CONTOURS OF CYBERCRIME

PERSISTENT AND EMERGING RISK OF ONLINE FINANCIAL FRAUDS AND DEEPFAKES IN INDIA

data LEADS

A DataLEADS Study on Online Financial Frauds and Deepfakes in India

ABOUT DataLEADS

DataLEADS is a globally recognised, award-winning digital media and technology company at the forefront of shaping conversations on information integrity, digital safety, and the future of data and AI ecosystems. Headquartered in India but global in its outlook and reach, DataLEADS has spent the last decade working to create strong information environments that empower individuals, institutions, and communities across Asia. Operating at the intersection of digital transformation, public policy, data and cutting-edge technology, DataLEADS firmly believes that access to accurate, authentic information is fundamental to human agency and societal progress. Over the past decade, DataLEADS has designed and delivered some of India's largest and most impactful initiatives on digital safety and security—reaching thousands of organisations and millions of citizens. Our Trust and Safety services span content moderation, red teaming for AI testing, deepfake detection, and end-to-end product and policy consulting. These services have been instrumental in enabling our partners—governments, platforms, and civil society—to identify risks, shape safer digital products, and protect their ecosystems from harmful content. Through high-level research, convenings, and strategic partnerships, we bring together policymakers, platforms, educators, and community leaders to co-create scalable, real-world solutions that safeguard people and institutions in a rapidly evolving digital world.

data LEADS

TABLE OF CONTENTS

EXECUTIVE SUMMARY

SYED NAZAKAT
Founder & CEO, DataLEADS

In modern finance and banking, it is not just cash or capital that keeps the financial systems healthy and alive – it is information and trust. The information, accurate and timely, helps individuals and institutions alike make sense of markets, value assets, and direct capital with intent. The trust – less visible, but more fragile – is the quiet belief that underpins the system is fair, reliable, and ultimately safe. It is the interplay between these two forces that allow economies to function, and markets to grow.

Over the past decade, only a few countries have tested this relationship as dramatically as India. The country has pulled off one of the most remarkable digital financial transformations in the world. The rise of the Unified Payments Interface (UPI) has made digital payments ubiquitous, extending financial access deep into rural areas where bank accounts were once scarce. Today, millions of previously unbanked citizens hold accounts, make payments, and engage with a digital economy they were long excluded from.

The numbers tell a staggering story.

According to data from the Government of India, UPI processed over 18 billion transactions worth more than ₹24.03 lakh crore ($279 billion USD) in June 2025 alone – a figure unthinkable a few years ago. Through the Jan Dhan Yojana, more than 550 million bank accounts have been opened in India – many by rural and semi-urban women. Today, over 80 percent of Indian adults have a bank account and 491 million use digital payments. It is a triumph not only of policy, but of scale, infrastructure, and sheer ambition.

Yet beneath this success story, there is a growing wave of threats and risks.

As India's digital ecosystem grows, so do its challenges. Across encrypted channels, cloned websites, and deepfake advisories, a sprawling network of online scams has taken root. According to the official data submitted to the Parliament, cybercriminals siphoned off ₹22,845.73 crore (USD ~2.66 billion) from the Indian market and people in 2024 alone – a staggering 206 percent increase over the previous year. Despite blocking over 9.42 lakh (0.942 million) SIM cards linked to fraudulent activity and 2.6 lakh (0.26 million) IMEIs used in scams, cybercrime continues to persist – often through coordinated and cross-border attacks.

Why? Because the perpetrators are often global while the enforcement remains stubbornly local. The efforts to combat cybercrime face a fundamental mismatch. The World Economic Forum estimates that a mere 0.05 percent of cybercriminals are ever prosecuted worldwide – an indictment of both capacity and coordination.

As digital threats grow more sophisticated and borderless, the real challenge is structural: how to construct a defence that is as agile and interconnected as the threat itself? Can smarter use of data and early-warning systems disrupt scams before they metastasise? And in an age of deepfakes and AI-driven deception, how do we keep pace with the next wave of digital threats?

In response, we're launching the Global Unit for Analysis of Risk and Digital Threats – GUARD. Born out of the need to make sense of rising digital deception, GUARD is a social listening framework and early warning platform designed to track emerging patterns of online scams, frauds and deepfakes across digital ecosystems. This report is a part of GUARD's efforts to examine persistent and emerging threats. Based on GUARD's research of 100 social media fraud cases over 90 days and 600 cases reported in the past year, the report reveals intricate workflows designed to evade detection and increase confusion.

Stemming from open-source intelligence, powered by AI verification, and grounded in public policy engagement, GUARD combines advanced technology with different social listening data. Its strength lies in providing institutions with foresight – helping them anticipate risks and build resilience against the hidden dangers of the digital financial world.

Special thanks are due to Rifat, whose tireless work in research laid the foundation for this report. Venkata Subrahmanian provided vital data and oversight. Md Ayan Haque automated the data extraction and built a repository of reported cases for analysis, while Sonia Bhaskar offered sharp editorial review, ensuring accuracy, clarity, and actionable insight.

Our Trust and Safety Team formed the operational backbone of this effort. Waseem Ahmad brought critical insights into the public policy dimensions of the report. Tej Kumar Daram developed the comprehensive research log sheet that structured much of the data capture, while Varadarajan Ananthakrishnan meticulously designed its data categories and reviewed key findings. We also extend our gratitude to all the experts whose insights shaped and strengthened this report.

Born out of a need to make sense of rising digital deception, GUARD is a social listening framework and early warning platform designed to track and detect emerging patterns of online scams, frauds and deepfakes across digital ecosystems.

The findings point not only to rising fraud, but to increasing complexity – and call for a shift from reactive responses to preventive strategies. The report outlines practical solutions, urging all stakeholders to set guardrails and help build a safer, more resilient financial ecosystem.

PREFACE

The increasing use of Unified Payments Interface (UPI), internet banking, emergence of crypto platforms, and fintech applications have transformed people's investment choices, how they save, invest and spend their money. Processing 85% of the country's digital payments, UPI recorded more than 18 billion transactions in June 2025[1]. With its full interoperability, UPI has truly democratised payments in India, bringing financial services within reach of every Indian through the mobile device. Digital payments grew from ₹162 crore (roughly $19 million[2]) in 2012-13 to ₹18,120.82 crore (approximately $2.1 billion) till January 2025. India accounts for nearly 50% digital payments globally[3].

But alongside this ease and advancement, a challenging trend of online financial frauds is seeing a spike. In 2024, according to the National Crime Reporting Platform (NCRP) managed by Indian Cyber Crime Coordination Centre (I4C), more than 1.9 million complaints of cybercrime were recorded – a dramatic increase from 1.56 million in 2023 and a tenfold spike from 2019 levels[4]. Financial frauds remain the leading cause of such crimes.

Citizens as a whole lost a record ₹22,842 crore ($2.6 billion) to cybercrime in 2024 alone — almost three times the loss of ₹7,465 crore ($868 million)[5] just a year ago in 2023, and almost ten times more than the reported ₹2,306 crore ($268 million) loss in 2022. Cumulatively from 2020 to 2024, cybercrime activities have cost individuals and businesses across the country more than ₹33,165 crore ($3.85 billion)[6].

The fallout of high mobile penetration, which has ensured that 55.3% of the population has access to the internet[7,8], is that India today has one of the largest number of people connected to social media platforms (491 million in January 2025[9]) in the world. The fact that social media platforms have emerged as a major source of information with massive user bases also means that there is a plethora of deceptive investment schemes, trading tips, the promise of unrealistic gains, trading courses and misleading advertisements. This growing threat is not just costing Indians money, but also denting their trust, sense of security and confidence in the digital financial system.

While India's financial sector is continuously upgrading itself and investing in digital technologies and security systems, the scammers on the online platforms are also constantly adapting and deploying a mix of online and offline outreach mechanisms and circuitous multi-platform techniques to dodge authorities and target unsuspecting online users. No one is immune, be it the literate, the tech savvy or the digitally challenged. From traditional Ponzi schemes to AI-driven fraud, online platforms have become the route to target both the needy and the greedy. I4C data suggests that Indian citizens are likely to lose over ₹1.2 lakh crores ($13.95 billion) in 2025[10].

The fact that social media platforms have emerged as a major source of information with massive user bases also means that there is a plethora of deceptive investment schemes, trading tips, the promise of unrealistic gains, trading courses and misleading advertisements.

This report **'Contours of Cybercrime: Persistent and Emerging Risk of Online Financial Frauds and Deepfakes in India'** seeks to highlight the complex web of online scams by examining and analysing the user journey and various touch points across platforms that are used to target and convert a user into a victim. The report takes a ground up approach to trace the workflow from the set up to the realisation and final evasion stage of online frauds. More importantly, it highlights the framework to combat the risks and mitigation strategies at scale to build a more secure online financial environment.

1. Press Information Bureau. (2025, July). India's UPI Revolution Over 18 billion Transactions Every Month, A Global Leader in Fast Payments. Retrieved July 24, 2025, from https://www.pib.gov.in/PressNoteDetails.aspx?NoteId=154912
2. Unless otherwise stated, a conversion rate of 1 USD = 86 INR has been used throughout this report
3. Press Information Bureau. (2025, March). Digital Payment Transactions Surge With Over 18,000 Crore Transactions in 2024-25. From https://www.pib.gov.in/PressReleaseIframePage.aspx?PRID=2110405
4. India Today. (2025, June). India's Rs 22,800 cr cybercrime hydra: How to defang it. From https://www.indiatoday.in/india-today-insight/story/indias-rs-22800-cr-cybercrime-hydra-how-to-defang-it-2747621-2025-06-28
5. Sansad. (2025, July). Unstarred Question No. 344 regarding cyber fraud. From https://sansad.in/getFile/loksabhaquestions/annex/185/AU344_PAfnK3.pdf?source=pqals
6. The Hindu. (2024, May). AI becoming formidable tool for cybercriminals: Report. From https://www.thehindu.com/news/cities/bangalore/ai-becoming-formidable-tool-for-cybercriminals-report/article69736569.ece
7. Press Information Bureau. (2025, June). Highlights of Telecom Subscription Data as on 31st May 2025. From https://www.pib.gov.in/PressReleasePage.aspx?PRID=2140179
8. DataReportal. (2025, February). Digital 2025: India. From https://datareportal.com/reports/digital-2025-india
9. Meltwater. (2025, March). Social Media Statistics for India [Updated 2025]. From https://www.meltwater.com/en/blog/social-media-statistics-india
10. The Hindu. (2024, October). Cyber fraud losses could amount to 0.7% of GDP, MHA study projects. From https://www.thehindu.com/sci-tech/technology/cyber-fraud-losses-could-amount-to-07-of-gdp-mha-study-projects/article68788093.ece

We have to be careful with new technology... These videos look very real and, therefore, we need to be very careful before believing the authenticity of a video or an image. India is emphasising a global framework for AI.

Prime Minister **Narendra Modi** at Smart India Hackathon December 20, 2023

METHODOLOGY

Our research methodology for this report involved tracking financial scams and misleading information through multiple stages of its lifecycle using primary investigation (through comprehensive social media monitoring across platforms and risk evaluation) and analysing secondary sources (through studying reported scams).

Over the 90-day period between April 1, 2025 and July 1, 2025, the GUARD team conducted a comprehensive analysis of the posts, ads containing content about investment tips, financial advice and schemes across different online platforms such as Instagram, YouTube, Facebook, X and closed messaging apps such as WhatsApp and Telegram. Further, the team also analysed the data from prominent fact-checking websites and cybercrime coverage in media outlets published from November 2024 to June 2025.

The primary focus was to find cases of misleading financial claims, online frauds, use of deepfakes and AI generated content, in addition to potential accounts that run scams and the modus operandi in terms of the user journey and call to actions. In addition to extensive social media monitoring to identify emerging use cases and methods deployed by the bad actors to target vulnerable users online, the team also did a thorough tracking of reported scam incidents across the country.

It also considered other relevant elements of individual posts including engagement metrics (views, shares, virality, comments), mention of currencies or money figures, potential geographic origins, and misuse of known brand and institution names. The team then assessed the reach and impact of the content to identify persistent and emerging patterns of the online financial scams on social media sites.

Finally, the GUARD team also looked into how AI techniques like deepfakes and voice cloning of politicians, business leaders or officials of regulatory authorities and organisations are being used to promote certain dodgy investment schemes or trading platforms. This multi-stage approach provides detailed insights into how scams spread and evolve, and can be tracked and identified, rather than being limited to 'how it happened' after cases of fraud have been reported and money has been lost.

CHAPTER-1

FROM TRUST TO TRAP

SOCIAL ENGINEERING AND PRIVATE MESSAGING FUEL FINANCIAL FRAUD

Among most tech products or solutions, social media platforms are the most dynamic, in terms of their evolution, their offering and how these are being deployed to serve different purposes. What started out as a public square for people to meet, exchange ideas, opinions and build connections, soon grew in stature and user bases to assume the size of many countries' populations. The range of these platforms is varied and the impact wide. With this growth the influence of these platforms on individuals' lives grew – from seeking personal validation in the form of likes, comments and shares, to being the source of entertainment and news, credible or otherwise, to entrusting faith in these platforms with one's hard earned money.

What role do social media platforms play in the increasing cases of online frauds and how do people end up losing money? That was one of the objectives of the extensive social media monitoring exercise undertaken by the GUARD team. The goal was to understand the user journey and what motivates users on these platforms to end up making payments to an unknown entity behind the veil of anonymity.

In a physical world, if strangers are not handed over money then why do people not exercise the same caution online? The study monitored most prominent social media platforms daily for a period of 90 days (April 1, 2025 – July 1, 2025). The platforms included Facebook, Twitter, YouTube, Instagram and these then led to the monitoring of closed messaging groups on WhatsApp and Telegram. In this process, 100 potentially fraudulent cases were identified. What emerged was the anatomy of online financial scams. The data

What role do social media platforms play in the increasing cases of online frauds and how do people end up losing money? That was one of the objectives of the extensive social media monitoring exercise undertaken by the GUARD team.

GUARD FINDINGS: PLATFORMS FROM WHICH 100 ANALYSED SCAMS ORIGINATED

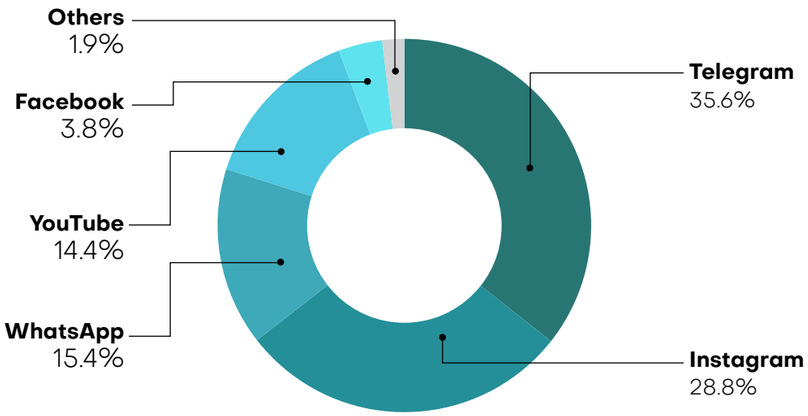


Figure 1: Platforms of origin of 100 analysed scams (Source: GUARD Research)

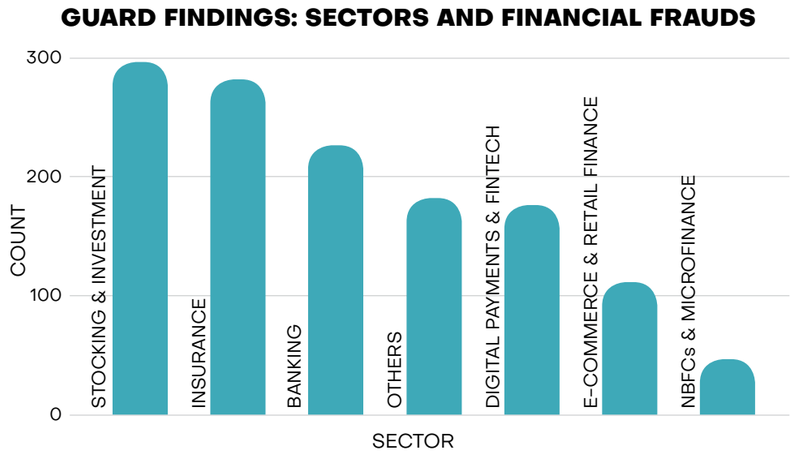


Figure: Sector-wise analysis of frauds analysed through secondary research - media publications and fact-checks (Source: GUARD Research)

In addition to primary research, GUARD also scanned 600 reported cases from multiple sources such as newspapers, online websites and government data to zero in on the role played by social media platforms in perpetuating financial scams.

revealed disciplined, yet deceptively simple funnels: from ads to pages, from pages to private channels, followed by fabricated and inflated proofs of profits shared persistently in these closed groups and then for more convincing, unverified identification proofs for payments made by others were shared.

In addition to primary research, GUARD also scanned 600 reported cases from multiple sources such as newspapers, online websites and government data to zero in on the role played by social media platforms in perpetuating financial scams.

Online financial frauds in India are targeting entire sectors, ranging from old economy banking and insurance to fintech, digital payments, and NBFCs, causing widespread disruption.

BANKING SECTOR

Banking fraud in India has surged dramatically, with the Reserve Bank of India reporting a nearly eightfold jump in value during the first half of the fiscal year 2025. Compared to the same period in 2024, the number of reported frauds in 2025 rose from 14,480 to 18,461, with the associated monetary loss escalating from ₹2,623 crore ($305 million) to ₹21,367 crore ($2.5 billion)[11]. For the full financial year, the fraud amounts nearly tripled to ₹36,014 crore ($4.19 billion)[12].

Private sector banks accounted for nearly 60% of the total fraud incidents, reporting 14,233 cases in 2024-2025. However, the highest monetary impact was borne by the public sector banks, with frauds totalling ₹25,667 crore ($2.98 billion) — a whopping 71.3% of the total fraud amount across all the banks[13].

INSURANCE SECTOR

The insurance sector, spanning life, health, vehicle, and general insurance, has become an increasingly attractive target for fraudsters. As companies digitise their operations and reach customers through online platforms, scammers are using misleading information, fake offers, and impersonation to defraud both customers and companies.

The fraudsters in many cases analysed by GUARD impersonate several well-known and credible insurance companies, including Reliance General Insurance, HDFC ERGO General Insurance, Kotak General Insurance, Chola MS General Insurance, Shriram General Insurance, Royal Sundaram General Insurance, IFFCO-TOKIO General Insurance, and Future Generali General Insurance. They produce forged policies using computers, printers, and fake QR codes that make the documents appear official. Since they often hand over the fake policy instantly — unlike genuine insurers who typically take up to 12 hours or more — many buyers fail to notice the warning signs.

The impact on victims is severe. Customers, upon discovering the fraud, often blame or doubt the real companies, thinking they were careless or complicit, even though the firms had no role in the scams.

STOCKBROKING, FINTECH AND INVESTMENT SECTOR

In GUARD's secondary research, fintech, stockbroking, and investment sectors together accounted for the highest volume of fraud cases, with nearly 500 out of the 600 documented incidents analysed. This concentration underscores how these interlinked domains - characterised by high digital penetration, rapid customer onboarding, and fragmented regulatory coverage - have become prime targets for scammers exploiting trust gaps, data vulnerabilities, and consumer urgency.

TIMELINE: FAKE SBI LIFE INSURANCE POLICY SCAM (INDORE, 2023–JAN 2025)[14]

➔ **Early 2023:** A retired teacher in Indore receives a call from "Rajiv Sharma," who claims to be an SBI Life Insurance official offering to help with her existing policies. She is asked to transfer online ₹1 lakh (roughly $1160) as processing fee for an upfront amount to claim policy benefits.

➔ **2023–2024:** Over the next two years, the scammers extracted a total of 34 online payments, amounting to ₹96 lakh (roughly $1,11,000) for taxes, policy processes, and promising future profit benefits.

➔ **January 2025:** Communications abruptly cease, and she realised she has been defrauded.

11. NDTV. (2024, December). Banking Frauds Rise In H1FY25, Amount Involved Jumps 8-Time: RBI Report. From https://www.ndtv.com/india-news/banking-frauds-rise-in-h1fy25-amount-involved-jumps-8-time-rbi-report-7336767
12. The Indian Express. (2025, May). Bank fraud amount jumps by three times to Rs 36,014 crore in FY25: RBI. From https://indianexpress.com/article/business/bank-fraud-amount-jumps-by-three-times-to-rs-36014-crore-in-fy25-rbi-10036075/
13. The Economic Times. (2025, May). Banking fraud cases fell but amount involved tripled in fy25. From https://economictimes.indiatimes.com/news/economy/finance/spike-in-loan-and-digital-frauds-rbi-data-reveals-frauds-jump-three-times-in-fy25/articleshow/121492721.cms
14. India Today. (2025, February). Indore woman loses Rs 96 lakh while paying for fake SBI life insurance policy. From https://www.indiatoday.in/technology/news/story/indore-woman-loses-rs-96-lakh-while-paying-for-fake-sbi-life-insurance-policy-2684846-2025-02-24

CHAPTER-1

From Trust to Trap: Social Engineering and Private Messaging Fuel Financial Fraud

CASE STUDY: NIDHI SAGAR INVESTMENT SCAM (DELHI, JULY 2024)[15]

Early July: Spots Facebook ad using Upstox brand. Clicks through and gets added to a WhatsApp group impersonating a popular stock trading app led by “Rajat… VIP”.

Early July: Installs app from a shared link and uploads Aadhaar + PAN to activate account.

Mid July: Nidhi makes 19 payments totaling ₹24.35 lakh (roughly $28,000) to 10 different bank accounts for shares and IPOs. App dashboard shows gains. She attempts to sell holdings. The platform blocks withdrawal and pressures her to take a loan.

Aftermath: Realises the operation was fraudulent. Faces financial loss and exposure of personal ID documents.

Takeaway: Social media to chat groups to off-platform apps is a high-risk funnel.

Red Flags: Brand impersonation; fund splitting across many accounts; request for personal data and ID proofs on unverified channels.

Investors are tricked through fake apps, websites, and social media promotions with high returns. Victims are shown fictitious gains but get delays or blockages when they attempt withdrawals. Unregulated "finfluencers" (influencers in the financial sphere) push investment, housing and crypto schemes with no disclosures, causing loss of investors’ life savings, inflicting enormous financial and emotional anguish.

OTHER SECTORS

Scammers often impersonate popular e-commerce platforms like Amazon, Flipkart, Myntra, Meesho, or Ajio to trick unsuspecting customers. They either create fake websites and apps that mimic the look and feel of the official brand or run fake promotional campaigns on social media. These fake offers often promise unbelievable discounts, cashback rewards, or exclusive “flash sales” to lure people in.

Some common online scams include:

- **Phishing Scams:** Messages (SMS/WhatsApp) falsely claim prize winnings or refunds from e-commerce sites. These direct individuals to links or fake customer care numbers, where fraudsters solicit sensitive details like UPI IDs or card numbers to steal funds.
- **Fake Product Listings:** Popular items are listed at unrealistically low prices on online marketplaces. Buyers pay in advance, but the seller vanishes, and the product is never delivered.
- **Payment Confirmation Scams:** Fraudsters send fake SMS messages or emails claiming an order couldn't be completed and asking to “verify payment”. Clicking these links can lead to theft of financial credentials or malware installation.
- **QR Code Scams:** Malicious QR codes are placed on billboards, posters, flyers, or even through email links or on e-commerce sites. Scanning these codes can lead to fake payment pages or malware downloads.

These scams thrive because they exploit the massive popularity of online shopping and digital transactions in India, where millions of payments happen daily. Many first-time or less digitally literate shoppers are unfamiliar with best practices for online safety, making them easy targets. Also, fraudsters use fast-evolving tactics like setting up fake websites, stealing logos, and leveraging social media ads to stay ahead of enforcement and detection.

The GUARD’s research exercise reveals a troubling trend: online financial scams have evolved from isolated scams to coordinated, multi-platform operations targeting every sector of the financial ecosystem.

LEARNINGS

According to data presented in the Parliament, in the financial year 2023–24, Maharashtra reported the largest amount lost to financial frauds – approximately ₹42.54 crore ($4.95 M), followed by Kerala – ₹30.42 crore ($ 3.54 M), Haryana – ₹ 29.57 crore ($3.44 M), Uttar Pradesh – ₹29.23 crore ($3.4 M), and Karnataka – ₹6.61 crore ($0.8 M). This pattern suggests the vulnerability in India’s densely populated and highly digitised economic regions[17].

The GUARD’s research exercise reveals a troubling trend: online financial scams have evolved from isolated scams to coordinated, multi-platform operations targeting every sector of the financial ecosystem. From banking and insurance to fintech and e-commerce, fraudsters are exploiting the pace at which India’s digital presence is accelerating.

CASE STUDY: FAKE PART-TIME JOB OFFER FROM AMAZON[16]

➔ **Initial Contact (via Telegram):** An engineer is approached by “Amazon recruiters” offering a “part-time product review task”.

➔ **Early Tasks & Small Payments:** He clicks on the provided links, places low-value test orders, and posts reviews. Scammers pay him modest commissions.

➔ **Advance Payment Request:** After a few weeks, he is told that larger “unlock fees” or “security deposits” are required to access higher-paying tasks.

➔ **Escalation of Payments:** He sends multiple advance payments, believing he will receive bonuses later. Total outlay reaches ₹50 lakhs ($58,000).

➔ **Discovery of Fraud:** Promised commissions and bonuses stop arriving and all contact channels go silent. He realises the operation was a sham.

15. Finance Outlook India. (2025, May). A Heightened Stock Market Frauds In India: Unraveling Three Recent Cases. From https://web.archive.org/web/20250218230538/https://www.financeoutlookindia.com/editors-column/a-heightened-stock-market-frauds-in-india-unraveling-three-recent-cases-nwid-2904.html

16. India Today. (2023, September). Engineer gets offer for part-time job in Amazon, loses Rs 52 lakh in the scam. From https://www.indiatoday.in/technology/news/story/engineer-gets-offer-for-part-time-job-in-amazon-loses-rs-52-lakh-in-the-scam-2442071-2023-09-29

17. Digital Sansad. (2024, December). Loss Due to Financial Frauds. From https://sansad.in/getFile/annex/266/AU2471_EyMojc.pdf?source=pqars

CHAPTER-2

ANATOMY OF CYBER SCAM

TECHNIQUES, TACTICS AND THE PSYCHOLOGY BEHIND ONLINE FRAUD

Dissecting the cases, collected through primary and secondary research and evaluating the risks through the framework that the GUARD team has developed, yielded what we call the 'Anatomy of The Online Financial Scam'. The framework looks at more than 25 parameters to assess the threat posed by the content appearing on different social media platforms before the user is diverted to the closed messaging groups.

The framework looks at more than 25 parameters to assess the threat posed by the content appearing on different social media platforms before the user is diverted to the closed messaging groups.

The parameters of the framework range from context of the post or ads to analysing the source that the post originates from or the ads lead to. Followed by the content analysis of the claims or the call to action, in terms of whether personal data is being sought and what level of personal and financial data is being collected, be it bank account details or Aadhar/PAN and other documents. Further the framework also takes into account if some QR codes or links are being shared for an app or other file downloads apart from the type of media content used and engagement metrics of the content, in terms of views, likes, shares, comments. The other aspect we looked at is the emotional quotient of the posts and use of prominent people to attract attention. Based on all these parameters and more, the risk evaluation of the posts were done and categorised into Severe, High, Medium, Low risk.

According to data from Ministry of Home Affairs' Indian Cybercrime Coordination Centre[18] (I4C), while the law enforcement agency has partnered with social media companies on intelligence sharing, the organisation said that in the three months from Jan-March 2024 alone, they received complaints where the big tech platforms are being misused.

Platform	January 2024	February 2024	March 2024
WhatsApp	15355	13696	14746
Telegram	8462	6567	7651
Instagram	6708	5940	7152
Facebook	6525	7190	7051
YouTube	1591	1156	1135

Table: Number of complaints received by I4C in January-March 2024 (Source: I4C)

18. Ministry of Home Affairs. (2024, December). Annual Report 2023-24. From https://www.mha.gov.in/sites/default/files/AnnualReport_27122024.pdf

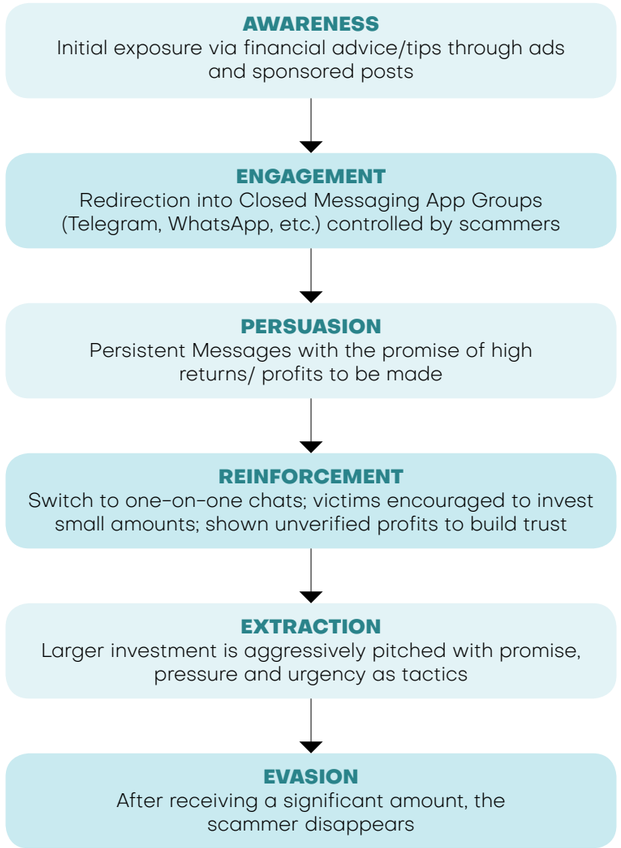
CHAPTER-2

Anatomy of Cyber Scam: Techniques, Tactics and the Psychology Behind Online Fraud

WHY TAKEDOWNS OFTEN FAIL

- ➔ Scams migrate fast to closed groups – public pages are only a funnel
- ➔ Same network runs many groups and pages at once – on multiple platforms
- ➔ Accounts rename and respawn repeatedly
- ➔ Victims are already hooked before any warning

ANATOMY OF A SCAM



Anatomy of the Online Financial Scam (Source: GUARD Research)

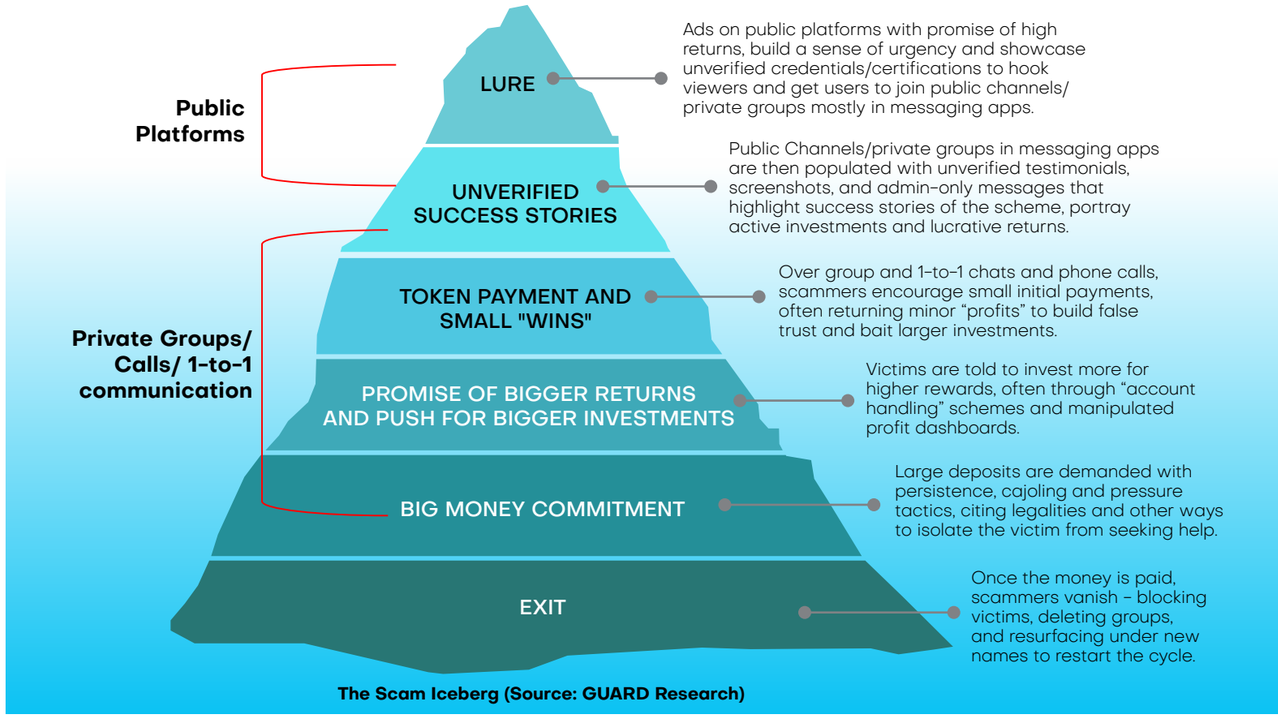
GUARD's primary research shows that taking a user from Awareness to Evasion stage takes time, effort, multiple followups, and trust building exercises. Only on the success of those, the push comes to up the stakes for larger gains.

GUARD's primary research shows that taking a user from Awareness to Evasion stage takes time, effort, multiple followups, and trust building exercises. Only on the success of those, the push comes to up the stakes for larger gains. It is not just about convincing the user about the efficacy of the scheme on offer, but constant nudging and reiterating the "benefit for you" angle and also assuring that the offer is genuine. Unverifiable ID proofs like Aadhar and PAN are shared with the user to give them a sense of security.

But the fact remains that like an iceberg, where only the tip is above water and the large chunk deceptively remains out of sight below the water, the entire user journey, from being drawn into the fraudster's net to most part of the process of social engineering takes place in the encrypted world of closed messaging groups and at a one-on-one level that remains under the radar and dodges most security and tracking mechanisms.

The trust building exercise is methodically transformed into financial loss through six distinct stages. Starting with a broad public lure and narrowing through privately providing ID proofs to elicit small, token payments to prove the validity of the claims, and escalating pressure, to isolate the targets and drive even-larger investments. The transition is swift from public platforms to closed networks, first into a group or channel and then immediately on to one-on-one personal chats. Every step after the first step, the initial lure stage, takes place in a closed environment – making it impossible for enforcers to effectively monitor these conversations and take timely preventive action.

THE ONLINE SCAM ICEBERG



The Scam Iceberg (Source: GUARD Research)

How trust is methodically transformed into financial loss through six distinct stages starting from public platforms and progressing to private ones – making scams difficult to track[19]

19. Disclaimer: Please note that this iceberg is a simplified, generalised representation intended to illustrate the typical progression of a scam. In practice, the number of stages, their sequence, the channels used, and the timing can vary substantially depending on the scammers' methods, platform policies and regional factors.

CHAPTER-2

Anatomy of Cyber Scam: Techniques, Tactics and the Psychology Behind Online Fraud

STAGES	TACTICS	RED FLAGS THAT PEOPLE IGNORE	WHAT WE SAW
I. THE LURE*	■ Paid advertisements on Instagram/Facebook ■ Enticing promises of profitable trading tips ■ Unrealistic earnings claims ■ Redirect to closed messaging groups ■ Use of trusted brand names and logos ■ AI-generated influencers or stolen expert images ■ Professional ad design	■ Ads disappear quickly ■ Too-good-to-be-true profit promises ■ Immediate redirect to private groups ■ Unauthorized use of established financial brand logos ■ Sense of urgency ("offer only valid today") ■ Limited availability claims ("only five slots left")	■ Groups with thousands to millions of followers ■ Frequent name changes and group recycling ■ Multiple pages staying live only for days ■ Common phrases: "from zero to hero," "offer only valid today" ■ Misuse of HDFC Securities and other trusted brands · Screenshot blocking in groups
II. THE STORIES:** Unverified Success Stories	■ Regular "admin" updates claiming daily profits ■ Screenshots with blurred payer details ■ Videos showing fake success stories ■ Hundreds of posts per day in some groups ■ Restricted messaging (admin control only) ■ Social proof through fake member testimonials ■ Bot-generated "thank you" messages	■ Blurred or incomplete payment screenshots ■ Members can't send messages (one-way communication) ■ Overwhelming volume of success posts ■ Generic testimonials without verifiable details ■ Screenshots taking is blocked ■ Lack of independent verification sources	■ Screenshots shared to show "active investing" ■ Fake success stories from supposed group members ■ Controlled group discussions with no member interaction ■ Bot-generated social proof messages ■ Frequent posts highlighting unrealistic returns
III. THE TASTE OF SUCCESS*:** Token Payment and Small "Wins"	■ "Free" tips that seem to work initially ■ Building rapport through one-on-one advice ■ Small initial investments to build trust ■ Flashy fake earnings dashboards ■ Fake withdrawal proofs ■ Claims of connections to big investors/celebrities ■ Gradual escalation of investment amounts	■ Initial "wins" that seem too convenient ■ Pressure to start with small amounts · Inconsistent documentation (PAN, Aadhaar, bank details) ■ Mismatched names on accounts vs. IDs ■ Fake SEBI registration certificates ■ UPI handles not matching provided names	■ Inconsistent credentials and documentation ■ Fake earnings screenshots showing rising profits ■ False withdrawal proofs to build confidence
IV. THE PROMISE: Promise of Bigger Returns and Push for Bigger Investments	■ Account handling services offering profit sharing ■ Requests for Demat account login details ■ Minimum capital requirements (₹50,000–₹2 lakh) ■ Daily follow-up messages and phone calls ■ Fake apps/websites showing rising profits ■ Pressure tactics about missing "big gains" ■ Encouragement of secrecy ("VIP access")	■ Requests for account login credentials ■ Unrealistic profit-sharing promises ■ High minimum investment requirements ■ Persistent daily contact and pressure ■ Secrecy requirements ■ Threats of missing opportunities ■ Fake customer support numbers	■ Around 50 groups investigated offering account handling ■ Daily motivational calls and messages ■ Fake apps showing artificial profit increases ■ Pressure to maintain secrecy from family/friends ■ Isolation tactics to maintain control
V. THE EXIT: Big Money Commitment	■ Withdrawal blocks and delays ■ Additional "fees" and penalty demands ■ Fake legal complications and tax threats ■ Fake customer support to handle concerns ■ Group shutdown and contact blocking ■ Complete disappearance of scammers ■ Rebranding under new names to restart	■ Inability to withdraw funds ■ Sudden fee demands for withdrawals ■ Legal threats and penalty claims ■ Customer support that doesn't resolve issues ■ Group/contact sudden unavailability ■ Loss of all invested money ■ Complete communication cutoff	■ Blocked withdrawals when victims try to exit ■ Additional fees demanded for fund release ■ Fake legal complications to justify delays ■ Complete loss of contact after money stops flowing ■ Groups deleted or abandoned ■ Victims left with financial and mental trauma

Figure*: some common phrases found in social media ads (Source: GUARD Research)

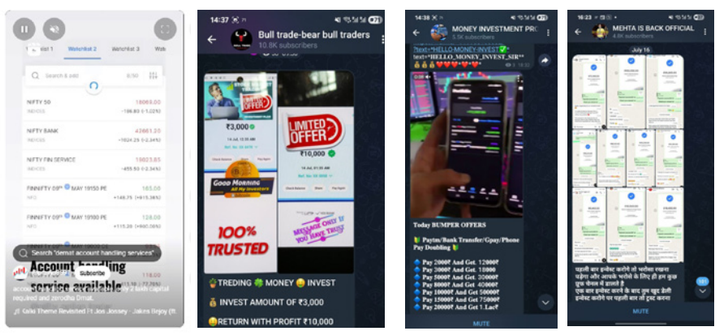


Figure: Misuse of trusted brand names to promote fraudulent schemes is rampant across social media platforms (Source: GUARD Research)**

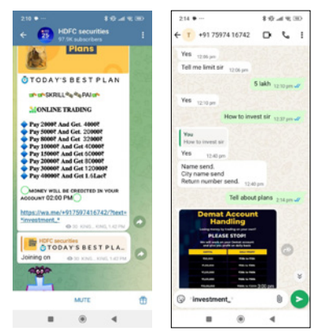


Figure*: A public Telegram group using HDFC Securities's name to lead users into private WhatsApp group, where they seek 'investments' (Source: GUARD Research)**

CHAPTER-2

Anatomy of Cyber Scam: Techniques, Tactics and the Psychology Behind Online Fraud

EXCERPTS FROM OUR 'CONVERSATIONS'

Call 1: Double your money!

(GUARD) analysts spotted a sponsored advertisement promoting quick profits through investment and trading on Instagram, which led to a Telegram channel filled with screenshots showing large profits in various investor accounts. This eventually led to WhatsApp voice calls, making the entire setup appear legitimate and professional.

GUARD Team: Sir, I want to know about the investment?

"Can you transfer ₹2,00,000 ($2200)? You will get good growth."

GUARD Team: So I will give you two lakh… and you will get me double?

"You will get ₹20,000-30,000 ($230-350) per day, as long as you will work with us."

GUARD Team: This means, if I get ₹20,000 every day, I will get ₹4,00,000?

"Yes".

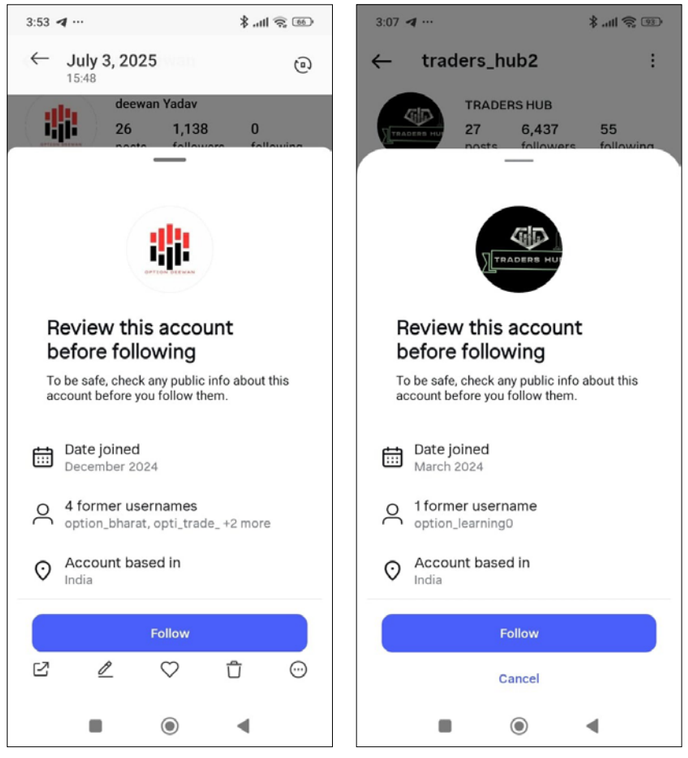
GUARD Team: Sir, is there any risk please tell me?

"There is no risk, if we do genuine work. Let's work for a few days, if you don't like you can logout."

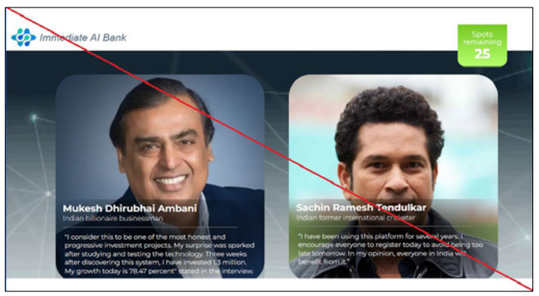
In light of GUARD's findings, Meta's new account-history safety notice, which allows people to track changes made to a username and other risk signals, is a welcome step. However, the timing and scope is a problem. The alert appears only after a user has already reached the page, while most scam activity shifts quickly into closed groups where platform safeguards are weaker.

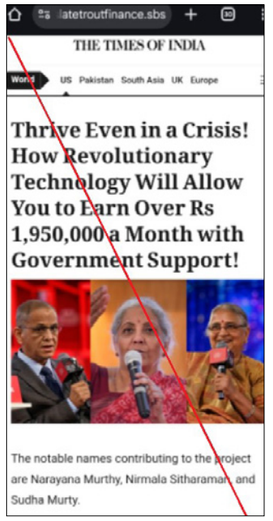
GUARD's investigations show that the same fraud operation can span multiple short-lived groups across several platforms. Shutting down a single page does little. Scammers rotate account names and links, move conversations to phone numbers, and maintain parallel channels so that takedowns simply redirect traffic. Platform response needs to anticipate cross-platform movement, detect repeated name churn tied to the same payment rails, and intervene earlier in the funnel.

The following selected dialogues drawn from GUARD researchers' exchanges with some of the group admins bring to life the scam



Meta's new feature that tells users about an account's history and potential red flags before following – highlighting the frequent name-changing tactic used by fraudulent accounts (Source: GUARD Research)





Celebrity endorsements in fake bank websites; cloned website of The Times of India with claims of government-backing (Source: GUARD Research)

iceberg outlined in the report. These verbatim snippets illustrate how scammers deploy urgency in language, fake proofs, and private-chat pressure to first build and then erode trust incrementally. Juxtaposing these conversations with the GUARD team's quantitative findings helped gain deeper insight into the psychological tactics at each stage. It also helped pinpoint moments for timely intervention.

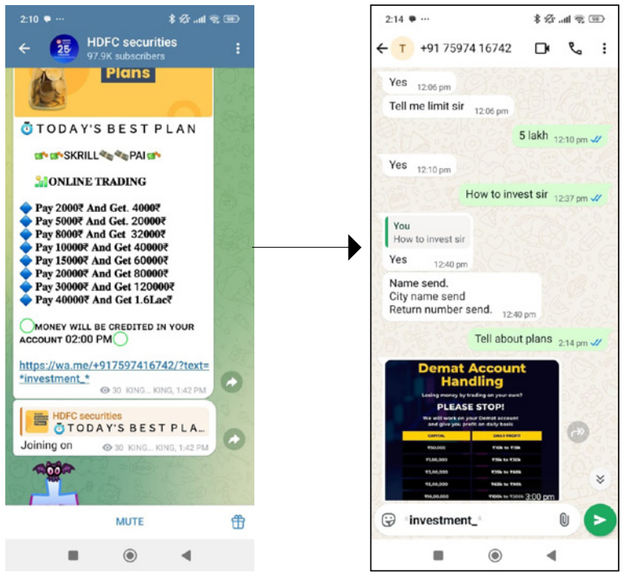


Figure: A public Telegram group using HDFC Securities's name to lead users into private WhatsApp group, where they seek 'investments' (Source: GUARD Research)

Call 2: Profits in 2-4 hours!

(The caller raises the promised return and introduces a brokerage fee, continuing the same high-pressure scam tactic. The investor expresses concern about fraud and seeks details about the investment vehicle.)

"You will get the return amount of ₹2 lakh ($2200). … but we will take some brokerage like 30% of profit."

GUARD Team: It means I will get around ₹90,000 ($1,040)?

"You will get profit in 2-4 hours."

GUARD Team: Where will it be invested, sir?

"We will do trade in US international market."

GUARD Team: Sir, I am scared, today frauds are common.

"You are right that is why I am asking you to invest ₹50,000 ($570) initially."

Call 3: It's a game of trust

This transcript captures a call in which an individual solicits a small "test" investment with promises of unrealistically high returns, a common tactic in investment scams. The caller uses unverifiable documents and exerts pressure to build trust before requesting funds.

GUARD Team: Namaste Bhaiya, I want to ask how much I will get and how much I have to invest.

"You can start from (₹) 1000–2000 ($11–$23), so that you can gain trust. You will get 10,000 from 2000 in an hour."

GUARD team: Where will this money be used?

"It will be invested in International markets"

GUARD Team: What will happen after I get ₹10,000 ($116.2) of this investment

"Then you can invest big amount and earn ₹50,000 ($581) per day"

GUARD Team: Sir, the global market is big, how will I trust you?

"I will send you a PAN card, Aadhar card and QR code. Sir, you have to trust the first time."

GUARD Team: Sir, the ID proof you sent has a photo of someone who looks like a kid?

"This is not a kid pic! The date of birth is 2003. Send money then I will talk to you."

(Hangs the call.)

KEY LEARNINGS

LEARNING	RECOMMENDATIONS
Multi-Platform Presence: Scammers operate across platforms such as Instagram, Facebook, YouTube, Telegram and WhatsApp, using them to build credibility and attract victims through fake testimonials and success stories.	Strengthen cross-platform monitoring and intelligence-sharing between platforms, regulators, and other stakeholders to detect and disrupt fraudulent networks early.
Funnel Strategy: Victims are first exposed to scams through public content or groups and then directed to closed messaging apps, which are handled by individuals posing as financial advisors or account managers.	Develop integrated takedown protocols that trace public posts to private channels and collaborate with messaging apps to block fraudulent numbers and accounts.
Emotional exploitation & persistent outreach: Scammers use urgency language ("premium plan," "account handling") and followup calls or texts to deepen trust.	Launch targeted public-awareness campaigns on common scam tactics; introduce oneclick reporting and verification and automated spam filters for unsolicited financial outreach.

The predictable patterns revealed in our investigation, from initial ads to final extortion, represent both the challenge and the opportunity. While scammers have systematised their approach, a coordinated defence mechanism is the need of the hour.

With 18 billion monthly UPI transactions creating unprecedented opportunities for both legitimate commerce and sophisticated fraud, the stakes have never been higher. The predictable patterns revealed in our investigation, from initial ads to final extortion, represent both the challenge and the opportunity. While scammers have systematised their approach, a coordinated defence mechanism is the need of the hour.

CHAPTER-3

AI AND DEEPFAKES IN ONLINE FRAUD

EMERGING THREATS IN THE DIGITAL FINANCIAL ECOSYSTEM

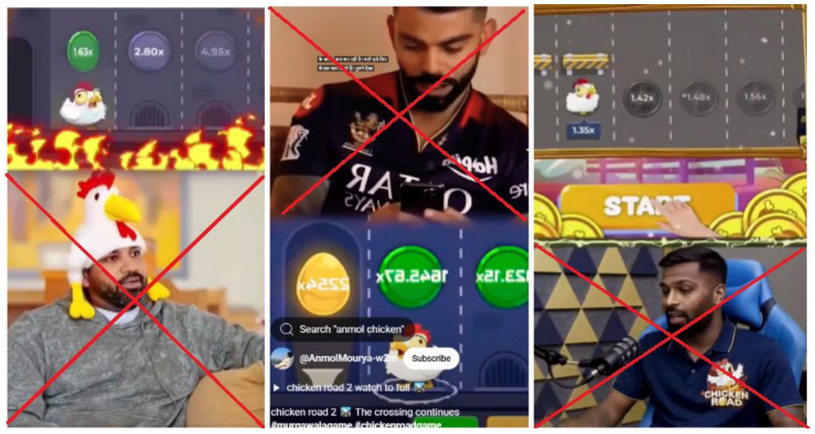
One of the victims, Veena, encountered a video on social media in which NR Narayana Murthy appeared to promote a trading platform with guaranteed high returns. Believing it to be genuine, she invested in the scheme, only to later realise that it was a scam[21].

Financial frauds perpetrated with AI and deepfakes are scaling fast in India, victimising all demographics and sectors, from senior citizens and students to startups and large companies. According to a response given in the Parliament by Bandi Sanjay Kumar, Minister of State (MoS), Ministry of Home Affairs, a staggering ₹22,845 crore ($2.66 billion) was lost to cyber frauds in India in 2024[20]. A lot of these frauds were AI-driven. Deepfake videos, voice clones, and AI-based phishing are making frauds more difficult to identify and simpler to amplify. Banks expose themselves to regulatory reprisal and brand harm; consumers risk losing life savings; insurers risk a deluge of false claims; and fintech companies, founded on trust in the digital space, are particularly at risk.

DEEPFAKE DECEPTION: FACT VS FICTION

Last year, two residents from Bengaluru fell victim to scams involving deepfake videos featuring Infosys Co-founder NR Narayana Murthy and Reliance Chairperson Mukesh Ambani. One of the victims, Veena, encountered a video on social media in which Murthy appeared to promote a trading platform with guaranteed high returns. Believing it to be genuine, she invested in the scheme, only to later realise that it was a scam[21].

Deepfake videos and voice clones of well-known personalities and public figures including Finance Minister Nirmala Sitharaman, Infosys Co-founder NR Narayana Murthy and his wife, Author and Philanthropist Sudha Murthy being used to promote fraudulent investments (Source: PTI News[22])



Deepfake videos and cloned audios of celebrity cricketers used to promote a gaming app (Source: GUARD Research)

In another instance, Ashok Kumar, a retired employee, was deceived after coming across a Facebook ad with a deepfake video of Indian businessman Mukesh Ambani. Convinced by the video's promises, he transferred ₹19,00,000 ($22,000) across various accounts. When the scammers stopped responding, Kumar realised he had been conned[23].

Sundararaman Ramamurthy, the Bombay Stock Exchange (BSE) Chief Executive, last year revealed that his deepfake videos had been employed by fraudsters to trick the general public and BSE staff. As a reaction, BSE released a public advisory for millions of retail investors in India to use sensibility when dealing with investment tips on social media platforms. These cautions came after incidents involving doctored videos and WhatsApp messages purporting to be Ramamurthy's spread far and wide, with fraudsters encouraging the onlooker to join fake investment groups. The application of hyper-realistic video, audio, and image editing highlights an increasing danger[24].

In 2023, PM Modi warned, "We have to be careful with new technology... These videos look very real and, therefore, we need to be very careful before believing the authenticity of a video or an image. India is emphasising a global framework for AI[25]."

GUARD's analysis of over 100 potential cases of misleading financial misinformation between April 1, 2025 and July 1, 2025 also finds that deepfakes of celebrities and public figures are still the most potent tool being used to promote bogus investment platforms and gaming applications.

20. The New Indian Express. (2025, July). Indian citizens lost over Rs 22,845 crore to online fraudsters in 2024: MHA in Lok Sabha. From https://www.newindianexpress.com/nation/2025/Jul/22/indian-citizens-lost-over-rs-22845-crore-to-online-fraudsters-in-2024-mha-in-lok-sabha
21. Hindustan Times. (2024, November). Two Bengaluru people fell prey to Narayana Murthy and Mukesh Ambani deep fake videos, loses close to Rs 90L. From https://www.hindustantimes.com/cities/bengaluru-news/two-bengaluru-people-fell-prey-to-narayana-murthy-and-mukesh-ambani-deep-fake-videos-loses-close-to-rs-90l-report-101730688063694.html
22. PTI News. From https://www.ptinews.com/fact-detail/Deepfake-Alert!-Nirmala-Sitharaman%E2%80%99s-Budget-interview-with-PTI%E2%80%99s-CEO---Editor-in-Chief-shared-with-superimposed-audio-to-falsely-claim-FM-endorsed-trading-platform/2685381; https://www.ptinews.com/fact-detail/pti-fact-check-no-narayana-murthy-sudha-murty-did-not-launch-any-investment-programme-digitally-altered-video-shared-with-fake-claim/1956611
23. Hindustan Times. (2024, November). Two Bengaluru people fell prey to Narayana Murthy and Mukesh Ambani deep fake videos, loses close to Rs 90L. From https://www.hindustantimes.com/cities/bengaluru-news/two-bengaluru-people-fell-prey-to-narayana-murthy-and-mukesh-ambani-deep-fake-videos-loses-close-to-rs-90l-report-101730688063694.html
24. Financial Times (2024, May). Deepfake 'menace' hurting retail investors, warns head of India's BSE bourse. From https://www.ft.com/content/c441d157-4e76-4374-b685-72c9b5213776
25. India Today. (2023, December). PM Modi deepfakes: Videos, photos, Smart India Hackathon, Rashmika Mandanna, Kajol. From https://www.indiatoday.in/india/story/pm-modi-deepfakes-videos-photos-smart-india-hackathon-rashmika-mandanna-kajol-2478140-2023-12-20; video available at https://www.youtube.com/watch?v=GpWS_v3BzsI

CHAPTER-3

AI and Deepfakes in Online Fraud: Emerging Threats
in the Digital Financial Ecosystem

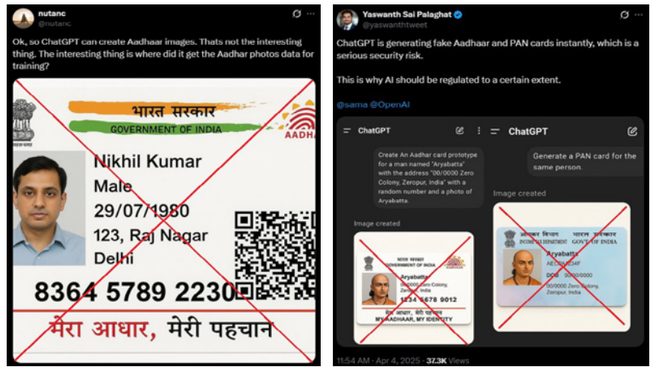
GUARD's research of claims going viral on online platforms, across different categories like news, politics, entertainment and others, which were published by various media organisations in India between November 2024 and March 2025, revealed that of the 4,800 stories analysed, 80 (nearly 2%) were finance-related.

GUARD's analysis of over 100 potential cases of misleading financial misinformation between April 1, 2025 and July 1, 2025 also finds that deepfakes of celebrities and public figures are still the most potent tool being used to promote bogus investment platforms and gaming applications. We found deepfakes of three popular sportspersons – Rohit Sharma, Virat Kohli, and Hardik Pandya – promoting fraudulent gaming platforms.

These videos collectively generated over 300,000 views across Facebook and YouTube, with one Hardik Pandya deepfake alone receiving 188,000 views. The realistic nature of these videos, combined with voice cloning technology, makes these extremely convincing for viewers.

Social media companies have strategically positioned themselves as 'platforms' rather than 'publishers', allowing them to avoid accountability for user-generated content. These platforms are also increasingly scaling back on fact-checking efforts and content moderation, creating a fertile ground for misleading information to thrive.

SEBI last year reported to the Parliament that it had flagged as many as 9,000 deceptive investment-related posts on platforms such as Facebook, Telegram, YouTube, and others, and asked these websites to take legal action against the offenders[26]. Meta now requires SEBI registration for all investment ads in India, targeting unregulated financial influencers on Facebook and Instagram[27].



AI-generated official documents (Source: X/@nutanc, X/@yashwanthtweet)

26. Financial Times (2025, February). India's stock market has an influencer problem. From https://www.ft.com/content/3be3c45b-6a05-4d1c-bf21-42d0d980d22c
27. The Economic Times. (2025, June). Meta puts finfluencers on notice; SEBI verification now mandatory for investment ads on Facebook, Instagram. From https://economictimes.indiatimes.com/markets/stocks/news/meta-puts-finfluencers-on-notice-sebi-verification-now-mandatory-for-investment-ads-on-facebook-instagram/articleshow/122159467.cms?from=mdr

Recent commercial development and advancements in generative AI technologies have made deepfakes significantly more advanced, accessible, and simple to create. A new threat stemming from advancements in AI image generation capability is the emergence and misuse of realistic-looking AI-generated government documents.

This increases the threat that bad actors will use deepfakes to impersonate executives and key staff members of an organisation, damage brands, defraud consumers, manipulate markets, or conduct information operations. The risk of disinformation, information circulated with the intention to mislead, has also increased with the advancement of these new technologies.

QUANTUM OF THE PROBLEM OF DEEPFAKES RELATED TO THE WORLD OF FINANCE

GUARD's research of claims going viral on online platforms, across different categories like news, politics, entertainment and others, which were published by various media organisations in India between November 2024 and March 2025, revealed that of the 4,800 stories analysed, 80 (nearly 2%) were finance-related. While finance remains a smaller slice of the overall misleading information online landscape, the nature of the content and platforms involved reveal worrying trends, particularly in the context of AI manipulation and social media amplification.

AI is beginning to play a significant role in this landscape with nearly 10% of the analysed content having AI-generated videos and AI-cloned audio, while around 3.75% had deepfake images. These formats, while fewer in number, present a serious threat due to their capacity to mislead and mimic authority with high realism and scale.

The presence of AI-driven content in nearly one-third of the total stories (combining deepfake video, AI audio, and images) indicates a critical shift in how financial scams are being perpetrated. These tools are increasingly being used to impersonate financial authorities, replicate trading apps, clone customer service voices, or simulate investment testimonials – making scams more believable and harder to detect. While it is becoming easier to generate realistic content using AI tools, there are no tech solutions yet that can comprehensively analyse and conclusively give a verdict of what is AI generated and what is not, across different forms of media and platforms used to generate the content. Currently, there are no standardised watermarks or AI content labelling tools that facilitate deepfake detection across different platforms. The detection tools at best are playing catchup.

AI is beginning to play a significant role in this landscape with nearly 10% of the analysed content having AI-generated videos and AI-cloned audio, while around 3.75% had deepfake images.

CHAPTER-4

BUILDING A TECH-DRIVEN DEFENCE

TOWARDS A 360° STRATEGY FOR DETECTION, PREVENTION AND AWARENESS

With increased reporting, monitoring and tracking there is a greater understanding of the modus operandi of different kinds of financial frauds, its genesis and trajectory. There is also now more effort being made on part of the Banking Financial Services Industry (BFSI), regulators, fintech companies and of course the government, to not just improve surveillance, tracking, cybersecurity, but also run awareness campaigns to inform consumers about the dos and don'ts, issue advisories and highlight grievance redressal mechanism. But the fact remains that with the onslaught of platforms distributing information, and with everyone turning content creators using the device in their hand, the basic guardrails, filtration and quality checks of content creation, production, and dissemination are missing. As a consequence, the users have access to more information than ever before but at the same time little understanding of its veracity or the capability to verify all that their feeds are being bombarded with. The constant quest for growing one's sources of income, higher returns, easier loans may stem from need or greed, but the plethora of options that the online platforms offer at the click of a button is both convenient and tempting.

GUARD deploys in-house AI tool and techniques and is working on agentic AI integrated workflows coupled with human oversight to speed up the detection and data analysis

CHAPTER-4

Building a Tech-Driven Defence: Towards a 360° Strategy for Detection, Prevention and Awareness

GUARD has also worked to develop a blueprint for an effective nationwide financial literacy campaign that reaches the last mile and factors impact evaluation into every stage, offering tailor-made solutions for a diverse and complex country like India.

Given the scale and spread of the problem, any isolated intervention on a standalone basis, to put checks and balances to stem the problem of online financial frauds will not have the required impact. The need of the hour is a concerted and collective effort to combat the problem at scale it deserves, by mobilising key stakeholders to devise a comprehensive plan to act in a swift and coordinated manner to prevent and not just counter financial fraud after it is committed.

GUARD has a strong social media monitoring unit to track, verify, evaluate and report content floating across different platforms. It deploys in-house AI tools and techniques and is working on agentic AI integrated workflows, coupled with human oversight, at every stage to speed up the tracking and analysis

Coupled with an inter-disciplinary team of OSINT experts, media professionals, public policy specialists, AI/ML engineers, data analysts, GUARD also has access to experts for specialised needs like deepfake detection. GUARD is in communication with institutions like IITs (whose research groups have developed multiple deepfake detection tools and techniques), and other research, tech, watermarking and labelling organisations that embeds forensic markers into content produced to facilitate identification, manipulation of content, to combat cloning and deepfakes.

NEED FOR COORDINATION AND COLLABORATION

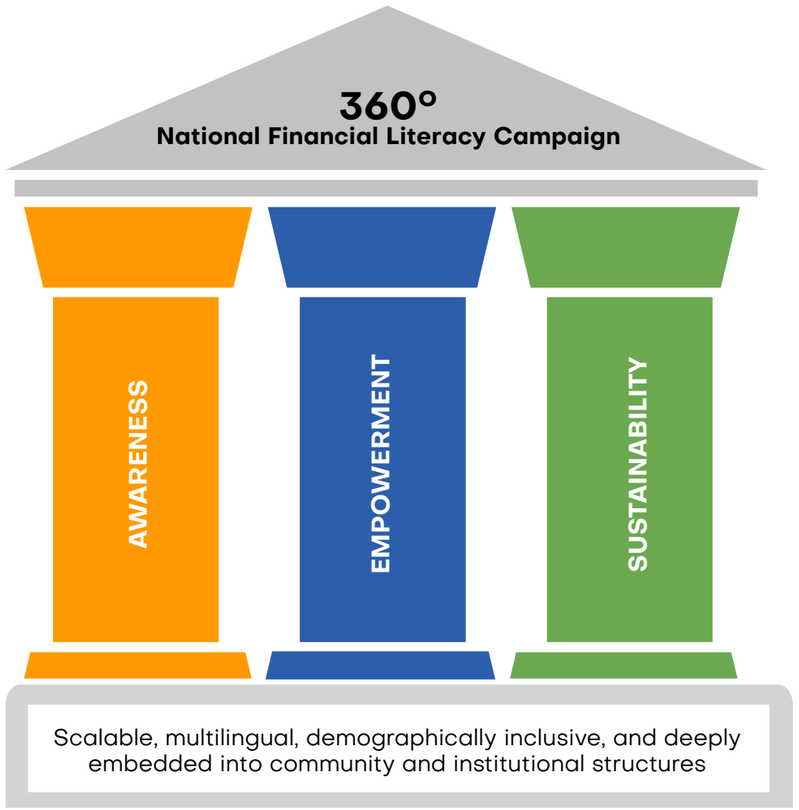
The model proposed by GUARD, envisages a series of solutions for different stakeholders. Risk evaluations and intelligence reports from GUARD can be made available to the consortium members at regular intervals. Some alerts may need to be real time if the risk is perceived to be high or immediate. Stakeholders can depute specific points of contact (POCs) to receive and take suitable action on the alerts sent, depending on the risk evaluation. The collective model will also ensure that the communication flows both ways. With members of the consortium not only having easy access to regular intelligence alerts on potential fraud threats but also contributing information for speedy verification and clarification of misleading information.

We recognise that while numerous marketing campaigns are run by financial institutions to raise awareness about the precautions to take while transacting online and the steps to follow in case of fraud, these efforts often operate in isolation and predominantly as a one-way flow of information. These initiatives may sometimes fall short in strategic depth, linguistic inclusivity, and technological flexibility, making it challenging to keep pace with rapidly evolving cyber threats.

What is required is not just periodic awareness drives, but an integrated and sustained national movement that embeds digital financial awareness into public consciousness, on the scale of a Swachh Bharat Abhiyan or the Polio eradication campaign. We welcome collaborations and partnerships to bring these visions to life.

GUARD has also worked to develop a blueprint for an effective nationwide financial literacy campaign that reaches the last mile and factors impact evaluation into every stage, offering tailor-made solutions for a diverse and complex country like India.

This concept proposes a 360° national campaign anchored in three core pillars: **Awareness, Empowerment, and Sustainability**. The model is designed to be scalable, multilingual, demographically inclusive, and deeply embedded into community and institutional structures.



Three Pillars of our Proposed Financial Literacy Campaign

This concept proposes a 360° national campaign anchored in three core pillars: Awareness, Empowerment, and Sustainability. The model is designed to be scalable, multilingual, demographically inclusive, and deeply embedded into community and institutional structures.

FUTURE-PROOFING AGAINST DIGITAL THREATS AND DECEIT

KEY TAKEAWAYS AND STRATEGIC IMPERATIVES FOR STAKEHOLDERS

India's digital financial revolution, while promising unparalleled convenience and inclusion, has inadvertently created fertile ground for misleading financial information and cyber fraud. As this report demonstrates, the human and economic costs are staggering, measured not only in rupees lost but in trust broken, lives disrupted and a growing erosion of confidence in financial systems.

This is not just a law enforcement issue; it is a systemic threat to India's economic resilience and digital aspirations. The evidence makes it clear: misleading information online is not a fringe concern – it is a central risk to financial health, institutional integrity, and national growth.

To counter the growing menace of online financial frauds, coordinated action is required from all stakeholders – government, regulators, financial institutions, technology companies, media, civil society and citizens. Based on the findings of this report, we recommend the following:

Effective oversight of online discourse requires combining continuous social media monitoring with rigorous risk evaluation by inter-disciplinary teams.

I. SOCIAL MEDIA MONITORING AND RISK EVALUATION

Effective oversight of online discourse requires combining continuous social media monitoring with rigorous risk evaluation by interdisciplinary teams across all platforms. By deploying the latest analytics platforms, natural-language processing engines and network-analysis tools, these teams can identify emerging patterns of fraud in real time.

Sustainable resilience stems from empowering communities to act as first-line defenders of their own financial security. Financial literacy campaigns should therefore go beyond broadcast messages and instead cultivate local ownership through village committees, self-help groups and peer-mentoring networks.

II. STRENGTHENING INTER-AGENCY COLLABORATION

Cyber-fraud and misleading financial information online cross both public and private sector boundaries, so robust inter-agency collaboration is essential. Formal information sharing protocols and joint task forces can align the efforts of law enforcement bodies, financial regulators, consumer protection agencies, telecom operators and civil society organisations. Regular coordination meetings and shared intelligence dashboards foster mutual trust, reduce duplication of effort and enable coordinated campaigns. In this way, each stakeholder contributes its unique expertise toward a cohesive national defence against digital threats.

III. INVESTING IN COMMUNITY-DRIVEN FINANCIAL LITERACY

Sustainable resilience stems from empowering communities to act as first-line defenders of their own financial security. Financial literacy campaigns should therefore go beyond broadcast messages and instead cultivate local ownership through village committees, self-help groups and peer-mentoring networks. Mobile-learning modules, community radio broadcasts and township workshops in regional languages reinforce practical skills such as verifying payment, URLs, recognising phishing lures and using grievance-redress apps to not just report frauds but also to verify posts, claims and ads they counter online.

IV. DEVELOPING INDIGENOUS LABELLING, WATERMARKING, AND DEEPFAKE DETECTION SOLUTIONS

Reliance on foreign proprietary tools can leave critical gaps in local threat response. Greater investment in home-grown R&D initiatives – spanning academic centres, technology startups and open-source communities – will yield tailored algorithms for labelling AI-generated content and detecting deepfakes in video and audio streams. By fostering collaboration between institutes such as IITs, industry partners and government labs, India can build adaptive platforms that understand regional languages, cultural contexts and device constraints.

V. REAL-TIME FRAUD AND DEEPFAKE REPORTING AND PREVENTIVE REDRESSAL MECHANISMS

To ensure barrier-free access, we propose 24/7 multilingual hotlines and WhatsApp tip-lines across India. Staffed by trained operators conversant in major regional languages, these channels will field reports of suspicious transactions, offer step-by-step verification guidance, and advise immediately on phishing or scam attempts. Complementing fraud reporting, a network of regional deepfake detection and reporting centers should be established. Each center, staffed by AI specialists, forensic analysts and legal advisors, will ingest user-flagged audio and video, apply automated screening blended with manual review, and publish public advisories within hours. A shared public portal will display active deepfake alerts and best practice guidelines, empowering media outlets and citizens to verify content authenticity and reduce the spread of manipulated media.

Strategic collaboration with private-sector and academic partners is critical. Coordination between technology firms, telecom operators, social-media platforms and research institutions will enable real-time intelligence sharing, harmonisation of reporting protocols, and integration of best-in-class detection tools. This will allow for a scalable response framework that safeguards every user and evolves alongside emerging cyber fraud and synthetic media threats.

KEY LEARNINGS

LEARNING	RECOMMENDATIONS
Fraud funnels are scripted and repeatable	Prevention can be systematised just as fraud is systematised.
Closed-group ecosystems and evasive tactics defeat single-platform takedowns and narrow, reactive actions	Counter-measures must operate cross-platform and persist after takedowns.
AI-driven deception is already mainstream	Detection and authentication tech are now baseline requirements Many banks and financial institutions will use AI. They need to find mechanisms to label it so that people are not misled.
Financial literacy is lagging	Campaigns must target the “confident but unaware”, not only the unbanked, and must be rooted in communities.
Regulation, enforcement and platform policy remain fragmented	A hub-and-spoke model of financial literacy campaigns is essential to reach the last-mile.

Strategic collaboration with private-sector and academic partners is critical. Coordination between technology firms, telecom operators, social media platforms and research institutions will enable real-time intelligence sharing, harmonisation of reporting protocols, and integration of best-in-class detection tools.

ABBREVIATIONS

AI	Artificial Intelligence
BFSI	Banking, Financial Services, and Insurance
BSE	Bombay Stock Exchange
GUARD	Global Unit for Analysis of Risk and Digital Threats
FY	Financial Year
I4C	Indian Cyber Crime Coordination Centre under Ministry of Home Affairs
IIT	Indian Institute of Technology
IMEI	International Mobile Equipment Identity
INR	Indian Rupee
KYC	Know Your Customer
NBFC	Non-Banking Financial Company
NCRP	National Cybercrime Reporting Portal
OSINT	Open Source Intelligence
PAN	Permanent Account Number
POC	Point of Contact
QR Code	Quick Response Code
RBI	Reserve Bank of India
SEBI	Securities and Exchange Board of India
SMS	Short Message Service
UPI	Unified Payments Interface

GUARD
GLOBAL UNIT FOR ANALYSIS
OF RISK AND DIGITAL THREATS

ABOUT GUARD

The Global Unit for Analysis of Risk and Digital Threats (GUARD) is a social listening unit and early warning platform of DataLEADS. GUARD's mandate is to analyse real-time data from online platforms, identify emerging patterns of financial fraud and AI-driven deception, assess risk impact and produce actionable intelligence for stakeholders. Leveraging DataLEADS' decade-long expertise in sociallistening and content monitoring, OSINT, deepfake detection and AI verification techniques, GUARD delivers insight to regulators, technology platforms, industry associations, banks, insurers and other financial institutions.

GUARD'S CORE APPROACH

- **360° risk monitoring:** Continuous surveillance across public social channels and closed messaging apps.
- **Interdisciplinary analysis:** Integrate AI/ML engineers, data analysts, policy specialists and media professionals.
- **Collaborative framework:** Partner with IIT research groups, forensic watermarking platforms and industry experts to enhance detection capabilities.

OUR SERVICES

- Early-warning tracking and alerts, verification of online financial claims including deepfake detection
- Community-anchored verification networks
- Intelligence-sharing platform & dashboards
- Pan-India multilingual financial literacy campaigns
- Agentic AI-integrated detection workflows
- To know more, write to us at **partnerships@dataleads.co.in**

data LEADS

DISCLAIMER: The report is the result of an extensive social listening and content monitoring exercise that completed in July 2025. It is not an exhaustive, quantitative cataloguing of all instances of misleading financial information online, but a qualitative analysis and insight based on the detailed open source intelligence gathering exercise. GUARD has taken every effort to ensure that the information shared in this report is accurate and verified, however, neither GUARD nor OW DataLEADS can accept responsibility or liability for how readers interpret or use the information in this report. All information provided in this report is accurate as of July 1, 2025, based on our knowledge at the time of publication.