to open accounts, launder the proceeds, and circumvent compliance controls. A Federal Reserve report shows that deepfake attacks rose twentyfold in just three years, with different schemes ranging from phishing calls, deepfakes used in video meetings (including a UK firm that lost $25 million after being called by a deepfake impersonating the CEO), and advertisements generated by AI that are fraudulent. Over one in ten reported attempts to defraud them using AI-generated voices and/or videos in a 2024 survey (Barr, 2025).

Simultaneously, enforcement is difficult due to the cross-border nature of these schemes. Fraud rings set up offshore call centers and tie (often with consent) calling targets to send funds to international money-mule networks, making it exceedingly difficult to trace the funds or arrest the fraudsters. As Button noted (2025), the cross-border nature of many scams makes it difficult for law enforcement agencies to track and catch those perpetrating the fraud. These developments raise fundamental values: persons have an emerging right to

security of person and property in response to new threats created by AI. Many international human-rights instruments already recognize these rights, while the U.S. Constitution does not have an explicit identification of a general right to security, due process and equal protection limit what harm the State can tolerate. Deepfake financial fraud is commercial deception rather than an expression of political speech (Kaushik et al., 2024); thus, it is outside of the protections of the First Amendment. Courts have consistently held that falsity, in the context of misleading a consumer or investor, is a sufficient predicate for regulation and liability.

Legal actions are rapidly expanding. A few states, such as Washington and Pennsylvania, have passed laws barring the use of deepfakes for defrauding, coercing, or stealing money. Penalties for these laws, though significant, provide a safe harbor for platforms and those involved in communications and marketing who did not know about the content and acted after notice. This indicates the emphasis on punishing the individual rather than burdening the intermediary with blanket affirmative obligations (Avsec and Michaels, 2025). At the federal level, there are bipartisan bills, much like the Preventing Deep Fake Scams Act of 2025, that call for joint enforcement between the Treasury and the financial regulators to regulate cross-border payment systems and anti-money laundering (AML) compliance, as those frameworks would be essential to any lasting solutions (Bracken, 2025).

Historically, fraud has evolved in tandem with the available technologies of communication and identity verification. If we look back across centuries, we see that fraud was only the agency of the human, primarily through means such as forged signatures, counterfeit seals or invalid contracts. The revolution of digital communication brought new devices of fraud such as phishing emails, fake websites, and compromised credentials: law enforcement has often responded to these, with inconsistent success. So-called deepfakes present us with the latest and arguably most pernicious phase of fraud along our trajectory toward deception. While previous forgeries of various forms could sometimes be detected through technological tools or, at a minimum, through expert inspection, deepfakes blur the line between what is real and what is false faster and more realistically than anything flown before. They take advantage of our (possibly innate) tendency to trust visual and auditory cues, which are only now becoming skewed as indicators of truth and authenticity.

Given even just these practical implications considered in isolation, the theoretical implications are significant. While traditional forms of fraud might be construed as singular incidents (even when repeated), deepfake fraud can be automated, and in its scale, it can involve thousands of potential victims at once. Beyond stealing financial capital, it undermines the trust to communicate as a good. Scholars have suggested that the "trust crisis" engendered by a society reliant on synthetic media might ultimately prove to be far more impactful than the loss of financial capital, because if every voice and every image can be immediately denied as fake, we may, in fact, be living in the "post-evidentiary age" as a society. In this sense, deepfake fraud - while novel - is not just a novel and complex form of deception; it is more of a systemic challenge to truth conceived as a good, in the public interest.

In this context, this study shifts to the questions that have implications for future policy and legal interpretations. The key issues consider whether online platforms have an affirmative obligation to restrict or prevent financial scams to deliver advertising or whether

this can even be done. Courts would need to consider what evidence is relevant when dwelling on the legitimacy or manipulation of deepfake media. Additionally, what regulators are able to exercise jurisdiction to adapt the issues of KYC and AML to AI fraud? Addressing these issues will utilize recent case law, guidance from regulators, empirical literature, and a comparative analysis of some states in the United States and other countries.

Platform Liability and Public-Duty to Prevent Fraud

(a) Section 230 and the Current Immunity Landscape

Under the current legal framework of the United States, online platforms have extensive protections with regard to liability for user-generated content (Asadi, 2023). Section 230 of the Communications Decency Act provides online intermediaries the protection from being treated as the publisher of third-party posts. Furthermore, courts have consistently interpreted Section 230 and used it as a bar for tort claims - defamation, invasion of privacy, product liability, etc. - anytime the harm occurred by user generated content. In the deepfake context this translates to online platforms (Facebook, YouTube, etc.) will typically not be liable for the posting of a malicious deepfake advertisement or video by a third party. An analysis by Edwards and Baker (2025) summarizes the law pertaining to tort claims in this context: Section 230 bars defamation and similar tort claims. E-commerce companies can use Section 230 as a bar to product liability and breach of warranty claims when the alleged harm is caused by a deepfake video posted by a user. Without new legislation, the response of victims of deepfakes will be that the law is very reluctant to impose a duty on platforms to screen third-party speech.

This immunity is not without limits, and as a judge in federal court recently noted, while the platform must be treated as publishers of content, if they do participate in the creation or sponsorship of content, this immunity wanes. In particular, *Forrest v. Meta* (N.D. Cal. 2024), Judge Pitts denied Meta's motion to dismiss a deepfake-advertisement lawsuit entirely based on Section 230 immunity. The court held that while Section 230 will bar tort claims against platforms, the plaintiff's amended complaint could plausibly demonstrate that Facebook's ad tools contributed to the content of the advertisements. More importantly, the court's order also allowed the plaintiff to argue that Meta breached a duty to operate in a commercially reasonable manner by failing to remove fraudulent ads. This was hailed by commentators as a clear dividing line from traditional media while noting allow us to argue that Facebook has the capability and the responsibility to prevent fraudulent ads from appearing on their platform (United States District Court N.D. California, 2024). The Forrest case has tentatively chipped a narrow opening in this immunity if a platform is actively monetized and targeting ads, a duty of or could be held liable for negligence in the platform's operation of the ads in the business context.

(b) Proposals for a Platform Duty (Legislative and Regulatory)

In keeping with this ongoing debate, proposed legislation suggests that platforms be shielded from liability for deepfakes and advertisements. A reintroduced bill, the No FAKES Act (2024–25), would establish a federal right of publicity for digital replica and require platforms to remove use of unauthorized voice/image replicas (Łabuz, 2023). It would importantly create the first ever federal right of publicity and find a narrow exception to the immunity of Section 230 in order to hold platforms liable for knowingly failing to remove deep fakes. Likewise, as reported by Politico, there is allegedly a bipartisan House bill to condition immunity under Section 230 based on platforms' attempts to prevent deepfakes (Citron; Wittes, 2018). The Federal Trade Commission (FTC) has indicated action, too: in early 2024, the FTC proposed an expansion of its impersonation-fraud rule under Section 5 to include platforms using AI. The proposed rule would hold a tech company liable if it knows or has reason to know that its generative AI is used by bad actors to impersonate someone for commercial purposes. This is likely to lead to liability for platforms that know they are providing means and instrumentalities of deepfake scams. Collectively, these proposals represent a shift in policy: platforms should not be provided with an absolute safe harbor for wrongful impersonation that is used for fraud.

None of the proposals have yet passed, and there are legitimate concerns over First Amendment and innovation implications. Specifically, there is concern that imposing a broad duty to prevent harm will lead to censorship and over-removal of content. The No FAKES Act itself was amended to guarantee that satire, commentary, and other similar socially valuable uses could not be removed legally (Gollin, 2025). State laws in Washington and Pennsylvania are more cautious: both laws criminalize malicious deepfakes, but expressly provide immunity for platforms/content hosts that do not intentionally facilitate the fraud. Washington's law, for example, provides an interactive computer service with immunity for forged content if the service promptly removed the content after notice that it was forged. Pennsylvania provides similar immunity for information service providers as long as they did not facilitate the creation and include prompt takedown or disclaimer on the fraudulent content. These provisions reflect a compromise: platforms will not be held responsible for every ad, but they will lose immunity if they have actual knowledge of the ad and do not properly respond.

Should platforms bear a public-law duty to prevent financial harm for scam ads? is still an open question. In terms of research, it is well-established that platforms have unique abilities (ad algorithms, user data, detection tools) to prevent harmful usages of the platform. In terms of ethics, the continuation of the economic model based on advertisement suggests that the platforms should acknowledge some social cost of preventing fraud (De Rancourt-Raymond and Smaili, 2023). The real question is whether Congress would require Section 230 or equivalent statutes to be amended by imposing some sort of duty. Outside of that amendment, the potential of private law remains limited. Until that occurs, the trend will remain incremental: while platforms are expected to police themselves (via Terms of Use, AI detection protocols), regulators continue to scrutinize specific categories of ads (e.g. crypto

scams), and courts like Forrest will continue to determine the bounds of negligence liability. In short, victims probably will have new causes of action (or at least defensible arguments) that a platform be reasonably expected to intervene to block the ad, but as long as the proposals remain just proposals, no general affirmative duty will emerge as law anytime soon.

(c) Comparative International Approaches

Despite U.S. discussions largely centering on Section 230 immunity and federal-state divisions, other countries have taken more aggressive steps to regulate synthetic media. The European Union's AI Act (2024) introduces conditional transparency obligations, which require that AI-generated content, including audio-visual deepfakes, be labelled as such if it is distributed to the wider public. While these transparency conditions are primarily directed at the issues of consumer protection and misinformation, they pertain directly to financial fraud, since AI-created synthetic identities who remain unlabelled as AI should lead to liability for both user and providers under EU law.

China has adopted a regulatory scheme that is often touted as one of the most stringent in the world, as well. In 2023, the Cyberspace Administration of China ordered that all forms of "synthetic media" must contain an indelible watermark indicating that the content was AI-generated. Service providers will be required to register users, verify identities, and create audit logs of records to allow for traceability. Violations can lead to fines or suspension of the ability to do business in the country. Critics have raised concerns about censorship, but this regulatory scheme provides an illustrative example where a centralized regulatory model seeks to address potential misuse at scale.

In South Asia, the Reserve Bank of India (RBI) circulated warnings to financial institutions regarding AI impersonation fraud advising institutions to implement enhanced customer verification and monitoring process regarding fraud. Even though the guidelines issued by the RBI are not yet enshrined as statutory law, the very emergence of guidelines displays recognition of deepfakes as a systemic financial risk in developing markets where adoption of digital payment systems is happening quickly.

These examples suggest that while U.S. law, at a federal level, remains exceedingly fragmented, there is increasing recognition of deepfake fraud both as a consumer protection issue and a financial integrity issue, as evidenced through these regulatory approaches, and points to possible avenues for transnational cooperation against deepfake fraud.

Evidentiary Standards for Alleged Deepfakes

As deepfake technology emerges in various litigation or possible enforcement, courts will face challenges regarding the authenticity of deepfake evidence. A new potential risk arising from deepfakes is that disputants will claim deepfake in order to challenge authentic audio and video evidence, or conversely, malicious actors will deny anything is real by claiming that it was a deepfake (Apolo and Michael, 2024). Legal commentators predict that we will

experience a tsunami of deepfake evidence and they wonder how the existing rules can handle this new threat (Amadi, 2024). The federal Rules of Evidence are adapting: the Advisory Committee proposed new Rule 901(c) governing AI-generated evidence in late 2024.

The draft Rule 901(c) involves two main presumptions – Firstly, the mere statement of this is a deepfake will not be sufficient. The objecting party will need to present a preliminary indication of falsity (forensic evidence of clues, expert analysis, metadata revealing anomalies) before the court can further investigate. Secondly, if a credible basis for suspicion of the deepfake has been established, the burden will shift to the proponent of the evidence in order to prove the authenticity using a higher standard than usual. The proponent of evidence will need to prove that it is more likely than not that the recording from which the audio or video originated was authentic, as opposed to using the ordinary prima facie showing under Rule 901(a).

In practical terms, courts will not take the proponent's or addendum's label on its face. As an example, the proposed rule expressly reads a party challenging [an item] on the ground that it was fabricated using AI must adduce evidence of fabrication to warrant inquiry. If this threshold showing is made, the item will be admitted only if the proponent demonstrates more likely than not is authentic. For example, if the defendant in the fraud suit says the recorded confession was a deepfake, the defendant would then need to produce some proof (for example, expert testimony) that indicated the confession was a deepfake (Pfefferkorn, 2019). If the deepfake attack was appropriate, the plaintiff would then have the burden to prove that the recording was from the being accused; the courts might also want to look to demonstrated chains of custody timeline, voice biometrics, known reliable transcripts, etc. This burden-shifting construct is still entirely theoretical (the proposed rule is not final), but ultimately illustrates how courts might allocate and assess the risk of disputed evidence.

Absent this proposed rule, courts will consider cases based on the existing evidence law. Courts can analogize cases ruling based on hidden alterations or digital edits. For example, lower courts have held that incredible edits in a video may justify exclusion of the video based on lack of trustworthiness. In some instances, in deepfake contexts, some judges have ruled that fake or edited video losses the confidence of the fact-finder; suggesting some level of judicial skepticism. Furthermore, some deepfake decisions (outside of the financial realm) have held deepfake images as inadmissible based on claimed privacy or defamation doctrines. There are still limited cases where images have been ruled admissible, some outside of fraud, under defamation- privacy doctrine ruling.

In addition to the courts, regulators and enforcers will have evidentiary choices to rely upon for AI deepfake evidence. A consumer complaint to a regulator cannot be terrible if the consumer contends an advertisement contained a deepfake of my voice, never mind that they will not do anything on their own, but in a litigation it will be an evidentiary determinate, where regulators may have to act. In an enforcement action context, regulators have begun treating deepfake complaints similar to fraud claims by requiring eye and/or ear witness statements, or expert analysis, or forensic reports before they accept claims of wrongdoing as factual. For AI technology to disclaim a benefit of the doubt, the standards of materiality and

preponderance require a heavier burden of proof than mere surmise or skepticism and will import presumptive weight at the same time, unless the suspect evidence that crossed the materiality standard does not yield corroborating evidence (Seng and Mason, 2021). The law is clearly establishing a heavier burden of proof for AI technology related fraud, related to a growing circle of evidence presented to courts to support such claims to balance the threat of false skepticism (the liar's dividend) against and the possibility of just letting fraud carry on unchecked.

Human Rights and Ethical Dimensions

Beyond evidentiary and procedural considerations, deepfake financial fraud raises broad issues of human rights related to personal dignity. The ability to clone an individual's face or voice, without their consent, implicates rights to privacy, autonomy, and identity. The Universal Declaration of Human Rights states that individuals have a right to security of person and property. The International Covenant on Civil and Political Rights states that individuals have the right to be free from unlawful attacks on their honor and reputation. Deepfake fraud violates both rights at once, monetizing the human likeness for nefarious purposes.

Ethically, the harm is not just about the financial loss. Victims are often left with feelings of empiricism, loss of power, and anxiety over their digital likeness that could be circulated for an indeterminate amount of time without their knowledge or control. These responses fit aptly within philosophical debates about the degradation of authenticity that define digital life. If a CEO's voice is cloned to order money to be transferred fraudulently, or a consumer's likeness is synced to a scam advertisement, the person becomes an interchangeable data point to be twisted to others' economic advantage.

In addition, the burden to police authenticity is heavily placed on the courts. Courts then may created burdens of disbelief, where good evidence is rendered no good based on suspicion and doubt. This "liar's dividend," does not only impact litigation but spreads into social and democratic discourse as well. These ethical frameworks call for greater protections, to reflect the need for human dignity through the consideration of protecting economic means.

Public vs. Private Regulation: Which Agency Leads and How AML/KYC Adapt

The intersection of technology and finance creates jurisdictional challenges in the new space of deepfake financial fraud. In the U.S., the most relevant regulators are in two categories: financial regulators (the Treasury, banking agencies, and securities regulators) and communications/consumer regulators (the FTC, telecom regulators, and broadcasting commissions). Each regulatory category has a different toolkit, and often a different interest.

(a) Financial Regulators and AML/KYC Responses

Financial institutions, especially banks and payment institutions, being the frontline targets and frontline defenders against deepfake schemes. FinCEN (the Treasury Department's financial intelligence unit) has taken an early leadership role in addressing deepfake fraud. An industry alert released by FinCEN in November 2024 warned bad actors are opening financial accounts using deepfakes created with generative artificial intelligence to evade identity verification (Fincen, 2024). The warning alerts banks to various risk indicators, such as a customer's photo ID appearing unreasonably young for its stated date of birth or a client employing a dubious webcam plug-in during live video verification. Therefore, banks are recommended to check metadata, use reverse-image searches of photos on ID verification, and apply enhanced due diligence to questionable transactions (Romanishyn *et al*, 2025).

FinCEN emphasized that these schemes are connected to critical Anti-Money Laundering ('AML') issues (cybercrime and fraud) and also remind organizations of their statutory responsibilities for reporting Suspicious Activity Reports ('SARs') (Fincen, 2024). Practically speaking, banks in the United States are expected to implement tighter Know Your Customer ('KYC') protocols, including AI tools to identify synthetic images of photo identification, and also be a strong current to require liveness in video verification, and flag accounts that exhibit repeated behavior opening multiple accounts with AI-generated documents, similar in nature and therefore suspicious. Additionally, federal banking regulators support this push. Fed Governor Michelle Bowman referred to deepfakes, this new cheating fraud scheme, as an accelerated form of identity fraud (Barr, 2025).

Also, in April 2025, Fed Governor John Barr stated that banks are frontline defenders to work toward increasing identity verification to include AI-powered technological advances such as facial recognition, voice analysis and behavioral biometrics. The Fed has formally indicated that existing multi-factor authentication also require the adoption of deepfake algorithms in order to detect minute differences from voice or video. Banks are additionally warned to evaluate recipient accounts: Barr advocated that banks should utilize advanced analytic tools to assess recipient suspicious payment patterns (e.g. unusually large payments over a short period of time, multiple similar accounts) before approving large payments (Barr, 2025).

Other U.S. financial agencies also consider deepfake fraud scams within their missions. The Securities and Exchange Commission ('SEC') has recognized the danger of AI disinformation and how it can cause market disorder (targeted erosion of confidence). SEC officials recommend that brokers and advisors better educate themselves through media-literacy training, and ensure adequate tools to identify potential AI fraud. While SEC enforcement is focused on securities markets - when representatives of brokers or dealers circulate false AI content (e.g. a fake or staged company video, or a fake statement by a CEO) that affects the price of securities - the agency's valuable guidance indicates that these actions are all within the anti-fraud jurisdiction of the SEC (Rohman et al., 2025). Similarly, the SEC supports the efforts of the private sector toward (FS-ISAC overall taxonomy

conceptualization, Deloitte reports) and toward establishing industry standards that would require firms to utilize deepfake detection safeguards.

Regardless of the blockages above, it is apparent that U.S. financial regulators (e.g. Treasury Department, CFPB, SEC, and the Fed) are mobilizing: the proposed Preventing Deep Fake Scams Act of 2025 would create a Treasury-sponsored task force, chaired by the Secretary of the Treasury and with the Fed, CFPB, OCC, FDIC, NCUA, and FinCEN as working task force members to study and address the potential for deepfake fraud. The proposed legislation also represents a consensus that anti-fraud enforcement should be led by the banking regulators in the U.S. and not the tech regulators. It also comes from the perspective that deepfake financial fraud schemes could involve illegal financial flows (e.g. money laundering, payment fraud) rather than traditional communications (Khan *et al*, 2024).

(b) Communications and Consumer Regulators

Regarding communications, the Federal Trade Commission (FTC) recently showed interest in fraud using AI technologies. In addition to the impersonation rulemaking already discussed for AI, the FTC considers false or deceptive ads using AI as an unfair practice under Section 5 of the FTC Act. In some cases, the FTC has already taken action against crypto or investment scams (not necessarily deepfakes specifically) related to false or deceptive advertising (Scharfman, 2024). The new proposed FTC rule could push liability upstream to people selling AI technologies used for impersonation. Similarly, state attorney general offices have the ability to use consumer protection laws to go after fraudulent ads using AI.

Communications regulators like the Federal Communications Commission (FCC) have a limited role in content policing (due to First Amendment and common carriage doctrines). However, they may become involved in situations like when telecom carriers must block robocall campaigns when the robocalls include deepfake voices. In fact, in a pending League of *Women Voters v. Kramer* (D.N.H. 2025) case included robocalls using a deepfake Biden voice, which has implications for the robocall regulations (even if it is primarily a political speech case). More broadly, telecom regulators around the world are working on call authentication (STIR/SHAKEN, CRIS), which will verify ID caller and could reduce deepfake voice calls. Again, many of these efforts are still in the early stages (Young, 2025).

Outside the U.S., other countries have taken different stances. For example, the UK's Office of Communications (Ofcom) is a regulator of broadcast and video content. Their rules about harmful misinformation could tangentially cover deepfakes and misinformation, but Ofcom has yet to put specific policies out regarding deepfakes. The Financial Conduct Authority (FCA) regulator in the UK has taken a leadership role in this area, with CEO Rathi providing early warnings to banks to improve defenses against fraud as AI develops (Joshi, 2025). The FCA's consumer- protection regulation encompasses unauthorized financial advertising.

Indeed, the FCA has exercised de facto platform authority as to ads. In 2021, the FCA and Google engaged together to develop a policy that banned unapproved financial adverts. Google is now in a position where the FCA has to approve ads for UK financial products, otherwise it prohibits paid ads. As the FCA states it, it is a duty to prevent financial harm,

meaning that platforms have to pull ads from scammy finance companies in order to stay compliant with regulators. The FCA had a large impact in diminishing illegitimate ads by almost 100%. The example of the FCA shows how a financial regulator and a platform can find common ground to enforce laws regarding financial promotions. It also presents a possible model, since regulators could require compliance from platforms using supplied lists of known authorized lists, disclaimers, or deposit screens for certain types of paid ads (Mason and Clarke, 2025).

(c) Adapting AML/KYC Frameworks

A principal regulatory concern is whether, and how, knowledge-your-customer (KYC) and anti-money laundering (AML) laws might be applied, and enforced, in respect of deepfakes. Traditional KYC procedures involve reliance on government-issued ID documents and in-person authentication (Seshadri, 2025). While AI forgeries could pervert these traditional authentication methods, regulators expect institutions to change KYC processes to keep pace with these risks. This alert from FinCEN reminds banks that they cannot open accounts without first verifying that the customer presented government-issued ID and the source of funds, and further indicates that these same AI tools can create falsified documentation of both forms (Fincen, 2024). It follows, therefore, that banks could be directed to conduct enhanced due diligence (EDD) for high-risk customers, which includes the use of multi-factor circumstantial identification checks, independent third-party documentary verification services, and ongoing transaction monitoring to identify unusual activity (e.g. layering transactions, unexpected geographical locations) as noted in the Federal Reserve's guidance.

How might this imply a formal regulatory change of KYC and AML laws? Currently, there is no direct mention of either AI or deepfakes in AML laws. However, the Bank Secrecy Act (BSA) and accompanying regulations obligate that reasonable procedures be established and deployed by financial institutions to verify identity. FinCEN's recent guidance arguably provides clarity that reasonable procedures now imply an ability to detect synthetic identities. In reality, regulators will expect banks to incorporate AI tools that analyze webcam feeds, voiceprints, and face biometrics for indicators of fraud. If a fraud occurs after a bank has not changed its KYC process, regulators may cite a failure to have implemented an enhanced KYC process as a compliance violation (Rybacka, 2024). The Financial Services industry has already begun remedying this: commercial vendors market the sale of deepfake detection modules for deployment as part of the larger KYC service, and some industries (FS-ISAC) have published control frameworks for deepfake risk.

Another consideration is cryptocurrency and fintech; cross-border payments (especially crypto payments) bring their own challenges to AML. The Senate bill and FinCEN also specifically call out crypto scam issues. While crypto exchanges are often decentralized, they also remain obligated to employ KYC/AML procedures. It is safe to assume that regulators will extend the scrutiny of KYC procedures to onboard customers using video as part of the process. The Fed indicates this is even possible by stating that a non-depository institution might conduct analytics of payors to verify and scrutinize payees. And lastly, flowing out of

travel rules for crypto transfers (FATF/FinCEN), it is possible that regulators may eventually imply a need for KYC protocols that identify deepfake impersonators for transactions flagged as potentially risky. However, this remains speculation.

Financial regulators associations are leading the charge: there are clear statutory obligations of their mandates (fraud, AML, market integrity) covering the unauthorized use of deepfakes. Communications regulators (the FTC/Ofcom), are filling in the gaps in enforcing the deceptive nature of deepfakes in advertising, impersonation, etc. but usually issue specific guidance rather than a law. Cooperation between both areas of enforcement may be the best approach...for example, FinCEN may communicate with the FTC or SEC to inform them of scams that are methods of ways the scam may be denoted in the methodology; or the FTC disseminating or coordinating enforcement with the DOJ or SEC; or state regulators (like the FCA, etc) sharing intelligence with international regulating bodies as it relates to deepfake fraud. There are missing links, but the broad strokes conclude that financial regulators in our discussion will tackle the issues related to deepfakes as strictly an AML/fraud issue.

Technological Countermeasures

Although laws are not changing quickly enough, technological defenses can provide immediate resilience against deepfake financial fraud. Digital watermarking is one potential solution that could be developed, and there are ongoing efforts like the Coalition for Content Provenance and Authenticity (C2PA) to implement systems that embed cryptographic signatures in media files to allow institutions to verify whether an image or video has an authentic origin. Watermarks can be manipulated by fraudsters, but if watermarking became commonplace (deep wallet), fraud would become costlier and courts would be provided with greater assurance of authenticating the image or video.

Financial institutions are assessing AI-powered fraud detection scanning for minute anomalies in voice tone, facial micro-expressions, or metadata. For example, behavioral biometrics—including typing rhythm, mouse movement, or interaction dynamics—can disclose if a widget user is an authentic account holder or synthetic imitation. Some banks are scaling these systems to authenticate during KYC onboarding and in transaction monitoring workflows.

Distributed ledger technologies like blockchain could facilitate immutable identity registries. If digital identity had an observable relation to truly immutable and verifiable ledgers, fraudsters would lose ability to create synthetic identities. Adoption of these systems is uneven, and these technological countermeasures highlight the value of security through a multi-layer approach which includes law, regulation, and innovation.

Private-Law Remedies and State Comparisons

In addition to public enforcement, victims of deepfake financial fraud may have private-law claims—like tort or statutory causes of action. The most straightforward route is to claim

fraudulent misrepresentation. Whenever a deepfake tricks someone into giving away money, the wrongdoer—and potentially others who facilitated the deceit—can be sued for deceit or conspiracy. States across the country have statutes regarding fraud and statutes, either common law or statutory, that would extend to deepfakes just like other types of misrepresentation (Shirish and Komal, 2023). For example, a CFO whose voice is cloned and instructs some person to make a false payment commits an actionable lie with the intent to defraud in same way that he would have for any misrepresentation. This is precisely what happened in the Arup v. Cooley matter that Michael Barr cited from his time at the Federal Reserve, where reportedly Arup was pursuing civil fraud claims after being impersonated using deepfake technology.

Victims have also experimented with torts based on privacy, such as invasion of privacy or false light. As Ali et al (2025) noted, victims of deepfake media sometimes pursue false light, invasion of privacy, defamation, [and] intentional infliction of emotional distress. These tort causes of action require demonstrating to the court that the deepfake placed the victim in false or offensive light. It may not be as applicable where the primary motivation is financial fraud, but it could become relevant in some situations, for instance, if a CEO's likeness were used in an advertisement online to imply he had endorsed the advertisement. In some states, like California, right of publicity statutes also exist to protect against the unauthorized commercial use of their likeness or image.

Another possible line of argument is negligence. A victim might argue some platform, bank, an intermediary failed to perform a duty of care to protect identity or assets from being exploited. As scholar has noted, the lack of meaningful safeguards against misleadingly deepfakes might give rise to any number of negligence claims (Malik *et al*, 2024). For example, a financial institution that allowed customer service reps solely relied on voice verification and did not have any secondary identification or verification measures could be subjected to a negligence claim for failing to protect customers from harm. States also have passed certain consumer protection statutes, (like the California Consumer Privacy Act) that create a private right of action when the cause of the harm is a failure to use reasonable security procedures. Consumers (plaintiffs) will likely try to stretch the private right of action to cover deepfake losses as result of the statute. Although courts are generally hesitant to recognize freestanding duty to protect against non- bodily(physical) harm if a special relationship has not been established.

Property-based theories retain their salience as well. Deepfake fraud may involve outright theft or misappropriation of funds, providing a basis for traditional claims of conversion or unjust enrichment (Spivak, 2018). Depending on the circumstances, a victim could seek recovery of the value of property that was (1) misappropriated or (2) unjustly enriched by the wrongdoer. While property concepts may not apply deepfake videos per se, it is likely in the case of misappropriation of likeness or intellectual property claims, particularly when a deepfake video misuses intellectual property such as copyrighted materials, trademarks, or trade secrets, all without permission. The hybrid nature of deepfake litigation is reflected with Forrest, for example, where the plaintiff claimed misappropriation of likeness and negligent supervision against Meta and others for deepfake advertisements exploiting her likeness,

while Meta successfully argued for immunity from a Section 230 defense. The court allowed the plaintiff's claims to proceed and suggested that Meta may be unjustly enriched for having profited so disproportionately from deepfake ads, even if the immunity afforded by Section 230 enjoys some prominence in the platform liability defense.

Comparative perspectives highlight stark contrasts between U.S. and foreign approaches. American state legislatures have predominantly criminalized and deterred deepfake harms instead of expanding private civil liability against intermediaries. Washington's new law, for example, grants good-faith platforms an immunity from civil claims, while Pennsylvania law offers an exemption if disclaimers are utilized. As a result, victims of deepfake scams tend to pursue the source fraudsters—who are likely located abroad or by simply not retrievable—or pursue novel or disparate claims such as civil RICO. Some foreign jurisdictions permit regulators or judges to require technology companies to proactively filter harmful content and material, although this framework generally conflicts with American free-speech rights. In most instances in the United Kingdom, victims may pursue defamation claims or, when operational, claims under the Online Safety Act itself; but, deepfake financial fraud is primarily addressed in as a criminal offense.

Of course, in practice, private enforcement is complicated. Deepfakes are anonymous, fast, and transnational, and address direct remedies take time and are family more complicated. In most cases, victims will first approach law enforcement, regarding civil claims as a tool of second resort or adjunct relief. Nonetheless, tort and statutory claims can play important roles. That is, knowledge and acceptance of tort law can serve as a deterrent like negligence, compel intermediaries to take action to improve protections, and serve as a complementary process if a prevention entity and law enforcement is involved. As one writer indicated multiple claims may arise where cybercriminals utilize deepfakes, for example, the CCPA provides [one] private right of action to victims alongside claims of negligence and state unfair competition statutes. Ambitious plaintiffs have attempted to use civil RICO to prove patterns and practices of wire (financial) fraud across state lines or jurisdictions (Furey, 2024).

Private law does offer some forms of private civil remedies for fraud, privacy, negligence, property, and intellectual property claims that can address fraud and financial fraud related to deepfakes; however, private law often amounts to civil claims depending on or overlapping with some idea of criminal prosecution, and a sufficient deterrence or form of protection as current and future plausibility; these private claims still do not substitute for effective and comprehensive regulation.

Conclusions

Deepfake financial fraud poses a significant challenge at the intersection of law, technology, and public policy. The rapid evolution of generative AI affords bad actors the ability to impersonate individuals, create identities, and exploit financial systems at a scale that existing legal and regulatory frameworks are struggling to contend with. Although courts

have begun to explore liability of platforms, evidentiary standards, and the scope of Section 230 immunity, progress has been inconsistent and remains in its infancy. Government regulatory action – both in the financial regulatory space and consumer protection – show trajectory, but there are gaps in coordination, lack of shared jurisdictional authority. Private law remedies are valuable, but their public order limitations are limited in geographical and substantive scope, in light of the transnational, anonymous, and rapid response of fraud schemes. Collectively, this underscores the need for a robust frame or framework in which to grapple with the growing challenge of technological risk while preserving rights and economic security.

Going forward, a multi-pronged approach is necessary. Policymakers should clarify the obligations of platforms, but should refrain from stopping legitimate speech or prohibiting what is hard to define, ideologically voluminous, and controversial. Regulators will likewise need to broaden or at least adapt AML/KYC standards to the threats brought on by AI. At the same time, evidentiary rules should be adaptable to balance the risks associated with the liar's dividend, in light of undetected fraud. Also, international cooperation is important because of the multinational scope of deepfake scams and jurisdictional limitations. Ultimately, deepfake financial fraud reveals the truth about the demand for a system that balances strong regulatory oversight and accountability through private law and technological protections. Together they may deter bad actors and reinstate public trust in financial and communications systems resulting in protection of individual victims and integrity of markets globally.

Bibliography

ALI, M.; FERNANDO, Z. J.; HUDA, C.; MAHMUTAROM, M. 2025. Deepfakes and Victimology: Exploring the Impact of Digital Manipulation on Victims. *Substantive Justice International Journal of Law*, **8**:1–12.

AMADI, P. 2024. Federal Rules Of Evidence In The United States Of America And The Challenges Of Authentication In The Age Of Deepfake Technology. *UNILAG Law Review*, **7**:308–336.

APOLO, Y.; MICHAEL, K. 2024. Beyond A Reasonable Doubt? Audiovisual Evidence, AI Manipulation, Deepfakes, and the Law. *IEEE Transactions on Technology and Society*, **5**:156–168.

ASADI, F. 2023. Digital Platforms and Intellectual Property Infringement: Exploring Legal Liability for User-Generated Content in the Context of Digital Media. *Legal Studies in Digital Age*, **2**:39–50.

AVSEC, A. J.; MICHAELS, M. M. 2025. *Forged Faces, Real Liability: Deepfake Laws Take Effect In Washington State And Pennsylvania*. Disponível em: <https://www.mondaq.com/unitedstates/new-technology/1668436/forged-faces-real-liability-deepfake-laws-take-effect-in-washington-state-and-pennsylvania>. Acesso em: 15 ago. 2025.

BARR, M. S. 2025. *Deepfakes and the AI Arms Race in Bank Cybersecurity.* Disponível em: <https://www.federalreserve.gov/newsevents/speech/barr20250417a.htm>. Acesso em: 15 ago. 2025.

BRACKEN, M. 2025. *Financial deepfake scams targeted in bipartisan Senate bill.* Disponível em: <https://cyberscoop.com/financial-deepfake-scams-targeted-in-bipartisan-senate-bill/>. Acesso em: 15 ago. 2025.

BUTTON, M.; HOCK, B.; SUH, J. B.; KOH, C. S. 2025. Policing cross-border fraud 'Above and below the surface': mapping actions and developing a more effective global response. *Crime, Law and Social Change*, **83**:1-27,

CITRON, D. K.; WITTES, B. 2018. The Problem Isn't Just Backpage: Revising Section 230 Immunity. *Georgetown Law Technology Review*, **2**:425-453.

DE RANCOURT-RAYMOND, A.; SMAILI, N. 2023. The unethical use of deepfakes. *Journal of Financial Crime*, **30**:1066–1077.

EDWARDS, C. J.; BAKER, B. L. 2025. Intermediary Liability and Future Challenges for Section 230. *The Business lawyer*, **80**:287–295.

FINCEN. 2024. *FinCEN Alert on Fraud Schemes Involving Deepfake Media Targeting Financial Institutions.* Disponível em: <https://www.fincen.gov/sites/default/files/shared/FinCEN-Alert-DeepFakes-Alert508FINAL.pdf>. Acesso em: 15 ago. 2025.

FS-ISAC. 2024. Deepfake Technology Poses New Threats to Financial Institutions; *FS-ISAC Provides Guidance.* Disponível em: <https://www.fsisac.com/newsroom/deepfake-technology-poses-new-threats-to-financial-institutions-fsisac-provides-guidance>.

FUREY, M. 2024. Sitting on a Throne of Lies: Using RICO and Wire Fraud to Hold Politicians Accountable and Demonetize Campaigns that Intend to Defraud. *Villanova Law Review*, **69**:649.

GOLLIN, T. 2025. A Deep Fake Dilemma: The Battle Over Keeping It Real. *DePaul Journal of Art, Technology & Intellectual Property Law,* **35**:1-32.

GUPTA, D. 2024. Generative AI and Deep fake s: Ethical Implications and Detection Techniques. *Journal of Science, Technology and Engineering Research*, **1**:45-56.

JOSHI, V. C. 2025. *Changing Dimensions of Financial Services and Banking Regulation.* Singapore: Springer Nature Singapore,

KAUSHIK, P.; GARG, V.; PRIYA, A.; KANT, S. 2024. Financial Fraud and Manipulation. *In* GUPTA, G.; BOHARA, S.; KOVID, R. K.; PANDLA, K. (Eds.). *Deepfakes and Their Impact on Business*. IGI Global. 173–196.

KHAN, R.; TAQI, M.; AFZAL, A. 2024. Deepfakes in Finance. *In* LAKHERA, G.; TANEJA, S.; OZEN, E.; KUKRETI, M.; KUMAR, P. (Eds.). *Navigating the World of Deepfake Technology*. IGI Global, 91–120.

ŁABUZ, M. 2023.Regulating deep fakes in the Artificial Intelligence Act. *Applied Cybersecurity & Internet Governance*, **2:**1–42.

MALIK, S.; SURBHI, A.; ROY, D. 2024. Blurring boundaries between truth and illusion: Analysis of human rights and regulatory concerns arising from abuse of deepfake technology. *In International Conference Series on Ict, Entertainment Technologies, and Intelligent Information Management In Education And Industry 2024*, AIP Publishing.

MASON, R.; CLARKE, J. 2025. Social media as a compliance risk for financial services: Exploring emerging risks and finding solutions to mitigate harm. *Journal of Financial Compliance*, **8**:229-245.

PFEFFERKORN, R. 2019. " Deepfakes" in the Courtroom. *Boston University Public Interest Law Journal*, **29**:217-245.

ROHMAN, F. Y.; RAFI, K.; GANESHAN, S.; DEEKSHITH, K.; VEENA, K.; VINODH, K.. 2025. The Influence of Artificial Intelligence On Information Integrity: A Media Literacy Approach For Young People. *International Journal of Environmental Sciences*, **11**:1022–1034.

ROMANISHYN, A.; MALYTSKA, O.; GONCHARUK, V. 2025. AI-driven disinformation: policy recommendations for democratic resilience. *Frontiers in Artificial Intelligence,* **8**:1–21.

RYBACKA, J. 2024. The Significance of the KYC Procedure from the Perspective of Banking Institutions in Poland and the Perception of this Policy from the Clients' Perspective–Based on an Empirical Study. *Finanse i Prawo Finansowe*, **3**:117–137.

SCHARFMAN, J. 2024. *The Cryptocurrency and Digital Asset Fraud Casebook*, Volume II. Cham: Springer Nature Switzerland.

SENG, D.; MASON, S. 2021. Artificial intelligence and evidence. *Singapore Academy of Law Journal*, **33**:241–279.

SESHADRI, C. 2025. How DPRK IT Workers Exploit Identity Management Vulnerabilities. *The RUSI Journal*, **170**:74–84.

SHIRISH, A.; KOMAL, S. 2023. A socio-legal inquiry on deepfakes. *California Western International Law Journa*l, **54**:517.

SPIVAK, R. 2018. " Deepfakes": The Newest Way to Commit One of the Oldest Crimes. *The Georgetown Law Technology Review*, **3**:321-339.

UNITED STATES DISTRICT COURT N.D. CALIFORNIA. 2024. *Forrest v. Meta*, Judgment June 17, 2024.

YOUNG, F. 2025. *A Deepfake Evidentiary Rule* (Just in Case). Disponível em: <https://library.law.uic.edu/news-stories/a-deepfake-evidentiary-rule-just-in-case/>. Acesso em: 6 ago. 2025.

Submetido: 25/06/2025
Aceito: 04/12/2025