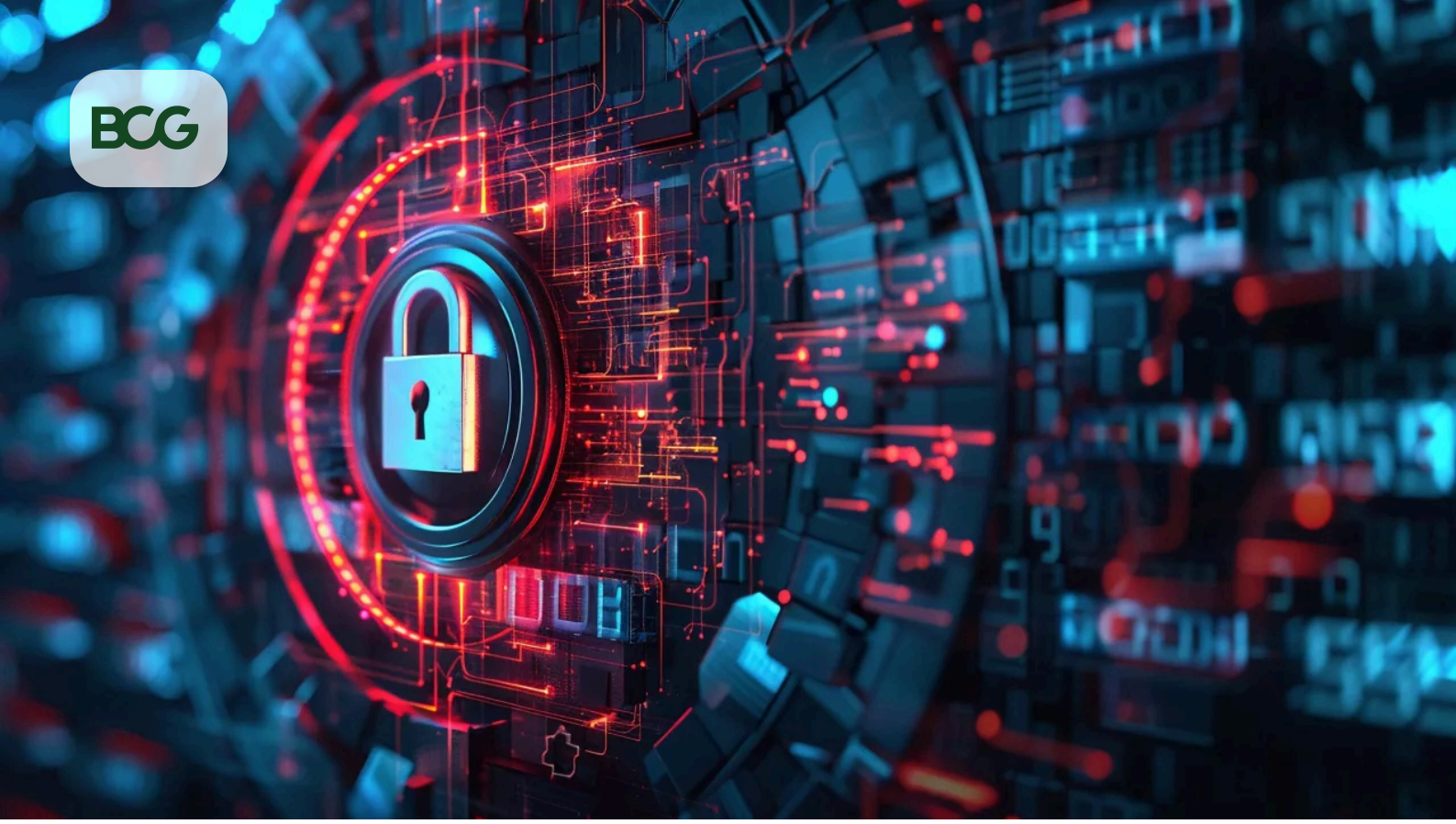


RISK MANAGEMENT AND COMPLIANCE

Agentic AI Will Industrialize Financial Scams. Are Banks Ready?

By Edzard Wesselink, Anjali Narendra, Stiene Riemer, Brian O'Malley, Matthew Barton, and Hanjo Seibert

ARTICLE JUNE 11, 2026 8 MIN READ

The ability of generative AI to convincingly create synthetic identities, cloned voices, fake documents, and realistic digital interactions has already led to a rise in fraud and scam cases. Over the next few years, GenAI tools that can create live deepfake videos will become highly accessible, driving a step change in how convincing these scams appear to victims. But the real

transformation will happen once agentic systems become capable of running scams and fraud end to end without human intervention.

At that point, BCG estimates the cost of running scams and fraud could fall by 90% or more, enabling adversaries to launch a much larger volume of attacks, test a broader range of tactics, and more rapidly adapt against new defenses. As a result, we could see a twofold—or more—increase in successful scam and fraud activity, with profound emotional and financial costs to those who are targeted.

For banks, this signals the arrival of a more persistent, "always-on" threat environment that requires a different defense model. The good news is that AI is also a powerful tool for preventing, detecting, and responding rapidly to scams and fraud. Banks can take action now to build adaptive, AI-enabled defenses, redesign operating models for scale, strengthen ecosystem coordination, and prepare for surge events.

> "The good news is that AI is a powerful tool for preventing, detecting, and responding rapidly to scams and fraud.

How Agentic AI Could Industrialize Financial Crimes

Financial scams and fraud are already a major problem today, with scams alone costing consumers and businesses around $442 billion annually, according to the Global Anti-Scam Alliance. The range of fraud and scam types is wide, with the cost to victims extending well beyond their bank account. (See the sidebar "The Broad Reach of Scams and Fraud.")

The Broad Reach of Scams and Fraud

Financial scams (which involve customers being deceived into voluntarily authorizing payments themselves) and fraud (in which deception is used to gain access to the victim's financial accounts) encompass a broad range of criminal activity, including:

- **Account Fraud.** Fraudsters steal account details of real customers through deceptive means like phishing links to fake banking websites, or fake customer service lines.
- **Application Fraud.** Fraudsters create synthetic IDs and fake documentation to apply for credit products, withdraw funds, and let the account go into arrears.
- **Invoice Fraud.** Fraudsters send fake invoices or hijack invoices and override legitimate banking details to redirect payments to fraudulent accounts.
- **Relationship Scams.** Scammers use fake personas to connect with victims via social media or other means, build a relationship, then convince the victim to transfer money to them. Includes romance scams run through dating apps.
- **Investment Scams.** Scammers pose as legitimate businesses and promise big returns on an investment. Sometimes they even return small amounts of money to prove the value of the investment, ultimately making off with the victim's funds.
- **Business Email Compromise Scams.** Targeting businesses, scammers pose as a supplier the company has done business with. Based on a fake invoice, they trick the business into paying them for services never rendered.
- **Money Recovery Scams.** Scammers pose as attorneys or other organizations that can help victims recover money lost via an earlier scam. They collect a fee—and, of course, return nothing.
- **Fake E-Commerce Scams.** Scammers create fake e-commerce websites that are clones of existing websites and let customers make orders but never send out the product.

In addition to the financial losses suffered by victims, research has found that these scams take a personal toll and can contribute to depression, anxiety, trauma, and isolation.

A Growing Risk. These financial crimes are not just devastating to the direct victims, but they also pose a real threat to banks and the broader financial system. Banks are often financially liable for fraud losses and, while they have not typically been required to compensate scam victims, jurisdictions such as Australia and the UK have moved to expand expectations and responsibility in relation to scams. Even where banks do not carry direct scam liability, they still face meaningful operational costs, as victims, and in some cases the scammers themselves, contact them to demand a response, seek reimbursement, or challenge payments.

Scams can also create knock-on effects, including disputes, complaints, credit-reporting issues, debt collection activity, and wider customer-servicing costs. That is why increased scam volumes could lead not only to higher losses and greater regulatory pressure, but also to a broader erosion of customer trust and confidence in the financial system.

A Looming Step Change. Certainly, the banking industry has had experience with technology-driven surges in criminal activity. For example, in some markets, the advent of online banking led to an explosion of fraud and scam activity. To understand how agentic AI could push these crimes to the next level, consider the impact on just one type of lucrative illegal activity: romance scams.

Imagine a new open-source model is released that can effectively plan and deliver a long sequence of actions over a multiday period. A criminal syndicate retrains the model to remove safeguards, teaches it to effectively run romance scams end to end, and mass-deploys thousands of agents on a bot farm with smuggled compute hardware.

The agents work together to scrape public social media profiles and identify potential victims, perform searches to build personal profiles of their hobbies and likes, and create fake accounts with tailored characteristics. Each agent experiments with different tactics to build trust and push victims to transfer money, and the most successful strategies are propagated between agents. Each transfer request is structured differently, so the agent swarm can determine what types of payments get caught by the bank's detection engine and can maximize the rate of successful transfers.

Now picture this threat is multiplied across all different scam types: armies of agents flooding search results with fake e-commerce websites, spamming video-sharing websites with crypto investment scams, and stealing credentials through clones of login pages.

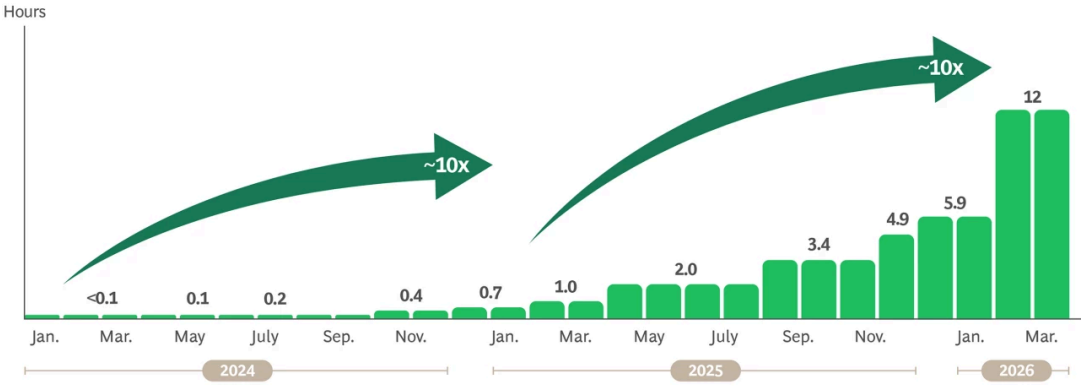
While current models cannot do this quite yet, the shift could come quickly. Agentic capabilities have improved tenfold per year since 2024. (See Exhibit 1.) If this trend continues, in the next one to two years models will have the capability to run day- or week-long scams. At that point, the ability of a model to perform multiple tasks along the end-to-end spectrum means that single individuals and small teams could generate sophisticated scams. This would democratize access

to capabilities needed to run scams and fraud and potentially draw in thousands of new scammers.

EXHIBIT 1

AI Models Can Perform Increasingly Long, Multistep Actions Independently

Highest monthly score on METR 1.1 benchmark, which measures the duration of software development tasks (in human equivalent hours) that AI can complete with a 50% success rate



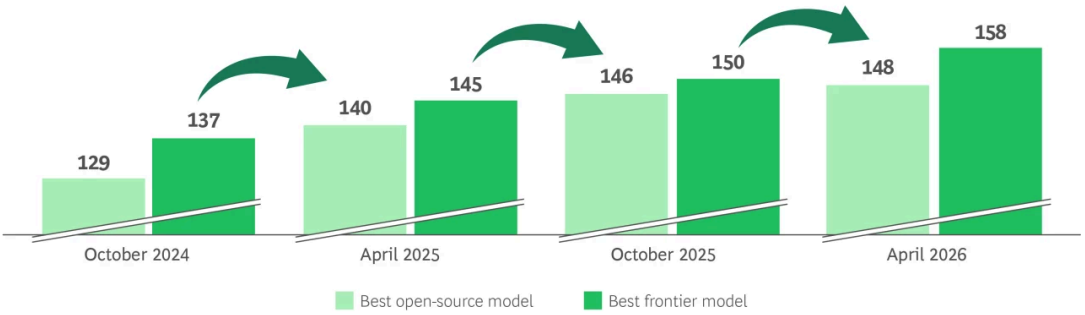
Source: METR, Time Horizon 1.1 benchmark.

Right now, of course, frontier models usually have built-in sophisticated safeguards to prevent misuse by criminals. However, open-source models typically catch up with leading capabilities within six to 12 months. (See Exhibit 2.)

EXHIBIT 2

New Capabilities in Frontier Models Typically Appear in Uncensored Open-Source Models in 6–12 Months

Highest indexed composite capability score of best AI model



Source: Epoch AI, Epoch Capabilities Index.
Note: Scores of the best open-source and the best frontier model on the first day of that month.

Given that it is easy to remove safeguards from open-source models, it is typically possible to download "uncensored" versions of new open-source models within days of their release that will permit any type of request. Scammers will likely get access to agentic models that can carry out

end-to-end scams within the next couple years, which gives banks a relatively narrow window to get ready.

How Banks Can Meet the Rising Scam and Fraud Threat

Acting early, before the operating model comes under strain, is far more effective than reacting once pressure builds. This makes it critical for banks to move now to build models that not only detect suspicious activity, but also freeze transactions, help customers exit dangerous situations quickly, and preserve trust when incidents occur.

Banks can take five actions to prepare.

Increase threat monitoring. As agentic AI developments change the threat landscape at a faster pace, banks need to increase their threat monitoring so they can track new trends and developments in a much more granular way. This allows banks to rapidly adapt defenses as the threat changes. This includes:

- Gathering richer behavioral and transactional data to train and continuously retrain detection and anomaly models
- Tracking capabilities of open-source, uncensored AI models.
- Building adversarial agents to red-team a bank's own detection models.
- Continuously analyzing social media along with customer interactions, calls, and complaints to detect new scammer tactics.
- Creating fake customers as scam targets to extract scammer information.

Leverage the head start afforded by frontier models. Banks should use the lead that frontier models have over open-source models to continuously improve scam and fraud prevention. Key steps they can take include:

- Building a strong internal AI team with early access to the latest models.
- Developing rapid approval pathways to get new capabilities into production quickly.
- Working with vendors that embed frontier capabilities into customer protection tools.

Build a responsive, scalable operating model. Banks will need an operating model that can effectively scale up to deal with volume spikes and respond within hours, not weeks. This includes steps like:

- Dynamically increasing controls on high-risk payments.
- Real-time analysis of calls to detect new potential scam and fraud patterns and use those to update detection engines.
- Automating or agentifying intervention, investigation, and recovery steps so they become scalable.
- Creating rapid-response governance to allow for daily model refinements.

Strengthen collaboration. As fraud and scam networks become more connected and more adaptive, financial institutions will need to work more closely with payment providers, platforms, telcos, social media companies, regulators, and law enforcement to identify and disrupt activity earlier. This includes:

- Strengthening and speeding up information sharing on scam and fraud patterns across the ecosystem.
- Accelerating the pace at which freeze and recovery requests are actioned.
- Coordinating responses across the ecosystem to break rings.

Design "fire breaks." Banks should prepare for periods when fraud and scam activity spikes and normal operating processes come under strain. In those moments, temporary measures may be needed to preserve control while the bank responds. For example:

- Developing surge playbooks with clear thresholds and escalation paths.
- Identifying customer journeys where temporary friction would have the greatest impact.
- Building controls that can be turned on flexibly, like step-up verification or transaction holds.
- Establishing clear rules for when certain types of transactions or account openings could be restricted.

The shift to agentic scams and fraud will likely be sudden. As attacks become persistent, adaptive, and highly personalized, today's defense models will come under increasing strain. Banks must move now to redesign how they prevent, detect, and respond to this threat. Those that do will be better positioned to protect customers and limit losses. Those that wait risk falling behind an adversary that is learning faster than they are.

Authors

Edzard Wesselink

Managing Director & Partner
Melbourne

Anjali Narendra

Project Leader
Sydney

Stiene Riemer

Managing Director & Partner
Munich

Brian O'Malley

Managing Director & Partner
Minneapolis

Matthew Barton

Managing Director and Partner
Philadelphia

Hanjo Seibert

Managing Director & Partner
Düsseldorf

ABOUT BOSTON CONSULTING GROUP

Boston Consulting Group partners with leaders in business and society to tackle their most important challenges and capture their greatest opportunities. BCG was the pioneer in business strategy when it was founded in 1963. Today, we work closely with clients to embrace a transformational approach aimed at benefiting all stakeholders—empowering organizations to grow, build sustainable competitive advantage, and drive positive societal impact.

Our diverse, global teams bring deep industry and functional expertise and a range of perspectives that question the status quo and spark change. BCG delivers solutions through leading-edge management consulting, technology and design, and corporate and digital ventures. We work in a uniquely collaborative model across the firm and throughout all levels of the client organization, fueled by the goal of helping our clients thrive and enabling them to make the world a better place.