ThreatMark

→ WHITEPAPER

AI VS FRAUD

The Future of Fraud Defense

Contents

Foreword

In 2025, AI is all around us—a buzzword, an assistant, a confidant, an opportunity, and, undeniably, a threat. And fraudsters are exploiting it to their advantage.

The news is hard to ignore. AI agents are becoming skilled scammers, AI-enhanced phishing and scam content is indistinguishable from that crafted by human experts, and AI-generated malware is evolving to evade detection—the list goes on.

This whitepaper examines the threats AI-enabled fraud poses to consumers and businesses worldwide. At the same time, it highlights how AI can serve as a powerful force in fraud prevention, helping financial institutions fortify their defenses. Naturally, any exploration of AI's role as a tool for combating fraud would be incomplete without addressing the evolving regulatory landscape that will shape its future adoption.

The potential is immense. Unlike traditional rule-based fraud detection, AI-powered systems offer unprecedented accuracy—even when detecting the rapidly expanding category of authorized fraud, a scheme that is notoriously difficult to catch using traditional methods.

At ThreatMark, we are committed to making AI a powerful ally in the fight against digital fraud because, as with any technology, its impact is determined by how we use it. However, keeping pace with evolving threats and opportunities requires continuous learning and adaptation.

That is the purpose of this whitepaper—to provide fellow fraud fighters and C-suite leaders in financial institutions with actionable insights and deepen their understanding of the current fraud landscape in the era of AI.

Wishing you a valuable read.

Michal Tresner
Co-founder & CEO, ThreatMark

1. Executive summary

This whitepaper examines AI's dual impact on the fraud landscape—both as a tool leveraged by fraudsters and as a critical asset for fraud prevention.

Technology is reshaping the fraud landscape, with AI at the forefront of both innovation and exploitation.

While many fraud trends, such as APP fraud, investment and crypto scams, and fraud industrialization, are not directly caused by AI, the technology significantly amplifies their impact—making fraud more persuasive, scalable, automated, and ultimately far more dangerous.

This acceleration is driven by AI's rapid advancement in fraud techniques. The number of deepfake videos is doubling every six months, while AI-generated voice cloning is becoming increasingly prevalent, with one in four people worldwide now encountering a voice scam—directly or via someone they know. AI also facilitates synthetic identity fraud, creating highly realistic fake identities that bypass verification processes, leading to billions in losses. Meanwhile, AI-generated malware is a nascent but growing threat, capable of adapting its behavior to evade detection. Additionally, AI-enhanced phishing enables hyper-personalized attacks that closely mimic legitimate communications, achieving success rates comparable to those of human-crafted scams.

As AI-enhanced fraud grows more sophisticated, traditional rule-based detection is proving increasingly ineffective, compelling financial institutions to adopt advanced AI-powered solutions for greater accuracy and adaptability.

Techniques such as behavioral profiling, anomaly detection, and predictive analytics strengthen fraud detection by analyzing user interactions, identifying suspicious patterns, and anticipating new attack methods. Meanwhile, the industry is exploring how Generative AI and General-Purpose AI could further transform fraud prevention.

While AI adoption in fraud detection is gaining momentum, several challenges remain. Privacy and compliance concerns, legacy systems, data integration issues, talent shortages, and implementation costs continue to hinder banks' efforts to enhance their fraud detection capabilities. To overcome these obstacles, financial institutions are increasingly turning to third-party solutions and cloud-based fraud detection platforms, which offer the flexibility to bridge the expertise gap, adapt to evolving fraud tactics, and maintain a strong cost-benefit ratio.

With the fragmented global regulatory landscape—the EU's AI Act tightening oversight, the US favoring a deregulated path, and PSD3 on the horizon—leveraging partnerships and intelligence-sharing remains essential to navigating this dynamic environment.

Collaboration with regulators, technology providers, and industry peers will be critical for staying ahead as AI's impact on fraud and its prevention continues to evolve—through 2025 and beyond.

2. Introduction: The fraud landscape

The evolving fraud landscape is deeply shaped by societal, economic, and technological shifts. While factors like the growing adoption of cryptocurrencies, rising cost-of-living pressures, and increasingly sophisticated criminal networks play a role, it is technological advancements—especially AI—that are redefining fraud on an unprecedented scale. From deepfakes to automation, AI is not just influencing fraud; it is transforming how it is executed, making it more sophisticated, scalable, and harder to detect.

2.1 Current fraud trends and AI's impact

Authorized push payment (APP) fraud

Authorized Push Payment (APP) fraud is rapidly increasing, fueled by a paradox—rather than hacking systems, fraudsters leverage technology to exploit human vulnerabilities. Unlike fraud methods that involve unauthorized access, APP fraud circumvents security measures by making victims the unwitting initiators of the transactions. These scams exploit deception, urgency, and psychological manipulation to trick victims into willingly transferring funds to fraudulent accounts.

Scam-related fraud has surged, with its share of overall fraud increasing by 56% and financial losses soaring by 121%. Scams now account for 23% of all fraudulent transactions in the US. With a 12% compound annual growth rate (CAGR), APP fraud alone is projected to reach $7.6 billion in losses by 2028, cementing its status as one of the fastest-growing financial threats worldwide.

AI's growing role in APP fraud

While APP fraud primarily relies on social engineering, technology is accelerating its growth in multiple ways. The share of APP fraud committed via instant payment rails is rising and is projected to account for 80% of total APP fraud losses by 2028.
Fraudsters are also increasingly leveraging AI to refine victim targeting, personalize attacks, automate scams, and enhance their overall persuasiveness. A major factor driving this surge is the growing use of deepfake technology—both video and voice—which significantly improves impersonation tactics and can be deployed across nearly all types of APP fraud (**see Section 3.1**).

Rising crypto-related fraud

Crypto scams exploit the hype and complexity of digital assets to deceive victims into transferring funds, revealing private keys, or investing in fraudulent projects. These schemes are becoming increasingly sophisticated, with pig-butchering scams emerging as a prime vector—long-term frauds that blend crypto and romance scams. Fraudsters build trust with victims, often through social media or dating platforms, before persuading them to invest in fake cryptocurrency schemes.

Since 2020, crypto scam activity has increased by an average of 24% annually, with 2024 revenue projected to surpass $12 billion—a new record. Pig-butchering scams alone have seen a 40% year-over-year increase, while cryptocurrency fraud continues to expand into new, specialized variations.

AI's growing role in crypto scams

Since crypto scams typically rely on social engineering, their use of AI follows similar patterns to APP fraud—enhancing effectiveness, automation, scale, and perceived legitimacy. This makes them increasingly difficult to detect for both victims and financial institutions. AI enables fraudsters to create ultra-realistic fake trading platforms, craft highly personalized phishing content, impersonate legitimate exchanges and wallet providers, and even generate deepfake videos of financial experts and celebrities to lure victims into scams.

Fraud industrialization

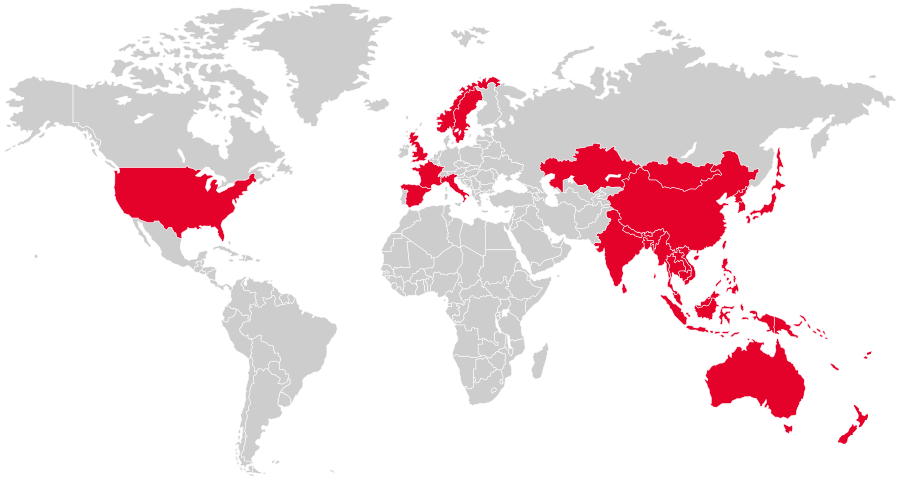
Fraud has evolved from small-scale actors into a highly organized, large-scale industry, where criminals operate with scalable, automated, and professionalized methods. Many fraud rings now function as legitimate businesses, complete with call centers, supply chains, and specialized roles. As the fraud ecosystem becomes more structured, entry barriers for scammers have dropped significantly.

Fraud-as-a-Service (FaaS) has further accelerated this shift by offering ready-made fraud tools, services, and expertise on underground markets—effectively serving as a one-stop-shop for cybercriminals. This model enables non-technical actors to execute sophisticated scams, such as phishing, without requiring advanced hacking or financial crime knowledge.

AI's growing role in fraud industrialization

AI is, without a doubt, accelerating the growth of the FaaS ecosystem. One example of this expansion is Huione Guarantee, an online forum and peer-to-peer (P2P) marketplace affiliated with Huione Group, which has become a central hub for cybercriminals seeking scam infrastructure. An analysis of crypto flows from 2021 to 2024 shows that Huione's scam service providers have experienced exponential revenue growth, with AI service vendors' earnings surging 1,900%—a clear indicator of the rapid expansion of AI-driven fraud facilitation.4

2.2 AI-enabled fraud across the globe: Regional trends and insights



Europe

Almost **43%** of detected fraud attempts on financial institutions now use AI. Of these, **29%** are successful.

Italy

In the first half of 2024, online fraud caused damages soaring to **€114 million—a 71% year-on-year increase**. In February 2025, scammers used an AI-generated voice to impersonate Italy's Defence Minister in an attempt to trick some of the country's top tycoons into wiring money overseas.

Spain

In December 2024, Spanish police dismantled a gang that used AI to scam companies and launder money. The criminals altered money mules' IDs, retaining their photos but replacing personal data with third parties' information.

France

The overall loss suffered by victims of financial fraud is at least **€500 million a year**. In a long-term romance scam, a French woman was conned out of **€830,000 ($850,000)** by deepfakes impersonating actor Brad Pitt.

United States

Generative AI could drive fraud losses from **$12.3 billion in 2023** to **$40 billion by 2027**, growing at a compound annual rate of **32%**.

United Kingdom

Nearly one in three UK adults suspects they were targeted by an AI voice cloning scam in the past year.

The Nordics

As much as **46%** of Norwegians and **48%** of Swedes are quite or very concerned that AI will increase the risk of fraud.

APAC

Between February and June 2024, mentions of deepfakes in cybercriminal Telegram channels and underground forums surged by **600%**.

Australia

In the past year, deepfake scams have targeted **20% of businesses** and **36%** of Australian consumers, resulting in estimated losses of tens of millions of dollars.

3. AI-enabled fraud: How fraudsters use AI to evade detection

Artificial intelligence is a broad term that encompasses various subfields and applications, including machine learning, deep learning, and generative AI (see Glossary). While AI's role in fraud is sometimes obvious—as with deepfakes—it can also be more subtle and harder to detect, such as in AI-enhanced phishing content. This chapter explores some of the most common AI-enabled fraud tactics.

3.1 Deepfake videos

AI type: deep learning, generative AI

Deepfake refers to using AI—specifically deep learning techniques and generative adversarial networks (GANs)—to create hyper-realistic, but fabricated media files.

Deepfakes are arguably the most recognizable and impactful AI-generated scams. These highly convincing, artificially created videos are a powerful tool for social engineering.

Deepfake videos are widely used in impersonation scams, from sophisticated CEO/CFO fraud to fake celebrity endorsements promoting get-rich-quick and crypto schemes. However, assuming that deepfake scams only target celebrities or corporate executives would be a mistake. With social media enabling widespread video sharing, virtually anyone can provide material for deepfakes—creating new opportunities for scammers to exploit.

Additionally, deepfakes pose a significant challenge to KYC processes by enabling fraudsters to bypass liveness detection. Using AI-generated deepfake videos, fraudsters can attempt to deceive verification systems that rely on facial recognition and movement-based authentication. While some liveness checks require users to perform specific actions—such as turning their heads or blinking—advanced deepfake technology can mimic these behaviors with increasing accuracy.

The data behind deepfake videos

The volume of deepfakes circulating online doubles every six months, with projections indicating around 8 million deepfakes will be shared in 2025.

In the financial sector, deepfake incidents surged by 700% in 2023 compared to the previous year.

Elon Musk is the most frequently impersonated celebrity in deepfake scams.

How a deepfake scam works: A real-life example

Nikki MacLeod, 77, lost £17,000 in a deepfake-powered romance scam. She met a woman named Alla Morgan, an alleged oil rig worker, in an online chat group. Initially skeptical, Nikki was convinced by highly realistic video messages, unaware they were AI-generated deepfakes and that Alla was merely a fabricated persona created by scammers to defraud her.

What started as small requests for Steam gift cards escalated to large bank and PayPal transfers. The deception deepened when a scammer posing as an HR representative claimed Alla needed funds for a helicopter transport from the rig.

The fraud only ended when Nikki's bank flagged a transaction and uncovered the scam. This case highlights how deepfake technology makes scams more convincing—even for cautious victims—and underscores the critical role of banks in fraud prevention.

> **"I dread to think how deepfakes are going to affect the world of romance fraud because it won't just be celebrities whose identities scammers use, but people from all professions."**

Becky Holmes
An expert on romance scams and author of Keanu Reeves Is Not In Love With You: The Murky World of Online Romance Fraud

3.2 Voice cloning

AI type: deep learning, generative AI

Voice cloning is an AI-driven technique used to generate a synthetic replica of a person's voice.

While deepfake videos have gained widespread attention, deepfake audio (or voice cloning) is just as dangerous, if not more so, as it is often more believable and significantly cheaper to produce, making it widely used by fraudsters. They leverage voice cloning in various impersonation and social engineering scams, from CEO/CFO fraud to grandparent scams, where criminals mimic a distressed family member to deceive elderly victims. There have also been reports of voice cloning being used for extortion.

The data behind voice cloning

One in four people worldwide has encountered an AI-generated voice scam, either directly or through someone they know.

More than one in three organizations worldwide (37%) reported being targeted by a deepfake voice scam.

A study found that in 27% of cases, people were unable to distinguish deepfake audio from real recordings.

How a voice cloning scam works: A real-life example

Anthony [last name withheld] lost $25,000 after scammers used AI voice cloning to impersonate his son and manipulate him into sending money. It started with a phone call: a voice identical to his son's claimed he had been in a car accident, injuring a pregnant woman. In reality, it was an AI-generated voice clone.

Minutes later, a scammer posing as a lawyer called, saying his son had been arrested and needed $9,200 for bail—urgently. When Anthony tried to call his son, it went to voicemail, reinforcing his belief that he was in jail.

Panicked, he withdrew the money from his bank, telling staff it was for solar panels to avoid suspicion. Soon after, another scammer called with devastating news: the pregnant woman had died, and bail had increased. Anthony withdrew another $15,800 and handed it to a stranger in an Uber. Only later, when the panic subsided, did he realize he had been scammed.

> **"Detecting voice cloning is incredibly difficult. It's not like installing an antivirus and being safe. Fraudsters can now clone voices with just seconds of audio, making this an ongoing challenge for financial institutions."**

Romano Ramanti
Certified Ethical Hacker, Swiss Banking Representative in Payment Scheme Fraud Prevention Working Group, European Payment Council

3.3 Synthetic identity fraud

AI type: deep learning, generative AI

Synthetic identity fraud occurs when fraudsters create a fake identity by combining real and fabricated personal information.

Synthetic fraud poses a significant threat to financial institutions. Fraudsters create synthetic identities, often called Frankenstein IDs, by exploiting data breaches—blending stolen details like Social Security numbers with fabricated names, addresses, and birthdates. These identities can bypass banks' application and verification processes, enabling criminals to open accounts, apply for loans, and commit various forms of fraud. With AI enabling more sophisticated schemes, this threat is expected to rise rapidly.

The data behind synthetic identity fraud

Synthetic identity fraud is the fastest-growing financial crime in the United States.

Synthetic identity fraud is projected to cause at least $23 billion in losses across the US by 2030.

The average payoff from synthetic identity fraud is estimated to range between $81,000 and $98,000.25

How synthetic identity fraud works: A real-life example

A Canadian man opened a bank account under one name and, a month later, attempted to open another under a different name at the same branch. A sharp-eyed teller at a downtown west Toronto bank grew suspicious and alerted police. Upon his arrest, the man confessed to the fraud, and officers found 20 genuine IDs under two fake names in his wallet, including multiple credit and debit cards, a driver's license, a social insurance card, and a Canadian citizenship card.

He admitted to being recruited into a scheme where his handler took him to driver's license facilities to obtain fraudulent IDs, which were then used to open bank accounts across Toronto. He was promised $5,000 per account.
His arrest launched Operation Mouse, a five-month investigation that uncovered $25 million in fraud losses linked to synthetic identities, with credit card bills and mortgages left unpaid.

Unlike traditional identity fraud, which is typically a one-time quick hit before being uncovered, synthetic identity fraud is a long-term scheme that can take years to develop. Fraudsters often start by applying for a store credit card using a fabricated identity, expecting rejection. However, the application itself creates a credit record, effectively validating the fake identity.

With this record in place, fraudsters can apply elsewhere, eventually securing a credit card. They make purchases and repay balances to build a strong credit profile, making the fictitious identity appear legitimate. Over time, this allows them to access larger loans, credit lines, and assets, leading to significant financial fraud.

"Synthetic identities are like Frankenstein's monster, stitched together from stolen and fake data to deceive banks and build credit over time. With AI making these fakes even more lifelike, fraudsters are exploiting financial systems at an alarming scale. At ThreatMark, we leverage advanced behavioral AI to detect these artificial identities before they cause harm, helping banks stay ahead of this evolving threat."

Adam Parry
Fraud Fighter at ThreatMark

3.4 AI-generated malware

AI type: generative AI, deep learning, machine learning

AI-generated malware refers to malicious software that leverages artificial intelligence to evade detection, adapt in real time, and automate attacks.

Unlike traditional malware, which follows predefined code and fixed behaviors, AI enables the creation of polymorphic malware—capable of adapting its behavior or appearance to evade detection. This adaptability makes it particularly effective against traditional antivirus measures, increasing the risk for financial institutions and their clients. Reports indicate that OpenAI's ChatGPT has already been jailbroken to generate polymorphic malware. ,

Moreover, AI enhances Malware-as-a-Service (MaaS) platforms by enabling automated malware customization and distribution, making attacks more scalable and accessible to less skilled criminals.

The data behind AI-generated malware

The UK's National Cyber Security Centre warns that AI could generate malware sophisticated enough to bypass current security filters, particularly when trained on high-quality exploit data.

Up to 60% of IT professionals worldwide cite AI-enhanced malware as the most concerning AI-driven threat for 2025.

Nearly half (48%) of CISOs in the UK and US consider AI-powered ransomware attacks a major concern.

How polymorphic malware works: A real-life example

Emotet (also known as Heodo) is a prime example of polymorphic malware, modifying its code with each access to evade detection. First appearing in 2014 as a banking Trojan, it evolved into a modular malware loader, commonly delivering information stealers, remote access trojans, and ransomware.

Spread primarily through automated phishing campaigns, Emotet emails have masqueraded as invoices, shipping notices, and COVID-19 updates, often containing malicious Word documents. Once opened, users were prompted to enable macros, triggering the malware's installation and compromising their systems.

Despite Europol's major takedown in January 2021, Emotet re-emerged later that year and remained active. By October 2022, Dark Reading reported its variants had multiplied, with researchers identifying over 21,000 invocation chains and 139 unique program chains, highlighting its rapid evolution.

As of February 2025, there is no clear evidence that fraudsters have integrated AI into Emotet or other malware. However, Emotet operates under a Malware-as-a-Service (MaaS) model, enabling cybercriminals to distribute malicious payloads through its infrastructure.

> **"Since polymorphic malware continuously alters its code while maintaining its core functionality, it renders traditional, signature-based security measures ineffective. To counter this threat, security solutions must employ advanced detection methods—such as heuristic analysis and AI-driven threat detection—that identify suspicious behavior rather than relying solely on known signatures."**

Lukas Jakubicek
Fraud Fighter & Lead Principal Solution Consultant at ThreatMark

3.5 AI-enhanced phishing

AI type: generative AI (LLMs, NLP)

AI-enhanced phishing is a cyberattack where fraudsters use artificial intelligence to generate highly personalized, automated, and convincing phishing content.

AI is transforming phishing into a more sophisticated and scalable cyber threat. Many security experts consider phishing powered by large language models (LLMs) to be one of the most pressing AI-driven risks today, as AI helps automate the entire phishing process.

LLMs enable fraudsters to generate highly personalized phishing content that seamlessly mimics brands or authorities while eliminating a common and telling sign of phishing—language deficiencies—and incorporating up-to-the-minute information from various sources to enhance credibility.

In addition to crafting deceptive messages, AI facilitates the rapid deployment of phishing sites, allowing fraudsters to create convincing fake websites at scale while using techniques that evade traditional security measures meant to detect them. These advancements make phishing attacks more convincing, cost-effective, and far more sophisticated than traditional methods.

The data behind AI-enhanced phishing

A study showed that AI-generated phishing content reaches similar success rates (65%) as phishing messages crafted by human experts.

While achieving equal or greater success rates, AI reduces costs of AI phishing attacks by more than 95%.

While the role of AI in phishing is often difficult to confirm, a report shows a 1,265% rise in phishing attacks potentially linked to AI tools.

The median time for a user falling for a phishing email is just 60 seconds.

How AI-enhanced phishing works: A real-life example

Sam Mitrovic, a Microsoft solutions consultant, received a Gmail account recovery request—a classic phishing tactic. Recognizing the scam, he ignored it. Forty minutes later, he missed a call from "Google" in Sydney, Australia.

A week later, the pattern repeated—another recovery request followed by a call. This time, Sam answered. A convincing American voice claiming to be from Google Support warned of suspicious activity on his account, asking if he had logged in from Germany. When he denied it, the scam escalated: the caller claimed an attacker had been accessing his account for seven days and had already downloaded his data.

While on the call, Sam Googled the number—it led to Google's business pages. Moments later, the scammer sent a realistic email from a Google-like domain, cleverly spoofed to appear legitimate.

The deception was nearly flawless, except for the "To" field containing a subtly altered address that wasn't actually a Google domain. But one detail finally exposed the fraud. The caller said "hello" twice—too perfectly spaced and pronounced. That's when Sam realized: it was an AI-generated voice. Had he stayed on the call longer, the next step would have been to approve the recovery request—handing over control of his account to scammers.

This case highlights how AI-powered vishing is becoming alarmingly realistic, capable of deceiving even tech-savvy professionals.

“AI is transforming phishing into an on-demand service. With Phishing-as-a-Service (PhaaS) and AI-Scams-as-a-Service (AISaaS), fraudsters can now launch hyper-personalized attacks at scale—automating deepfake calls, AI-generated emails, and real-time scam interactions. These attacks adapt dynamically, making traditional defenses ineffective. The only way to counter AI-driven phishing is with AI-driven detection—identifying behavioral anomalies, attack patterns, and real-time deception tactics.”

Lubos Klinko
Fraud Fighter & Head of Cyber Fraud Fusion Center at ThreatMark

4. AI Use Cases in Fraud Detection: How AI systems can detect and mitigate fraud before it occurs

As fraud tactics grow more sophisticated and AI becomes embedded in every stage of attacks—from victim targeting and execution to money laundering—traditional rule-based detection systems are falling short. To effectively detect and prevent fraud, banks must fight fire with fire. Research by PYMNTS Intelligence shows that 71% of financial institutions now leverage AI and machine learning, as these systems offer advanced capabilities to uncover complex schemes, detect patterns and anomalies, and predict emerging fraud trends.

4.1 Behavioral profiling

AI enables behavioral profiling by analyzing how users interact with digital platforms, creating detailed usage patterns. By examining extensive customer data—such as device behavior (typing speed, mouse movements), transaction history, and session patterns—behavioral profiling detects anomalies and deviations from established norms. This approach provides deeper, context-aware, and more personalized fraud protection, enhancing detection accuracy and risk assessment.

Behavioral profiling can reduce reliance on traditional multifactor authentication (MFA) by serving as a passive, frictionless authentication layer that can replace weaker MFA factors—such as SMS codes—while maintaining a smooth user experience. By continuously analyzing user behavior, behavioral profiling can also help ensure regulatory compliance, such as meeting Strong Customer Authentication (SCA) requirements under PSD2.

Behavioral profiling is effective against AI-driven fraud because it analyzes dynamic, individualized interaction patterns rather than relying on static identifiers like passwords or personal data, which can be compromised. By evaluating a wide range of behavioral factors, it can detect inconsistencies that reveal unnatural behaviors, expose synthetic identities, identify unauthorized account access, and flag anomalies in user interactions that may indicate scam manipulation.

ThreatMark

ThreatMark's approach

ThreatMark's Behavioral Intelligence Platform leverages a proprietary AI-driven risk engine that not only builds individual user profiles but also correlates behavior across multiple devices and sessions. This multi-dimensional approach allows banks to detect fraud attempts that involve account takeovers across different devices, reducing false positives and improving fraud detection rates. Beyond fraud detection, the Behavioral Intelligence Platform can also reduce authentication costs—such as SMS verification expenses—by up to 90% while helping financial institutions meet SCA requirements under PSD2.

4.2 Anomaly detection

AI enables behavioral profiling by analyzing how users interact with digital platforms, creating detailed usage patterns. By examining extensive customer data—such as device behavior (typing speed, mouse movements), transaction history, and session patterns—behavioral profiling detects anomalies and deviations from established norms. This approach provides deeper, context-aware, and more personalized fraud protection, enhancing detection accuracy and risk assessment.

Behavioral profiling can reduce reliance on traditional multifactor authentication (MFA) by serving as a passive, frictionless authentication layer that can replace weaker MFA factors—such as SMS codes—while maintaining a smooth user experience. By continuously analyzing user behavior, behavioral profiling can also help ensure regulatory compliance, such as meeting Strong Customer Authentication (SCA) requirements under PSD2.

Behavioral profiling is effective against AI-driven fraud because it analyzes dynamic, individualized interaction patterns rather than relying on static identifiers like passwords or personal data, which can be compromised. By evaluating a wide range of behavioral factors, it can detect inconsistencies that reveal unnatural behaviors, expose synthetic identities, identify unauthorized account access, and flag anomalies in user interactions that may indicate scam manipulation.

ThreatMark

ThreatMark's approach

ThreatMark's Behavioral Intelligence Platform leverages a proprietary AI-driven risk engine that not only builds individual user profiles but also correlates behavior across multiple devices and sessions. This multi-dimensional approach allows banks to detect fraud attempts that involve account takeovers across different devices, reducing false positives and improving fraud detection rates. Beyond fraud detection, the Behavioral Intelligence Platform can also reduce authentication costs—such as SMS verification expenses—by up to 90% while helping financial institutions meet SCA requirements under PSD2.

4.3 Predictive analytics

While anomaly detection helps identify irregular activities as they occur, predictive analytics takes fraud prevention further—analyzing historical patterns to anticipate and mitigate threats before they escalate.

Predictive analytics leverages deep learning, graph analytics, and time-series forecasting to model emerging fraud trends before they materialize. By identifying risk patterns within transaction flows, AI can detect fraud networks and anticipate fraudulent behaviors, allowing banks to take proactive action before an attack occurs.

Unlike reactive fraud detection, AI-powered systems dynamically adapt to new threats. This approach helps financial institutions uncover emerging fraud schemes, reduce financial losses, and improve detection accuracy while minimizing false positives.

ThreatMark

ThreatMark's approach

Machine learning and AI empower ThreatMark's Behavioral Intelligence Platform to detect both known fraud patterns and emerging threats, continuously refining its accuracy as it processes more data. This adaptive learning enables the platform to stay ahead of evolving fraud tactics, ensuring it remains highly effective and responsive to the latest threats. As a result, these capabilities increase fraud detection rates by up to 70% compared to traditional FDS.

Beyond identifying emerging threats, the Behavioral Intelligence Platform also strengthens fraud defenses by gathering intelligence on attackers' infrastructure, devices, tools, locations, payment methods, and attack vendors, enabling financial institutions to build proactive defenses against future fraud. By disrupting fraud infrastructure and crime ring operations targeting other bank customers, the Platform helps reduce fraud losses, delivering long-lasting, strategic impact.

4.4 Emerging trends

Generative AI

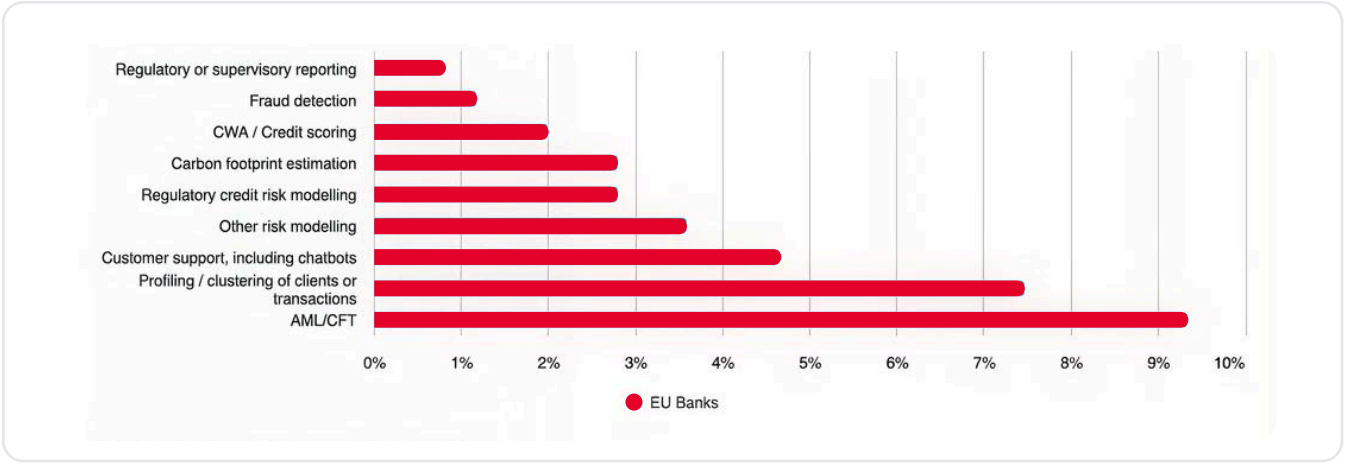
The banking industry is one of the sectors poised for the greatest impact from generative AI, which could deliver an additional $200 billion to $340 billion in annual value if fully implemented, according to McKinsey. While customer service, code generation, and productivity enhancement are cited as the most imminent generative AI opportunities, fraud prevention is also a key potential use case.

Generative AI could help banks better integrate unstructured data, enabling more efficient identification of new patterns and anomalies at both the customer and market levels. There are also creative applications designed to distract fraudsters, such as "AI Granny Daisy"—a custom-made, human-like chatbot that answers calls in real time, keeping fraudsters on the phone as long as possible to prevent them from targeting real victims.

While Daisy is primarily a creative PR tactic to raise awareness about scam threats and AI's potential—rather than a standalone fraud prevention solution—it underscores the diverse ways generative AI can be leveraged to combat fraud: by disrupting scammers' workflows, collecting intelligence on scam tactics, or even integrating with fraud detection systems.

GPAI

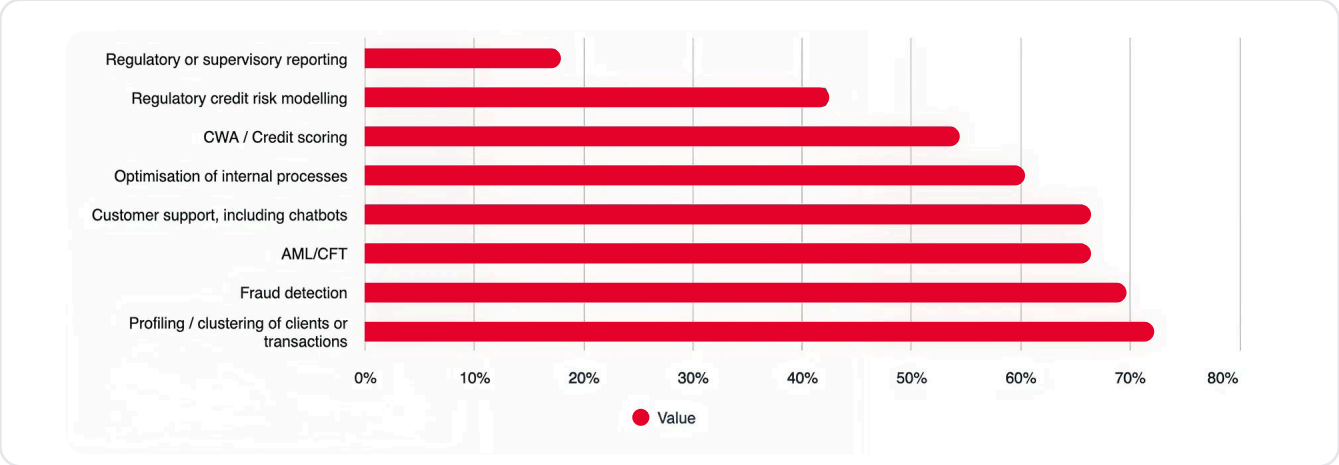
One emerging example of AI's potential in banking is General-Purpose AI (GPAI), though its adoption remains in the early stages. However, as banks explore its capabilities, concerns around explainability and data security, regulatory compliance will dictate its practical applications. These concerns reflect broader regulatory efforts to balance AI-driven innovation with financial security, as seen in the EU's AI Act. Around 10% of EU banks are already testing GPAI for various use cases, including AML/CFT, as well as client and transaction profiling and clustering.



Source: EBA Risk Assessment Questionnaire44

5. Challenges in adopting AI

In 2023, financial firms invested $35 billion in AI, with spending projected to hit $97 billion by 2027. While AI extends beyond fraud management, its adoption comes with challenges. Overcoming them ensures AI-driven fraud detection reduces losses, boosts efficiency, strengthens compliance, and improves user experience while delivering strong ROI.



Source: EBA Risk Assessment Questionnaire44

5.1 Personal data and privacy

McKinsey reports that 43% of professionals see personal privacy as a major risk of generative AI, underscoring the challenge of balancing AI adoption with data protection. Since fraud detection often operates across jurisdictions, compliance with data protection and financial regulations such as GDPR, CCPA, and PSD2 is crucial.

To mitigate these risks, financial institutions should adopt privacy-preserving AI techniques like federated learning, differential privacy, and encryption to protect sensitive data without compromising confidentiality.

ThreatMark

ThreatMark's approach

ThreatMark recognizes the critical need to balance AI-driven fraud detection with stringent data privacy regulations such as GDPR and PSD2. To address these concerns, ThreatMark leverages privacy-preserving AI techniques that ensure robust fraud detection while maintaining compliance and user trust.

Advanced encryption and secure data handling

ThreatMark integrates end-to-end encryption and secure data anonymization techniques to safeguard sensitive financial information from breaches. The platform is designed to operate without directly handling personal data—only the bank itself knows the identity of the end user. This approach enables high-performance fraud detection, ensures compliance with regulatory frameworks, and maintains strict privacy and security standards.

5.2 Technical and operational challenges

Quality of input data

The effectiveness of AI-driven fraud detection depends heavily on the quality, consistency, and integration of input data. Poor data can skew AI models, leading to reduced overall efficiency. Key challenges include:

Data fragmentation and silos

Many financial institutions operate with legacy banking platforms and disconnected data sources, making it difficult to integrate information across multiple channels. Siloed data can create gaps in fraud detection, as AI models may lack access to necessary digital cues, such as transaction context or behavioral signals.

Transparency and data bias

AI models learn from historical fraud data, which can carry biases. If the data disproportionately flags certain regions or demographics, AI may reinforce discrimination. This raises compliance risks under GDPR, CCPA, and the upcoming EU AI Act, which mandate fairness in automated financial processes.

False positives vs. false negatives

Inaccurate data can block legitimate transactions or let fraudulent payments slip through. Striking the right balance requires high-quality, standardized data and adaptive AI models.

Data standardization

Clean, properly labeled data is crucial for accurate AI-driven fraud detection. Challenges include integrating diverse data sources; handling real-time data streams with varying formats and inconsistencies; and eliminating erroneous data that can mislead AI models.

The data challenge also forces banks to decide between in-house fraud detection or third-party solutions. In-house machine learning models can be tailored to specific markets, but third-party models—trained on diverse, large-scale datasets—offer greater adaptability to the rapidly evolving fraud landscape and the global nature of financial crime. Banks must carefully weigh these factors when deciding on their fraud detection strategy.

Skilled professionals

The rapid integration of advanced AI technologies in financial services, especially for fraud detection, has significantly increased the demand for skilled professionals. However, the supply of qualified talent has not kept pace, leading to a widespread AI talent shortage across the industry. In 2024, the hiring gap was estimated at 50% of all required AI roles. The growing demand for AI talent in banking has intensified competition among financial institutions, driving a bidding war for skilled professionals.

Deploying third-party anti-fraud systems can help financial institutions bridge the AI talent gap, especially since fraud detection requires deep expertise in both technology and financial crime. Partnering with third-party vendors provides stability by reducing reliance on in-house risk engineers, mitigating the disruption caused by key talent leaving for competitors.

Integration obstacles

Implementing AI-driven fraud detection in banks presents several integration challenges, particularly when dealing with legacy infrastructure. Banks also face IT resource constraints, as AI models require significant computational power and ongoing system updates.

A strategic approach is to utilize third-party AI vendors and cloud-based fraud detection platforms. These solutions enable seamless integration with existing banking systems while ensuring compliance with evolving regulations. By leveraging external expertise, banks have a significant opportunity to use AI to accelerate the transition away from legacy applications—previously too costly to replace—and drive forward their digitalization efforts.

ThreatMark

ThreatMark's approach

ThreatMark provides a robust AI-driven fraud detection ecosystem that tackles the technical and operational barriers financial institutions face, including data quality, skill shortages, and integration complexities.

Data quality

By aggregating and standardizing data from multiple sources, ThreatMark enhances data quality and minimizes bias, helping financial institutions break down silos and build a more comprehensive fraud detection framework.

Fraud expertise

For banks struggling with the AI talent gap, ThreatMark offers pre-trained fraud models that eliminate the need for extensive in-house expertise, making advanced fraud detection more accessible. The platform's collaborative fraud intelligence network allows institutions to share insights and stay ahead of emerging threats without compromising sensitive data.

Cloud-native and versatile integration

ThreatMark's cloud-native fraud detection platform seamlessly integrates with legacy banking systems using versatile integration options (secure API or API-less), reducing implementation complexity.

5.3 Resistant banking industry

Cost of implementation

As banks navigate regulatory pressures and increasing competition from fintech firms, cost management has become a priority. At first glance, investing in advanced fraud protection may seem expensive—especially for institutions that have not yet felt the full impact of rising digital scams. However, fraud's impact extends beyond immediate losses. Banks must also account for its effect on customers—particularly affluent individuals and businesses, who are prime targets—along with the lasting damage to reputation, which can lead to long-term financial and competitive disadvantages.

Carefully weighing the costs and benefits of advanced anti-fraud solutions is essential. Once again, leveraging third-party vendors can be a cost-effective alternative, allowing banks to adopt cutting-edge fraud prevention technologies at a manageable cost. These solutions offer full-service support, continuous updates, and scalability, ensuring financial institutions stay ahead of evolving fraud tactics without excessive operational overhead.

Real cost of fraud

The loss of clients

According to an ACI Worldwide survey, one in four fraud victims globally—and over 30% of APP scam victims in the US—chose to leave their financial institution after experiencing fraud.[2]

The expense of acquiring new clients to replace those lost

Losing customers to fraud means that banks must invest heavily in customer acquisition.

Cost of reimbursements

With consumer protection laws increasingly requiring banks to cover losses from APP fraud, failing to invest in robust fraud prevention addressing rising threats could lead to significant financial liabilities.

The impact of a damaged reputation

Fraud damages a bank's reputation, making client retention and acquisition harder while requiring costly PR and marketing to rebuild trust.

AI skepticism

A key challenge in implementing AI-based fraud detection systems is overcoming mistrust toward the technology. McKinsey's research on generative AI adoption highlights this reluctance: while 65% of businesses integrated GenAI in 2024, only 8% of financial services leaders reported using it regularly—a figure unchanged from the previous year. Concerns over accuracy (63%), compliance risks (42%), and lack of explainability (40%) are key factors holding organizations back.

Despite ongoing caution, AI is steadily gaining industry trust, shifting from a perceived risk to a mainstream solution—particularly in fraud prevention. Today, approximately 70% of financial institutions leverage AI and machine learning to combat fraud. As the fraud landscape continues to evolve, institutions that remain skeptical risk falling behind, while early adopters gain a significant competitive edge.

Compliance concerns

The banking industry operates under strict regulations, making AI implementation in financial institutions a complex process that requires compliance. Privacy and transparency laws significantly impact AI-related decision-making, while new regulations—such as the EU's AI Act (**see Section 6.1**)—are further shaping the landscape. These evolving compliance requirements may make banks hesitant to adopt AI-based anti-fraud systems before clear regulatory guidelines are established, fearing they may need to adjust or overhaul their strategies later.

However, by engaging early in compliance discussions and integrating regulatory considerations into AI development, financial institutions can align with evolving regulations while maintaining innovation and fraud prevention effectiveness.

ThreatMark

ThreatMark's approach

ThreatMark's AI-driven fraud detection operates as a service tailored to the needs of diverse financial institutions—from Tier 1 banks seeking a solution that integrates seamlessly with their existing AI models to Tier 2 and 3 banks requiring a cost-efficient, ready-to-use fraud detection provider. As a cloud-native, scalable solution, the platform aligns with evolving regulatory requirements, ensuring banks' investments in fraud prevention remain future-proof.

6. The present and future of AI in fraud mitigation

The role of AI in banking and fraud mitigation is shaped by multiple factors, with regulatory frameworks playing a particularly critical role. Regulations influence not only the evolution of AI technologies but also their applicability within the banking sector. Aligning AI implementation with compliance requirements is thus one of many essential steps in adopting advanced fraud detection. For this reason, Section 6.3 provides a comprehensive guide to AI-driven fraud detection implementation.

6.1 Overview of the regulatory landscape concerning AI

EU's AI Act

Some regions have opted for a more cautious and regulated approach to AI development. The European Union's Artificial Intelligence Act (AI Act), adopted in August 2024, marks a major milestone in AI regulation, with implications for the financial services sector.

The AI Act adopts a risk-based approach to AI regulation, categorizing AI systems into four risk levels: unacceptable, high, limited, and minimal risk. Each category comes with specific regulatory obligations, ensuring that higher-risk applications are subject to stricter oversight and compliance requirements.

In financial services, high-risk AI systems—such as credit scoring and insurance risk assessment—are subject to stricter oversight due to potential bias or discrimination. However, as AI is already widely used in fraud detection, the AI Act is not expected to impose significant additional legal obligations in these areas. Supervisors will evaluate whether current frameworks are sufficient or if additional guidance is needed, factoring in proportionality, fairness, explainability, and accountability for specific use cases. In particular, bias in fraud detection models is becoming a growing concern from a Model Risk Management perspective. This raises challenges for both internal controls and external regulatory expectations.

AI regulations in the US

The US is taking a starkly different approach. On January 23, 2025, President Trump signed an executive order rolling back AI regulations to spur innovation, calling for a review of existing rules and a new AI action plan.

The current US administration is also scaling back AI oversight, with budget cuts hitting the National Institute of Standards and Technology (NIST) and its AI Safety Institute (AISI). Critics warn that dismantling AISI could undermine AI safety rather than boost US competitiveness.

Looser regulations may speed up AI advancements but amplify risks, from biased decisions to increased cybercrime. With fewer safeguards, fraudsters could more easily exploit AI tools —deepfakes, voice cloning, and more—fueling an already escalating fraud epidemic. This lack of oversight may also make it more difficult to enforce liability on financial institutions when AI-related fraud occurs.

Global response

As the US moves toward deregulation, its influence may set a precedent that redefines global AI governance, potentially shifting the balance between innovation and oversight.

Amid growing global divergence in AI regulation, UK ministers have delayed plans to introduce AI rules—a move likely aimed at supporting broader economic stimulus efforts. This decision affects a proposed bill that would have required AI companies to submit large-scale models, such as ChatGPT, for testing by the UK's AI Safety Institute.

Meanwhile, countries such as Canada, Japan, and Australia are advancing their own AI regulatory frameworks, many of which align more closely with the EU's emphasis on accountability, transparency, and ethical considerations rather than the pro-innovation, deregulated approach of the current US administration.

6.2 Other upcoming regulatory changes affecting fraud prevention

PSD3

The Payment Services Directive 3 (PSD3), a key EU regulation, aims to enhance payment market efficiency while tackling pressing issues like payment fraud—directly impacting fraud detection measures.

While PSD3 does not explicitly mention AI, its mandates—such as expanded liability for authorized fraud, enhanced transaction monitoring, and fraud data sharing—align with AI-driven detection systems. These solutions help financial institutions meet rising regulatory demands efficiently, strengthening both compliance and overall resilience against fraud.

How PSD3 proposal affects fraud detection*

Fraud data sharing

PSPs can share fraud-related data —including personal identifiers, transaction details, and fraud tactics—to enhance industry-wide fraud prevention, provided that proper security protocols are in place.

IBAN/name matching

Payment service providers (PSPs) must verify the recipient name and IBAN before processing credit transfers.

Authorized fraud compensation

PSPs must reimburse victims of impersonation fraud where fraudsters misuse the name, email, or phone number of any relevant public or private entity—such as a bank or the police—to deceive consumers.

Shared liability model

Liability for impersonation fraud may extend beyond PSPs to electronic communications service providers (ECSPs) and online platforms, requiring all parties to implement fraud prevention measures.

Strong customer authentication (SCA)

Independent factors from the same category are allowed, but dynamic codes remain mandatory. PSPs must support accessible SCA methods.

Transaction monitoring

PSPs must deploy advanced transaction monitoring systems to detect fraud and support SCA compliance. These systems should have the capability to monitor both outbound and inbound transactions, with the authority to block payments when fraudulent activity is strongly suspected.

* As of February 2025, PSD3 is still under review by the European Parliament and the Council of the European Union.

Regulatory shifts in US fraud liability

With the change in US leadership, efforts to increase bank liability for authorized fraud remain uncertain. The proposed Protecting Consumers from Payment Scams Act (Aug 2024) sought to hold banks accountable for reimbursements, but the Consumer Financial Protection Bureau (CFPB)—which would have enforced these rules—was effectively shut down in February 2025.

For banks, this removes the risk of mandatory reimbursements—at least for now, as consumer pressure persists—but regulatory instability still poses challenges. Policy shifts complicate long-term planning, and the CFPB's closure leaves key areas of consumer protection—such as financial literacy and fraud awareness—largely unsupervised.

6.3 Seven steps for banks to implement AI-driven fraud detection systems

Adopting AI-powered fraud detection is more than a technological shift—it is a strategic necessity for banks looking to combat increasingly sophisticated financial crime. To ensure a seamless and effective implementation, financial institutions should follow a structured approach.

1

Defining objectives and assessing the fraud risk landscape

Before implementing AI-driven fraud detection, banks must first gain a comprehensive understanding of their fraud risk landscape. A data-driven approach is crucial at this stage— the more insights a bank gathers on its fraud exposure, the better it can tailor its AI strategy to address specific pain points. This process involves assessing:

- The most prevalent fraud types affecting the institution (e.g., APP fraud, mule detection, synthetic identity fraud)
- Current fraud detection and prevention rates, along with a comparison to industry benchmarks
- Emerging fraud trends, particularly those leveraging AI to mimic legitimate behavior, such as deepfake-powered social engineering scams
- Existing vulnerabilities in fraud prevention systems and potential weaknesses as fraud tactics evolve

2

Setting clear goals for AI implementation

AI-driven fraud detection should be implemented with clear, measurable objectives that allow banks to assess its impact and optimize the system over time. Key goals typically include:

- Reducing false positives—minimizing unnecessary transaction declines while improving fraud detection rates
- Enhancing fraud detection accuracy—leveraging AI's ability to identify sophisticated fraud patterns in real time
- Lowering operational costs—automating detection and reducing reliance on manual fraud investigations
- Delivering risk-appropriate consumer friction—ensuring AI optimizes security without negatively impacting user experience

3

Ensuring data readiness and accessibility

The effectiveness of any AI fraud detection system hinges on data quality and accessibility, ensuring that AI systems operate transparently, securely, and effectively. Banks should:

- Identify and integrate relevant data sources—transaction history, behavioral analytics, device intelligence, network data, and other contextual information.
- Continuously monitor and remediate data quality issues. Inaccurate or incomplete data can lead to poor model performance.
- Break down data silos to create a unified fraud detection framework. AI performs best when it has a holistic view of customer activity across multiple touchpoints.
- Ensure compliance with data privacy regulations and regional banking laws. AI models must be trained using ethically sourced and legally compliant datasets.

4

Build vs. Buy: Choosing the right AI solution

A key decision for banks is whether to develop an in-house AI fraud prevention system or partner with a third-party provider. Before choosing, banks must assess their long-term strategy, operational capabilities, and budget, as both options come with advantages and challenges:

In-house development:

- Provides complete control over system design and data.
- Might appear cost-effective in the long term, but maintenance costs typically reach 30–40% of the initial build cost annually—effectively requiring reinvestment of the entire project cost every 2–3 years.
- Requires deep domain expertise in fraud detection, AI/ML, and real-time system architecture—areas that are complex and require constant adaptation to emerging threats.

Third-party AI fraud detection providers:

- Offer ready-to-deploy, continuously updated solutions that adapt to new fraud tactics.
- Include maintenance, regulatory compliance, and expertise in fraud-specific AI modeling.
- Can be more cost-effective than building in-house, especially for banks lacking specialized AI and ML teams.
- Provide less control, as banks must adapt to the vendor's predefined approach and capabilities.

5

Prioritizing explainability and regulatory compliance

AI-powered fraud detection must be transparent, auditable, and compliant with evolving regulations. Explainability is more than a compliance requirement—it enables fraud teams to interpret and refine AI outputs for continuous improvement. To meet regulatory expectations, banks should:

- Ensure AI decisions are explainable to fraud analysts, regulators, and compliance teams. Black-box AI models can raise trust and accountability concerns.
- Align with key regulatory frameworks, including Basel regulations, European Banking Authority (EBA) guidelines, PSD3, and local financial authorities.
- Implement governance frameworks to mitigate risks associated with AI bias, model drift, and ethical concerns.

6

Monitoring performance and adapting to emerging threats

AI-driven fraud detection is not a "set-it-and-forget-it" solution. Banks must continuously monitor performance and refine models to keep pace with evolving fraud tactics. This requires:

- Establishing key performance indicators (KPIs) to measure AI's effectiveness, such as fraud detection accuracy, false positive rates, and operational efficiency.
- Leveraging real-time monitoring dashboards to detect anomalies and rapidly respond to threats.
- Updating AI models proactively to counteract AI-powered fraud tactics, such as adversarial machine learning attacks used by criminals to evade detection.

7

Fostering cross-bank and third-party collaboration and intelligence sharing

Collaboration and fraud data sharing are essential in fighting financial crime and are increasingly backed by regulations like PSD3. Since a threat to one bank can quickly become an industry-wide risk, financial institutions must adopt proactive strategies beyond their own operations. To enhance defenses, banks should participate in collaborative intelligence-sharing initiatives, including:

- Participation in fraud intelligence networks to exchange AI-driven insights and threat patterns.
- Partnerships with law enforcement, industry bodies, and peer banks to combat large-scale fraud schemes.
- Supporting open banking initiatives while ensuring secure API connections to prevent third-party exploitation.
- Collaboration with third-party anti-fraud solution providers to leverage external expertise.

ThreatMark

ThreatMark's approach

At ThreatMark, we are dedicated to supporting banks in their fight against fraud—not only through advanced fraud detection technologies but also by fostering collaboration. Our Banking Fraud Summits serve as a platform where financial institutions can exchange insights, strategies, and best practices to stay ahead of emerging threats.

ThreatMark is pioneering a new approach to collaboration and intelligence that enables banks to proactively manage fraud risk and stay ahead of the threats inherent to the evolving digital economy. This is powered by a community-centric approach to securely, compliantly, and efficiently share data, generating value through reciprocity and shared insights.

As fraud tactics evolve, so must fraud prevention. That's why we prioritize research and innovation, keeping our Behavioral Intelligence Platform at the forefront of fraud detection to identify even the most sophisticated attacks.

But stopping fraud isn't enough—we aim to disrupt it at its source. By mapping entire fraud networks, from attack infrastructure to money laundering, we help banks achieve lasting fraud loss reduction and build long-term resilience against financial crime.

7. Conclusion

As examined in this whitepaper, AI is transforming the way fraud is both perpetrated and prevented—equipping criminals and fraud fighters with new capabilities. However, it is crucial to remember that AI is merely a tool. While it enhances certain fraud tactics—deepfake-powered impersonation being a prime example—fraud's fundamental purpose remains the same: deceiving customers and stealing their money. Old scams, wrapped in new and more deceptive disguises.

The greater challenge lies in AI's ability to scale fraud at an unprecedented pace, lowering costs and reducing the expertise needed to execute attacks. As a result, fraud is becoming more sophisticated and scalable, making attacks harder to detect and enabling a higher volume of targeted attempts.

Encouragingly, AI is proving to be just as powerful on the defense. More banks are adopting AI-driven solutions, such as behavioral intelligence, to detect and prevent fraud before it happens. AI is also transforming other areas, enhancing customer services and optimizing operations. For customers, this could mean a future of more secure, personalized, and seamless banking experiences.

Regulators, too, are working to keep pace, and their approach will continue to shape the future of AI adoption. For now, regulatory stances remain divided—Europe favoring a stricter, more controlled framework, while the US takes a more open-ended approach. How these differing tactics evolve will define AI's role in financial security in the years ahead.

Regardless of regulatory differences, one fundamental principle remains unchanged: the digital financial ecosystem must be safe, transparent, and fair. The same standard should apply to AI models.

At ThreatMark, we are committed to developing AI-powered fraud detection as a formidable tool in the fight against fraud.

However, we also recognize that fraud is a complex, industry-wide challenge that requires more than cutting-edge technology—it demands collaboration and shared intelligence. As fraud fighters—whether in banks, credit unions, or anti-fraud service providers—we must never lose sight of our ultimate goal: working together to disrupt fraudsters and protect people from the devastating harm fraud inflicts.

Glossary

AI (Artificial Intelligence) – Technology that enables machines to perform tasks that typically require human intelligence, such as decision-making, pattern recognition, and automation.

APP fraud (Authorized Push Payment Fraud) – A type of fraud where individuals are deceived into authorizing the transfer of funds to a scammer's account.

Behavioral intelligence – The process of gathering, analyzing, and interpreting behavioral data to identify patterns and detect anomalies, particularly in fraud prevention.

CCPA (California Consumer Privacy Act) – A US state law granting California residents rights over their personal data, including access, deletion, and opting out of its sale.

CPRA (California Privacy Rights Act) – An update to the CCPA that expands consumer privacy rights, strengthens data protection rules, and establishes the California Privacy Protection Agency.

DL (Deep Learning) – A subset of machine learning that uses neural networks to analyze complex patterns in large datasets.

ECSP (Electronic Communications Service Provider) – An entity that provides digital communication services, such as internet, messaging, or voice services.

Frankenstein ID – A synthetic identity created by combining real and fake personal data, often used in fraud schemes.

GANs (Generative Adversarial Networks) – A type of AI model where two networks compete to generate realistic synthetic data.

GDPR (General Data Protection Regulation) – An EU law governing data privacy and security, regulating how personal data is collected and processed.

GenAI (Generative AI) – AI that creates new content, including text, images, and videos, based on training data.

GPAI (General-Purpose AI) – AI capable of performing multiple tasks across different domains without requiring task-specific programming.

LLMs (Large Language Models) – AI models trained on large text datasets to generate human-like responses and process language.

ML (Machine Learning) – A subset of AI that allows systems to learn from data and improve predictions without explicit programming.

NLP (Natural Language Processing) – A field of AI that enables machines to understand, interpret, and generate human language.

PSD2 (Revised Payment Services Directive) – An EU regulation that enhances payment security, promotes open banking, and mandates strong customer authentication.

PSD3 (Third Payment Services Directive) – A forthcoming EU regulation expected to strengthen security, competition, and fraud prevention measures.

PSP (Payment Service Provider) – An entity that enables the processing of electronic payments, including card transactions, bank transfers, and digital wallets.

References

2024 Data Breach Investigations report. Verizon. https://www.verizon.com/business/resources/reports/2024-dbir-data-breach-investigations-report.pdf

Abrams, L. (2025, January 30). Time Bandit ChatGPT jailbreak bypasses safeguards on sensitive topics. Bleeping Computer. https://www.bleepingcomputer.com/news/security/time-bandit-chatgpt-jailbreak-bypasses-safeguards-on-sensitive-topics/

ACI Worldwide. (2024, October 1). Scamscope report: APP scam trends | ACI Worldwide. https://www.aciworldwide.com/scamscope-report-app-scam-trends

ACI Worldwide ScamScope Projects APP scam losses to hit $7.6 billion by 2028 | ACI Worldwide. (n.d.). ACI Worldwide. https://investor.aciworldwide.com/news-releases/news-release-details/aci-worldwide-scamscope-projects-app-scam-losses-hit-76-billion

AI Act. (2025, March 20). European Commision. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

AI Act and its impacts on the European financial sector. European Insurance and Occupational Pensions Authority. https://www.eiopa.europa.eu/publications/ai-act-and-its-impacts-european-financial-sector_en

AI voice cloning scams could catch millions out - with over a quarter of UK adults targeted in the past year. (2024, September 18). Starling Bank. https://www.starlingbank.com/news/starling-bank-launches-safe-phrases-campaign/

AI Watch: Global regulatory tracker - Australia. (2024, December 16). White & Case LLP International Law Firm, Global Law Practice. https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-australia

AI Watch: Global regulatory tracker - Japan. (2024, July 1). White & Case LLP International Law Firm, Global Law Practice. https://www.whitecase.com/insight-our-thinking/ai-watch-global-regulatory-tracker-japan

Akselrod, O., & Venzke, C. (2025, February 11). Trump's efforts to dismantle AI protections, explained | ACLU. American Civil Liberties Union. https://www.aclu.org/news/privacy-technology/trumps-efforts-to-dismantle-ai-protections-explained

Artificial Intelligence and Data Act. (2023, September 27). Government of Canada. https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act

Bousquette, I. (2024, April 3). Deepfakes are coming for the financial sector. Wall Street Journal. https://www.wsj.com/articles/deepfakes-are-coming-for-the-financial-sector-0c72d1e5

Bunn, A. (2024, September 22). Artificial Imposters—Cybercriminals turn to AI voice cloning for a new breed of scam. McAfee Blog. https://www.mcafee.com/blogs/privacy-identity-protection/artificial-imposters-cybercriminals-turn-to-ai-voice-cloning-for-a-new-breed-of-scam/

Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L., & Zemmel, R. (2023, June 14). The economic potential of generative AI: The next productivity frontier. McKinsey & Company. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier#introduction

Corporate governance principles for banks. (2015, July 8). BIS. https://www.bis.org/bcbs/publ/d328.htm

Courea, E. (2025, February 24). UK delays plans to regulate AI as ministers seek to align with Trump administration. The Guardian. https://www.theguardian.com/technology/2025/feb/24/uk-delays-plans-to-regulate-ai-as-ministers-seek-to-align-with-trump-administration

Crypto Scam Revenue 2024: Pig butchering grows Nearly 40% YOY as fraud industry leverages AI and increases in sophistication. (2025, February 13). Chainalysis. https://www.chainalysis.com/blog/2024-pig-butchering-scam-revenue-grows-yoy/

Cyber Security in the age of offensive AI (2024, July 1). Netacea. https://netacea.com/reports/cyber-security-in-the-age-of-offensive-ai/

Dangelo, M. (2023, December 6). Needed AI skills facing unknown regulations and advancements - Thomson Reuters Institute. https://www.thomsonreuters.com/en-us/posts/technology/needed-ai-skills/

Devising and detecting phishing emails using large language models. (2024). IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/document/10466545

EBA opinion on new types of payment fraud and possible mitigants. (2024). In European Banking Authority. https://www.eba.europa.eu/sites/default/files/2024-04/363649ff-27b4-4210-95a6-0a87c9e21272/Opinion%20on%20new%20types%20of%20payment%20fraud%20and%20possible%20mitigations.pdf

Fact Sheet: President Donald J. Trump takes action to enhance America's AI leadership. (2025, January 24). The White House. https://www.whitehouse.gov/fact-sheets/2025/01/fact-sheet-president-donald-j-trump-takes-action-to-enhance-americas-ai-leadership/

Financial firms hated US consumer watchdog, but rapid unraveling creates limbo. (2025, February 13). Reuters. https://www.reuters.com/world/us/financial-firms-hated-us-consumer-watchdog-rapid-unraveling-creates-limbo-2025-02-13/

Finding value in generative AI for financial services. (2023). MIT Technology Review. https://www.technologyreview.com/2023/11/26/1083841/finding-value-in-generative-ai-for-financial-services/

Four steps to get ahead of AI-Enhanced Cyberattacks, IT pros top worry in 2025. (2024, September 23). GetApp.com. https://www.getapp.com/resources/ai-enhanced-cyberattacks-top-worry-it-professionals/

Generative AI is expected to magnify the risk of deepfakes and other fraud in banking. (2024, December 10). Deloitte Insights. https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2024/deepfake-banking-fraud-risk-on-the-rise.html

Giuffrida, A. (2025, February 10). AI phone scam targets Italian business leaders including Giorgio Armani. The Guardian. https://www.theguardian.com/world/2025/feb/10/ai-phone-scam-targets-italian-business-leaders-including-giorgio-armani

Gollom, R. M. a. M. (2014, March 4). How 'synthetic' identity fraud costs Canada $1B a year. CBC. https://www.cbc.ca/news/canada/how-synthetic-identity-fraud-costs-canada-1b-a-year-1.2554429

Goodwin, L. (2024, December 20). "AI deepfake romance scam duped me out of £17k." https://www.bbc.com/news/articles/cdr0g1em52go

Gozzi, L. (2025, January 15). French woman duped by AI Brad Pitt faces mockery online. BBC. https://www.bbc.com/news/articles/ckgnz8rw1xgo

Granda, C. (2024, October 18). Fraudsters use voice-cloning AI to scam man out of $25K. ABC7 Chicago. https://abc7chicago.com/post/scammers-use-voice-cloning-artificial-intelligence-ai-swindle-man-25k-los-angeles-police-talk-how-avoid/15441538/

Heiding, F., Schneier, B., & Vishwanath, A. (2024, May 30). AI Will Increase the Quantity — and Quality — of Phishing Scams. Harvard Business Review. https://hbr.org/2024/05/ai-will-increase-the-quantity-and-quality-of-phishing-scams

Jacobson, N. (2024, February 26). Deepfakes and their impact on society. CPI OpenFox. https://www.openfox.com/deepfakes-and-their-impact-on-society/

Katz, L. (2024, November 15). Introducing Daisy, the 'AI Granny' outwitting scammers. Forbes. https://www.forbes.com/sites/lesliekatz/2024/11/15/introducing-daisy-an-ai-granny-outwitting-scammers-one-call-at-a-time/

Kerr, D., & Bhuiyan, J. (2025, February 22). Crypto and big tech's backing pays off as Trump makes tech-friendly moves. The Guardian. https://www.theguardian.com/technology/2025/feb/22/crypto-big-tech-trump

Lalchand, S., Srinivas, V., & Gregorie, J. (2023, November 23). Using biometrics to fight back against rising synthetic identity fraud. Deloitte Insights. https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2023/financial-institutions-synthetic-identity-fraud.html

Lemos, R. (2022, October 10). Emotet rises again with more sophistication, evasion. Dark Reading. https://www.darkreading.com/threat-intelligence/emotet-rises-again-with-more-sophistication-evasion

Mai, K. T., Bray, S., Davies, T., & Griffin, L. D. (2023). Warning: Humans cannot reliably detect speech deepfakes. PLoS ONE, 18(8), e0285333. https://doi.org/10.1371/journal.pone.0285333

Merino, R. (2024, December 3). National Police smash gang that used artificial intelligence to scam businesses in Spain. Sur in English. https://www.surinenglish.com/malaga/artificial-intelligence-organisation-used-launder-money-from-20241203075640-nt.html

Mitrovic, S. (2024, August 9). Gmail Account takeover: Super realistic AI scam call. SamMitrovic.com. https://sammitrovic.com/infosec/gmail-account-takeover-super-realistic-ai-scam-call/

References

Nelson, N. (2024, October 10). AI-Powered cybercrime cartels on the rise in Asia. Dark Reading. https://www.darkreading.com/threat-intelligence/ai-powered-cybercrime-cartels-asia

O'Regan, E. (2024, May 30). Over 40% of fraud attacks on European financial institutions now use AI. The Brussels Times. https://www.brusselstimes.com/1068616/over-40-of-fraud-attacks-on-european-financial-institutions-now-use-ai

Propson, D., & Parker, D. (2025). Artificial intelligence in financial services. In World Economic Forum & Accenture. https://reports.weforum.org/docs/WEF_Artificial_Intelligence_in_Financial_Services_2025.pdf

PYMNTS. (2024). The State of Fraud and Financial Crime in the U.S. 2024: What FIs Need to Know. https://www.pymnts.com/wp-content/uploads/2024/11/PYMNTS-State-of-Fraud-and-Financial-Crimes-November-2024.pdf

PYMNTS Intelligence. (2024). Protecting accelerated disbursements from fraud. In Money Mobility Tracker. https://www.pymnts.com/wp-content/uploads/2024/04/PYMNTS-Money-Mobility-Tracker-April-2024.pdf

Rocchi, S. (2024, July 31). Truffe online e frodi informatiche: in 6 mesi sottratti 114 milioni. RaiNews. https://www.rainews.it/articoli/2024/07/truffe-online-e-frodi-informatiche-in-6-mesi-sottratti-114-milioni--0fb03387-b468-4726-bc6e-6e9de729779e.html

Security challenges rise as QR code and AI-Generated Phishing proliferate | Recorded future. (2024, July 18). https://www.recordedfuture.com/research/qr-code-and-ai-generated-phishing-proliferate

Seven in 10 financial institutions use AI and ML to combat fraud. (2024, March 26). PYMNTS.com. https://www.pymnts.com/news/security-and-risk/2024/seven-in-10-financial-institutions-use-ai-and-ml-to-combat-fraud/

Shimony, E., & Tsarfati. (2023, January 17). Chatting our way into creating a polymorphic malware. CyberArk. https://www.cyberark.com/resources/threat-research-blog/chatting-our-way-into-creating-a-polymorphic-malware

Singla, A., Sukharevsky, A., Yee, L., Chui, M., & Hall, B. (2025, March 12). The state of AI: How organizations are rewiring to capture value. McKinsey & Company. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

Special topic – Artificial intelligence | European Banking Authority. https://www.eba.europa.eu/publications-and-media/publications/special-topic-artificial-intelligence

Synthetic identity fraud. KPMG. https://kpmg.com/us/en/articles/2022/synthetic-identity-fraud.html

The authorities are taking action to combat the massive phenomenon of financial scams catching out an increasing number of individuals. (2024, December 26). AMF. https://www.amf-france.org/en/news-publications/news-releases/authorities-are-taking-action-combat-massive-phenomenon-financial-scams-catching-out-increasing

The Gartner IT Security Approach for the Digital Age. (2017, June 12). Gartner.com. https://www.gartner.com/smarterwithgartner/the-gartner-it-security-approach-for-the-digital-age

The malware-as-a-service Emotet. (2021). ANSSI. https://www.cert.ssi.gouv.fr/uploads/CERTFR-2021-CTI-003.pdf

The near-term impact of AI on the cyber threat. (2024). National Cyber Security Centre. https://www.ncsc.gov.uk/report/impact-of-ai-on-cyber-threat

The state of deepfakes 2024. (n.d.). In Sensity. https://sensity.ai/reports/

Tietoevry. (2024, September 18). Fraud attempts reached an "All-Time High" this summer. https://www.tietoevry.com/en/newsroom/all-news-and-releases/press-releases/2024/09/new-survey-reveals-18-of-norwegians-and-15-of-swedes-fell-victim-to-financial-fraud-last-year--fraud-attempts-reache/

Vartbabedian, M. (2024, December 17). AI can take the slog out of compliance work, but executives not ready to fully trust it. Wall Street Journal. https://www.wsj.com/articles/ai-can-take-the-slog-out-of-compliance-work-but-executives-not-ready-to-fully-trust-it-7cd60a16

Wall Street banks are poaching rival AI talents. (2023, November 28). Bloomberg. https://www.bloomberg.com/news/articles/2023-11-28/goldman-raided-by-recruiters-in-wall-street-fight-for-ai-talent

What is Deepfake fraud in accounts payable, and how can you prevent it? (2024, December 9). Medius. https://www.medius.com/blog/what-is-deepfake-fraud-in-accounts-payable-and-how-can-you-prevent-it/

Williams, S. (2024, October 10). Deepfake scams cost Australian businesses millions, research reveals. CFOtech Australia. https://cfotech.com.au/story/deepfake-scams-cost-australian-businesses-millions-research-reveals

World's most dangerous malware EMOTET disrupted through global action | Europol. https://www.europol.europa.eu/media-press/newsroom/news/world%E2%80%99s-most-dangerous-malware-emotet-disrupted-through-global-action