TYPES AND CASE STUDIES OF AI-ENABLED FINANCIAL FRAUDS

Ishita Ahluwalia, LL.M. (Corporate, Banking & Insurance Law), Amity Law School, Amity University Campus, Noida

Dr. Ekta Gupta, Associate Professor, Amity University, Noida.

ABSTRACT

The speedy incorporation of Artificial Intelligence (AI) in the finance industry has resulted in increased efficiency, better decision making, and customer satisfaction. Nevertheless, such advancement has equally made it easier for fraudsters to commit sophisticated financial crimes. AI-based financial fraud entails acts where fraudsters take advantage of AI technology in deceiving people or systems financially.

This paper focuses on some of the common types of AI-based financial frauds that include deepfakes, artificial intelligence-based phishing, algorithmic trading manipulation, identity theft using biometric spoofing, and automatic money laundering. To give an example, fraudsters can use AI-based deepfake technology to deceive victims into believing that they are interacting with executives or other people. The same technology enables fraudsters to steal personal information from individuals, which may eventually be used for financial exploitation purposes.

Further, some of the cases that show the practical implications of these frauds have been examined. In one case, an energy company in Europe suffered huge losses because of a deepfake voice that mimicked the voice of its chief executive officer. In another case, there have been instances where bots using artificial intelligence have been seen manipulating stock markets through high-frequency trading tactics. The use of AI has led to frauds even in the area of synthetic identities.

In summary, although AI offers great potential in revolutionizing the finance sector, AI also poses great potential danger because it makes financial scams more complex and widespread. A comprehensive solution that entails changes in both legal and technological fronts is necessary.

Keywords: Artificial Intelligence (AI); Financial Scams; AI-powered Fraud; Deepfake Scam; Phishing Schemes; Identity Theft Scam; Algorithmic Trading Fraud; Money Laundering; Cybersecurity; Fraud Detection System; Machine Learning; Financial Regulations; Online Banking Security.

INTRODUCTION

The deployment of AI technology within the financial sector has led to a paradigm shift within the dynamics, pace, and evolution of financial crimes. The industry has experienced a change from static, rule-driven forms of fraud, such as simple signature counterfeiting or phishing through templated emails, to dynamic, generative, and self-evolving fraud schemes. This chapter proposes a detailed taxonomy of financial crimes that utilize artificial intelligence, organized not only according to the attack vector but based on the specific architecture used within machine learning algorithms. These attacks can be generally categorized under three categories, namely, Synthetic Media and Impersonation (GANs), Algorithmic Market Manipulation (reinforcement learning + HFT), and Cognitive Engineering (Language Models + Social Engineering).

Architectures for Deepfake Impersonation and Synthetic Media

Among the many forms of AI fraud, those using deep learning algorithms to create highly realistic synthetic media represent a particularly challenging threat vector. They threaten to break down the trust layer which has formed the bedrock of financial transactions, where audio-visual data was assumed to be immutable proof of identity.

Generative Adversarial Networks (GANs) and Visual Synthesis

The underlying principle that enables visual deepfakes through technology is the Generative Adversarial Network (GAN). GAN architecture involves a duel between two neural networks: a Generator, denoted by G, and a Discriminator, denoted by D. The Generator produces images generated from random noise and aims to replicate the distribution statistics of a dataset, for instance, the bank customer faces dataset. On the other hand, the Discriminator tries to differentiate the fake images generated by the Generator from the real ones, assigning probabilities.[1]

- This adversarial training is continued until the generator can produce data that cannot be differentiated from reality by the discriminator. This technique has been further developed using models like StyleGAN and Deep Fake, allowing the fine-grained manipulation of facial attributes (latent space manipulation), thus creating “synthetic identities” or face swapping.

[1] Goodfellow, Ian, et al. "Generative Adversarial Nets", *Advances in Neural Information Processing Systems* 2672 (2014); See also, Snippet for GANs in fraud detection.

- **Face Swapping and Injection Attack:** The fraudsters use autoencoder algorithms to map the facial landmarks of the victim onto those of the perpetrator in real time. This step is essential in evading V-KYC processes. The injection attack refers to the direct input of the generated stream into the banking application, without the involvement of camera hardware.[2]
- **Diffusion Models:** Apart from GAN, there have been other developments using diffusion models (for example, Stable Diffusion). The method uses text prompts to gradually denoise the initial random distribution in generating images. It is employed in fraud for creating static documents such as passports, driving licenses, or utilities receipts that mimic the appearance of real physical documents to evade OCR-based verification.[3]

Audio Synthesis and Voice Cloning

The equivalent of deepfakes for the ear rather than the eye is AI voice cloning technology, which has fundamentally changed "vishing." Earlier techniques used for voice generation involved days of training data. With the latest few-shot models, all that is needed is a mere three-second sample.[4]

- **Process:** The system uses the spectral envelope, pitch, tempo, and accent characteristics of the speaker. Mapping these characteristics to a text-to-speech system enables the creation of completely new sentences in the target's voice.
- **Emotional Delay:** Modern AI systems are also able to apply particular emotions to their artificial voices, like anxiety, fear, and even anger. This is especially deadly in cases of "Grandparent Scam" or "CEO Fraud", as the scammer will try to mimic someone the victim knows or trusts in a distressed manner.

[2] Cashfree Payments, "Deepfakes: The New Weak Link in Video KYC" (2025), available at: https://www.cashfree.com/blog/deepfake-fraud-video-kyc-secure-id/.
[3] CameraForensics, "The Dark Reality of Stable Diffusion" (2024), available at: https://www.cameraforensics.com/blog/2024/02/08/the-dark-reality-of-stable-diffusion/ ; See also, Snippet.
[4] McAfee, "Artificial Imposters: Cybercriminals Turn to AI Voice Cloning for a New Breed of Scam" (2024), available at: https://www.mcafee.com/blogs/privacy-identity-protection/artificial-imposters-cybercriminals-turn-to-ai-voice-cloning-for-a-new-breed-of-scam/.

Algorithmic Market Manipulation and Automated Predation

As mentioned earlier, while deepfakes are used to subvert identities, the process of algorithmic market manipulation involves targeting the price discovery layer of the financial market environment, utilizing HFT algorithms and AI agents to conduct predatory operations.[5]

AI-Driven Spoofing and Layering

Spoofing is the process of making fictitious orders to generate an appearance of market depth. Even though this is illegal, spoofing is hard to identify when done using AI.

- **Reinforcement Learning (RL):** RL agents differ from static programs that make decisions using "if-then" commands because RL agents learn how to control prices through experience obtained in simulation settings. The RL algorithm tries to maximize profits by making and canceling orders in milliseconds in a way that causes a reaction among other algorithms or human traders.
- **Layering:** This refers to the placement of numerous limit orders at various price levels within the same direction of trade. For example, placing a series of bids creates a false appearance of strength in the market. The market price starts moving up because of the "false" strength, after which AI places a real order on the opposite side of the market before cancelling the bid limit orders.[6]

Front-Running and Pump-and-Dump

- **Digital Front Running:** AI machines analyze data feed including news feeds and social media sentiments faster than markets can assimilate them. AI can predict a big order coming from institutions and “front run” the trade by buying up the stock milliseconds before that order comes in.[7]
- **Pump-and-Dump with Generative AI:** Scammers employ LLMs to create thousands of fraudulent articles and social media posts that create hype around a particular microcap stock. Botnets spread this information to create "social proof" for retail investors, who eventually fall into this trap. Thereafter, the algorithm initiates a sell-

[5] Yadav, "AI algorithmic market manipulation techniques spoofing layering mechanism", *Journal of International Law* (2024).
[6] Lin, "Detecting Spoofing and Layering in High Frequency Trading", *Journal of Financial Crime* (2025).
[7] VerityAI, "Algorithmic Trading Oversight & AI Governance" (2025).

off. The application of artificial intelligence enables the generation of new, logical, and convincing stories on an unprecedented scale.[8]

Cognitive Engineering: AI-Enhanced Phishing

This is the industrialization of social engineering. LLMs reduce the barrier to creating deceptive content of higher quality.

Spear Phishing and Whaling

Phishing messages can usually be detected through their poor grammar and lack of specificity. AI-powered phishing involves using a generative model to scrape an individual's online presence (LinkedIn, Twitter, company bios). The AI then creates a highly customized message (Spear Phishing) that cites particular colleagues, projects, or events.[9]

- **Whaling:** When targeting senior executives (Whaling), the AI can mimic the writing style of the CEO or a business partner, analysing past public emails or letters to replicate syntax and tone. This dramatically increases the success rate of Business Email Compromise (BEC) attacks.

Digital Arrest and Coercion

This new form of vector attack combines impersonation techniques with psychological warfare. Artificial Intelligence technologies are deployed to produce fraudulent legal documentation (warrant or court order) as well as a virtual video environment (police station). The targeted individual is made to believe he or she is "digitally arrested," thereby necessitating the transfer of funds in order to prove his/her innocence.[10]

[8] Akshansh, "Jane Street and Expiry Day Trap: Unpacking SEBI's Crackdown", *Oxford Business Law Blog* (2025).
[9] PwC, "Impact of AI on Fraud and Scams" (2024).
[10] VisionIAS, "Supreme Court Empowers CBI to Investigate Digital Arrest Scams", *Current Affairs* (2025).

CASE STUDY 1: DEEPFAKE FRAUDS IN LOAN APPROVALS – GOVERNANCE AND DEFENCE

Digitalisation of the loan verification process – more precisely, V-KYC – was supposed to be efficient and effective. Instead, it introduced a new form of cyber security risk to banks through deepfake attacks. In this part, the defensive strategies of HDFC Bank are compared with Arup's tragic experience to demonstrate two sides of AI in banking.

"Vigil Aunty" Campaign by HDFC Bank: Deepfakes for Defence

During the 2023-2024 period, HDFC Bank, the biggest private Indian bank, faced increasing threats caused by AI-powered scams. Instead of waiting until something happened, the institution initiated a proactive campaign known as "Vigil Aunty - End of Scam Sale". It is important to note that this is an interesting case of using deepfake technology in order to fight against deepfake scams.

"Lulumelon" Hoax

Together with creative agencies, HDFC Bank created a fictional fashion company called "Lulumelon". For promoting it, the bank used its own version of deepfakes with the cloned faces and voices of Bollywood actor Nora Fatehi.[11]

• **Deepfake Technology:** In this scenario, the deepfake avatar of Nora Fatehi offered "too-good-to-be-true" discounts on Instagram. The AI voice cloning technique was in perfect synchronization with the movement of the lips, making it seem like the celebrity was endorsing the brand.

• **Psychological Manipulation:** The fraudster was targeting users' Optimism Bias and FOMO (Fear of Missing Out). The target users believed the visual proof of the endorsement by the celebrity and opened links to access discounts.

• **Learning Outcome:** Instead of being directed to a payment gateway, the target audience was taken to the educational page. The user could watch a video detailing how he/she had become a victim of deepfakes and receive information about detecting deepfakes (e.g., unnatural blinking, audio distortion).

[11] Gan.ai, "Case Study: HDFC Bank's End of Scam Sale" (2024).

Governance Consideration: This is an excellent example of how fraud risk management should evolve in banks. Previously, organizations relied on static advisories for their clients. However, the "Vigil Aunty" campaign indicates that combating AI fraud involves educating customers using a similar technique. If banks can employ deepfake techniques within weeks in their campaigns, criminals can do the same.

The Arup Incident (2024): A Failure of Visual Trust

As much as HDFC used the deepfake technology for educational purposes, the one that occurred in 2024 where an employee of Arup, an international engineering firm, fell victim to scammers after sending HK$200 million (equivalent to $25 million) is a clear example of how vulnerable firms can be.[12]

Incident Reconstruction

The fraud started with a phishing email allegedly sent by the company's UK-based Chief Financial Officer (CFO), asking for a confidential transaction. After the employee raised concerns, the fraudsters moved on to using a video conferencing platform.

• **Deepfake Video Conferencing:** The employee was surprised to see the CFO and several other key figures join the conference. The quality of the audio and video was excellent; the facial expressions seemed natural, and the voices sounded like those of the executives.

• **No Humans Allowed:** Incredibly, all other participants in the conference were deepfakes, and the fraudsters utilized artificial intelligence to manipulate digital models of the executives, possibly based on video clips available from corporate webinars and YouTube channels.

• **The Action:** Confident in the presence of his bosses, the employee did not follow the required procedure and completed 15 transfers of funds to five separate bank accounts. The scam came to light after only a week, when the employee contacted the headquarters.

Analysis of Governance Lapses

Arup's example highlights how "Visual Trust" is not a proper security control.

[12] Gross Shuman, "Case Study: $25 Million Deepfake Scam Sends a Wake-up Call to Corporate Cybersecurity" (2024).

1. The digital footprint risk is the likelihood that the fraudsters may have collected audiovisual content of the executives from publicly accessible sources. This leads to the emergence of a novel risk factor where being visible in public is a potential threat.

2. **Out-of-band authentication was absent:** The main governance flaw is the lack of a required out-of-band authentication step (a phone call, a chat, etc.) for high-value transactions started on video.

3. **Liveness check failed:** Had the video conferencing platform had a live facial recognition feature based on blood flow pixel analysis or surface texture detection, the deepfake content could be easily detected.

Vulnerabilities in Indian V-KYC Frameworks

Deepfake threat scenario in the Indian environment will be limited to loan on-boarding using Video KYC technology.

- **Injection Attacks:** In this kind of attack, fraudsters tamper with the data flow for V-KYC technology. Rather than streaming their own live feed, they will use an already pre-rendered deepfake or face-swap video to open new accounts using stolen identities.
- **Creation of Mule Accounts:** Such accounts, referred to as mule accounts, serve as the infrastructure for money laundering. Several reports from Hyderabad and Bangalore have mentioned instances in which the fraudsters used voice clone impersonating family members for such attacks.[13]
- **Regulatory Gap:** The RBI requires banks to conduct "liveness checks." However, many non-banking financial institutions have software that cannot detect 2D deepfakes. The failure of 3D liveness detection is a systemic issue.

[13] Poonawalla Fincorp, "Deepfake Frauds & AI Scams Explained" (2025).

CASE STUDY 2: ARTIFICIAL INTELLIGENCE IN INVESTMENT FRAUD THROUGH FOREIGN TECHNOLOGY PLATFORMS

This section will study how artificial intelligence is exploited in investment fraud by using foreign technology platforms to circumvent domestic regulatory restrictions. Two vectors will be studied separately in this context, including the algorithmic manipulation by firms such as Jane Street and the extortion by Chinese loan apps.

Algorithmic Manipulation: The Jane Street Case (2024-2025)

The Jane Street case, reported by the Securities and Exchange Board of India after investigation in mid-2025, is a perfect example of a complex manipulation of market microstructure through algorithmic trading.

- **The Manipulation:** Algorithmic Trading SEBI investigated an allegation of manipulation of the Bank Nifty index on options expiry days against Jane Street, a global prop trading firm.
- **Manipulative Strategy:** The firm was accused of building up large long positions on cash days (buying the index component securities) while taking large short positions in derivatives (hedging the position or betting on the fall of the index).[14]
- **The "Dump":** In the final minutes of trading on expiry day, the algorithms would execute a synchronised sell-off of the cash positions. This flood of sell orders created artificial downward pressure on the index.
- **The Profit:** The induced drop in the index value dramatically increased the profitability of the firm's "short" options positions. SEBI alleged that Jane Street made an illicit profit of approximately ₹735 crore (approx. US$88 million) in a single trading session (January 17, 2024).

The AI Component and Market Microstructure

This method made use of HFT programs which could perform thousands of transactions per second.

[14] DataIntellect, "SEBI vs Jane Street: Index Manipulation Scandal and a RegTech Wake-up Call" (2025).

- **Predictive Modelling:** It seems like the algorithms must have used machine learning to estimate the "market impact" of their selling activities, calculating precisely what volume they required to move the index by the specified number of points to make the options profitable.
- **Grey Box Manipulation**: Such an incident is a classic example of why it is difficult for regulators to oversee AI-based stock trading. While the company insisted on the fact that it was just a hedging activity, the pattern and time correlation with the expiration period analysed by SEBI implied that it was intentional, making it a PFUTP violation.
- **Consequences:** SEBI put out an interim order of confiscating ₹4,843 crore worth of profits, implying that the regulators are beginning to look at the intent programmed into their trading algorithms.

Mass-Market Extortion: Chinese Loan Apps and AI Morphing

On the retail front, the case of the "Chinese Loan App" provides a vivid example of the Digital Debt Trap that may be formed using AI and foreign technology solutions.

- **The Modus Operandi:** Illegal apps usually offered via side-loaded APK files or temporarily available in app stores would provide micro-loans instantly to poor borrowers.
- **Data Collection:** Once installed on the device, these applications request an excessive number of permissions allowing full access to contact and photos folders.
- **AI Manipulation (Extortion Tool):** If the debtor misses the deadline for loan repayment, the recovery agency will apply Generative AI to "morph" the victim's face to naked or pornographic photos.
- **Mass Automated Extortion:** Generated pictures will be sent back to the debtor accompanied by threats to share them in the user's contact list. Using AI, agents will be able to generate thousands of personalized humiliating images each day.[15]

The Financial Laundering Trail

The ecosystem relies on a complex web of financial intermediaries.

[15] BBC News, "Chinese Loan App Scam Exposed by Undercover Investigation" (2024)..

- **Crypto-Laundering:** Investigations by the Enforcement Directorate (ED) revealed that funds extorted from Indian victims were funnelled through shell companies (e.g., Shinebay, Truekindle). These funds were then converted into cryptocurrencies (USDT, Bitcoin) and transferred to wallets in China and Hong Kong.[16]
- **Payment Aggregator Vulnerabilities:** The initial transactions often flowed through Indian payment gateways (Paytm, Razorpay). These entities were subject to ED raids for failing to detect suspicious transaction activity in these merchant accounts.[17]
- **The Fall of Paytm Payments Bank:** A significant fallout was the RBI's 2024 crackdown on Paytm Payments Bank. The regulator cited "persistent non-compliance" regarding KYC and the existence of lakhs of accounts without proper identification. These lapses were structurally linked to the ease with which loan app syndicates could operate mule accounts within the banking system.[18]

CASE STUDY 3: CROSS-BORDER FRAUDS UNDER GLOBALISATION

Globalization has led to a borderless financial world, which is exploited by fraudsters who use artificial intelligence to defraud victims from various jurisdictions. This subchapter discusses the "Digital Arrest" and the development of Nigerian scams.

The "Digital Arrest" Phenomenon

The "Digital Arrest" is a psychometric fraud that emerged in India in 2024-2025 and caused losses of several hundred crores. Involving impersonation of governmental bodies, this scam places victims under "virtual" arrest.[19]

The Anatomy of the Scam

- **The Bait:** They get a phone call from a VoIP number (which seems to be a +91 number) saying their parcel has been intercepted at customs because it contained either drugs or forged passports or that their Aadhaar is associated with money laundering.

[16] News18, "ED Exposes China-Linked Financial Fraud Network" (2025)..
[17] NDTV, "Chinese Loan Apps Case: Paytm, Other Payment Gateways Searched by ED" (2022)..
[18] Times of India, "Why RBI May Have Banned Paytm Payments Bank" (2024)..
[19] Al Jazeera, "What are Digital Arrests? The Newest Deepfake Tool Used by Cybercriminals" (2024)..

- **The AI-enabled Setting:** The scammer escalates the conversation to a video call using Skype/WhatsApp. Using AI technology, they create an artificial backdrop that looks like that of a police station or CBI office. For advanced scams, the scammers overlay the video with deepfakes to impersonate an IPS officer or even a judge.
- **Panoptical Psychological Warfare:** The victim is asked to leave their camera switched on around the clock for "surveillance." The isolation from any outside contact keeps the victim from seeking any help from relatives or legal advice. The fabricated AI setting convinces them that they are dealing with the government authorities.
- **Fake Documents:** Using Generative AI, the scammer can create fake FIR, arrest warrant, and asset freeze order documents that look entirely legitimate.
- **Cross-Border Coordination**

  The victims may be Indian, but the scammers may be based out of "cyber-slavery" dens in Southeast Asia (Cambodia, Myanmar, Laos), or they could even have a base in Nigeria. But the money is funneled to Indian bank accounts first before being transferred overseas using cryptocurrencies.

- **The Evolution of the Nigerian Email Scheme: From 419 to AI Vishing**
  Nigeria, the country synonymous with email scams, or "419," is evolving into the world's center of AI voice phishing, known as "Vishing".[20]

The "Grandparent Scam" 2.0

- **The "Grandparent Scam,"** which includes the caller claiming to be the grandchild and saying he or she needs help, has been around for years.
- **AI Voice Cloning:** Before this technique, a scammer would say the call had a bad connection and could not be heard. Now, they employ technology to clone a voice. With just a few seconds of audio from the victim's Instagram account, TikTok video, etc., a scammer can create an exact copy of the voice.[21]

[20] OCONUS Investigations, "Nigerian Cybercrime: From 419 to AI Voice Cloning" (2025)..
[21] American Bar Association, "AI Cloned Voice Scam" (2025)..

- **Emotional Manipulation:** The AI can modulate the voice to sound panicked, crying, or injured. This emotional realism bypasses the victim's critical thinking. A senior citizen hearing their grandchild sobbing is biologically hardwired to assist, making this an incredibly effective vector.
- **Scale:** Reports show a 700% increase in deepfake attacks within the financial sector, where artificial intelligence has made it easier to commit the scam without knowledge of the native tongue. Instead, the AI becomes the flawless interpreter and vocalizer.

IMPACT ANALYSIS: ECONOMIC LOSSES AND EROSION OF TRUST

Fraud conducted using AI technology is much more than just a simple annoyance; it represents a major danger that results in significant economic loss, and the trust foundation is broken.

Economic Losses

As can be seen from the financial statistics, the issue of increased risk is growing.

- **India:** According to reports, India suffers losses worth roughly Rs. 22,845 crore due to cyber frauds in 2024, which is a substantial increase compared to the preceding year. Overall loss due to cyberfraud between 2020-2025 amounts to more than ₹52,000 crore.[22]
- **Global:** The trend is mirrored globally. In the US, Generative AI is projected to increase fraud losses to $40 billion by 2027, with a 32% CAGR.[23] In the cryptocurrency sector, scam revenue reached $17 billion in 2025, with AI-enabled scams being 4.5 times more profitable than traditional scams due to their higher conversion rates and scalability.[24]

Erosion of Consumer Trust

However, beyond the numbers on the balance sheets, AI fraud is unsettling consumer confidence.

- **Trust Deficit:** According to FICO's survey, 66% of consumers in India now believe banks should compensate scammed customers, showing a change in attitude: the banks,

[22] News9, "India's Cyber Fraud Losses Top ₹52,000 Crore Between 2020-2025" (2025).
[23] Deloitte, "Deepfake Banking Fraud Risk on the Rise" (2024).
[24] Chainalysis, "Crypto Scams 2026 Report" (2025).

not the customers, are responsible for security.[25] In addition to this, 57% of individuals choose a bank on the basis of improved fraud prevention mechanisms.

- **Friction During Adoption:** Fear of "Digital Arrests" and UPI-related fraud has created friction during the adoption process. Though the number of transactions via UPI are quite high, the extent to which elderly and conservative segments are using it is threatened by a large amount of stories related to loss of savings due to video calls.
- **The "Liar's Dividend":** One such paradoxical effect of deepfakes can be termed as "Liar's Dividend." With increased use of deepfakes, individuals or corporates may refute any genuine proof against them by arguing that it was generated through deepfake technology. Meanwhile, videos sent by banks (for identity verification) face skepticism amongst customers.

COMPARATIVE INSIGHTS: REGULATORY RESPONSES IN THE US VS. INDIA

This varies because the response is based on a variety of laws as well as philosophical differences between jurisdictions with respect to economic and legal systems.

The US Response: SEC and "AI Washing"

While the SEC in the US has taken a position on disclosures, actively going against any form of "AI washing," which involves misrepresentation regarding a company's abilities in terms of artificial intelligence in order to acquire funding.

- **Legal Action Taken:** In 2024, Delphia (USA) Inc. and Global Predictions Inc. were charged by the SEC for claiming to utilize "predictive AI" and "collective data" to manage funds when, in fact, they did nothing of the sort. The two companies together paid an overall fine of $400,000.[26]
- **Philosophy of Regulation:** The SEC believes that false claims about AI are an infraction under the Investment Advisers Act (Material Misrepresentation). This is

[25] FICO, "2024 Scams Impact Survey: India" (2025).
[26] SEC, "SEC Charges Delphia and Global Predictions for AI Washing" (2024).

intended to save the market from getting caught up in its own "hype cycle." If a company claims any benefits of artificial intelligence, they must prove it.

- **Conflicts of Interest:** There have also been SEC proposals for eliminating conflict of interest problems related to the use of Predictive Data Analytics (PDA). This is because there may be an inherent bias in the use of algorithms towards making money for the company rather than helping the client.[27]

The Indian Approach: RBI and SEBI

Approach in India is comparatively more interventionist, concentrating on the infrastructure for payments and transactions.

- **Digital Hygiene at RBI:** The RBI prescribes Master Directions which lay down mandatory technology requirements. The "Master Direction on Digital Payment Security Controls" of the RBI (2021) lays down the requirement for real-time fraud detection, behavioural analysis, and stringent V-KYC standards.[28] This is a prescriptive model: the regulator dictates the security architecture.
- **SEBI on AI/ML:** The Securities and Exchange Board of India (SEBI) has released consultation papers regarding "Responsible Usage of AI/ML," mandating intermediaries to disclose their AI-based applications. In contrast to the US, SEBI is greatly worried about stability in the market and has shown a willingness to directly interfere in the practice of trading, as illustrated in the case of Jane Street.[29]
- **The "MuleHunter" Initiative:** To combat the specific threat of mule accounts used in AI frauds, banks and the RBI are collaborating on AI-driven detection systems (like "MuleHunter.AI") to flag accounts based on transaction velocity and profile mismatches.

[27] KKC, "The Rise of AI in Financial Advice and the SEC's Proposed Safeguards" (2025).
[28] KKC, "The Rise of AI in Financial Advice and the SEC's Proposed Safeguards" (2025).
[29] SEBI, "Consultation Paper on Guidelines for Responsible Usage of AI/ML in Indian Securities Markets" (2025).

Table: Comparative Regulatory Frameworks: US vs. India

Feature	United States (SEC/FTC)	India (RBI/SEBI)
Core Philosophy	**Disclosure & Market Integrity.** Focus on an accurate representation of AI to investors.	**Systemic Security & Consumer Protection.** Focus on securing the payment rails.
Key Enforcement Concept	**"AI Washing"** (Punishing fake AI claims).	**"Digital Hygiene"** (Mandating MFA, V-KYC standards).
Algorithmic Regulation	Focus on **Conflict of Interest** (Algo vs. Client) and ensuring unbiased advice.	Focus on **Market Manipulation** (Spoofing, Layering) and preventing volatility.
Deepfake Response	**FTC Voice Cloning Challenge** (Incentivising technical solutions).[30]	**I4C/MHA Advisories** (Public awareness campaigns and blocking).
Approach to Innovation	"Innovation with Accountability" (Rules on *how* AI is used/marketed).	"Controlled Innovation" (Sandboxes, strict reporting requirements).

Source([31])

[30] FTC, "FTC Voice Cloning Challenge" (2024)

[31]https://www.ijpsjournal.com/article/Comparative+Study+Between+Regulatory+Quality+++System+Of+India+And+USA (Last visit 04/04/2026

BIBLIOGRAPHY

PRIMARY SOURCES:

A. ACTS

1. Indian Acts

- *The Banking Regulation Act, 1949.*
- *The Bharatiya Nyaya Sanhita, 2023.*
- *The Bharatiya Sakshya Adhiniyam, 2023.*
- *The Digital Personal Data Protection Act, 2023.*
- *The Foreign Exchange Management Act, 1999.*
- *The Information Technology Act, 2000.*
- *The Reserve Bank of India Act, 1934.*
- *The Securitisation and Reconstruction of Financial Assets and Enforcement of Security Interest Act, 2002 (SARFAESI Act).*

2. Foreign Acts

- *Corporate Transparency Act* (United States).
- *Dodd-Frank Wall Street Reform and Consumer Protection Act, 2010* (United States).
- *Economic Growth, Regulatory Relief, and Consumer Protection Act, 2018* (United States).
- *EU AI Act* (European Union).
- *Gramm-Leach-Bliley Act, 1999 (GLBA)* (United States).

B. RULES, REGULATIONS, AND NOTIFICATIONS

1. Reserve Bank of India (RBI) Guidelines

- *Master Direction on Digital Payment Security Controls, 2021.*
- *Master Direction on Information Technology Governance, Risk, Controls and Assurance Practices, 2023.*

- *Master Direction on Know Your Customer (KYC) Direction, 2016.*
- *Master Direction on Outsourcing of IT Services, 2023.*
- *Guidelines on Digital Lending, 2022.*
- *Circular on Limiting Liability of Customers in Unauthorised Electronic Banking Transactions, 2017.*

2. Other Regulatory & Government Notifications

- *Department for Promotion of Industry and Internal Trade (DPIIT), Press Note 3 (2020 Series).*
- *Executive Order 14110 on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* (United States).
- *Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021.*
- *New York Department of Financial Services (NYDFS) Cybersecurity Regulation (23 NYCRR 500).*
- *SEBI Consultation Papers on "Responsible Usage of AI/ML".*

C. CASES LAWS

1. Indian Case Law

- *Internet and Mobile Association of India v. Reserve Bank of India*, (2020) SCC OnLine SC 275.
- *Justice K.S. Puttaswamy (Retd.) v. Union of India*, (2017) 10 SCC 1.
- *Satender Kumar Antil v. Central Bureau of Investigation*, Recent Order (Supreme Court).
- *Shayara Bano v. Union of India*, (2017) 9 SCC 1.
- *State Bank of India v. Pallabh Bhowmick*, (Guwahati High Court).
- *Suresh Chandra Singh Negi v. Bank of Baroda*, (Allahabad High Court).

2. International and Regulatory Cases

- *In Re: Delphia (USA) Inc. and Global Predictions Inc.* (SEC Administrative Proceeding, 2024).
- *In Re: Jane Street* (SEBI Investigation/Interim Order).
- *R v. Zhimin Qian* (United Kingdom).

D. REPORTS AND INTERNATIONAL INSTRUMENTS

1. Official Reports (India)

- *Indian Cyber Crime Coordination Centre (I4C), Annual Statistics.*
- *Ministry of Home Affairs (MHA), Recommendations of the Committee on Digital Arrests.*
- *NITI Aayog, National Strategy for Artificial Intelligence.*
- *Rangarajan Committee Reports on Computerisation in Banks (1980s).*
- *Reserve Bank of India, Report on Trend and Progress of Banking in India (2023-24).*
- *Reserve Bank of India, Annual Report (2023-24).*
- *Reserve Bank of India, Financial Stability Reports (FSR).*
- *Report of the Committee on FREE-AI (Fairness, Accountability, Trust, and Ethics).*

2. International Standards and Reports

- *Financial Action Task Force (FATF), Digital Transformation Strategy.*
- *Financial Stability Board (FSB), Reports on Artificial Intelligence and Machine Learning in Financial Services (2024).*
- *International Organisation of Securities Commissions (IOSCO), Recommendations on AI (2024).*
- *NIST AI Risk Management Framework (AI RMF).*
- *Organisation for Economic Co-operation and Development (OECD), Principles on AI.*
- *United Nations Commission on International Trade Law (UNCITRAL), Model Law on*

Automated Contracting (2024).

SECONDARY SOURCES:

- Eastnets, *Global Fraud Report 2024*.
- Hutchinson, T., & Duncan, N., *Defining and Describing What We Do: Doctrinal Legal Research*.
- Rodrik, Dani, *The Globalisation Paradox*.
- TransUnion, *2025 Global Fraud Report*.
- Zweigert, K., & Kötz, H., *Introduction to Comparative Law*.
- IDRBT, *Research on Adversarial Machine Learning*.