moveris™ | epcor

WHITE PAPER

Agentic AI and the Liveness Detection Gap

How autonomous AI agents overcome identity verification systems, and what to do about it

Moveris · in collaboration with EPCOR

FRAUD EVENTS RISING

67%

of financial institutions and fintechs reported an increase in fraud events

AI-ENABLED CRIME

91%

of financial institutions reported more financial crimes committed with AI

PAYMENTS FRAUD

79%

of organizations saw attempted or actual payments fraud in 2024

PUBLISHED AUGUST 2026

SECTION 01

The verification assumption is broken

For more than a decade, digital identity verification has rested on the deceptively simple premise that if a system can confirm that a human being is physically present, that human can be trusted, or at least held accountable. The industry built an entire strategic ecosystem around this premise, including secure account creation processes, third-party bot detection tools, CAPTCHA checks, two-factor authentication, and more. The underlying assumption that these tasks were easy for a person and hard for a machine is now obsolete.

DOCUMENTED CASE · AUTONOMOUS CHECKOUT, $119 ORDER

The team at Moveris recently conducted an experiment and documented a case that captures the new reality with uncomfortable precision. With the cooperation and consent of an online retail partner, the team at Moveris created an AI agent to operate autonomously on behalf of a user. Without any subsequent human intervention, the agent purchased a $119 product through a checkout flow that included browsing the retailer website, adding the product to a cart, creating a purchaser account, signing in, completing two-factor authentication, filling out shipping and payment forms, placing the order, and overcoming third-party bot detection tools, all without a human in the transaction loop. The customer service representative who handled the subsequent inquiry was, herself, a human. The “customer” who made the purchase was not.

This is not a theoretical vulnerability. It is a documented operational reality, and its implications extend far beyond e-commerce. The same class of agents that can navigate a product checkout can also navigate a loan application, a wire transfer authorization, a healthcare onboarding flow, or an account recovery process. Anywhere that identity verification checks stand between an AI and a desired outcome, the check is now a puzzle to be solved rather than a barrier to be respected.

SECTION 02

Defining the agentic threat

The term “AI agent” has been used loosely enough to obscure the genuine novelty of the current moment. We define an AI agent as an autonomous software system that can perceive its environment through one or more input modalities (visual, textual, auditory), reason about that environment using a large language or vision-language model, form multi-step plans to achieve a specified goal, and execute those plans by calling external tools, APIs, or interfaces, including standard web browsers and application interfaces, without requiring human confirmation at each step.

This is categorically different from a scripted bot. Earlier-generation fraud automation relied on hardcoded decision trees and predictable interaction patterns. Modern agentic systems are adaptive, context-sensitive, and capable of navigating novel interface configurations they have never encountered before. They read instruction text, respond to dynamic prompts, and adjust their behavior when initial approaches fail.

The scale of deployment

Agentic AI is not a fringe capability. As of 2025, leading technology platforms (and bad actors) have released production-grade agentic frameworks that allow non-technical users to deploy goal-directed AI across web interfaces. These agents are being used legitimately for productivity automation, research assistance, and transactional workflows. The same infrastructure, however, is also accessible to actors with fraudulent intent.

Research published in 2025 demonstrates that vision-language-model agents can navigate e-commerce platforms, evaluate products against specified criteria, and complete purchases, thereby exhibiting behavior that is, in many respects, more consistent and goal-directed than human shoppers.[1] The fraud-relevant implication is not that AI agents shop differently than humans, but that they shop at all, and that the verification systems standing between agents and high-value transactions were not designed to stop them.

SECTION 03

Why this matters now

The timing of this challenge is particularly consequential. Fraud-as-a-service ecosystems have made agentic attack tooling accessible to a broad range of actors. Simultaneously, the regulatory environment has increased pressure on financial institutions to demonstrate robust KYC and AML compliance. Organizations caught between sophisticated AI attackers and heightened regulatory scrutiny face a verification gap that conventional liveness detection cannot close.

AI compresses the distance between capability and misuse. A technique that once required specialized malware, human fraud rings, or custom automation can now be performed by a general-purpose AI agent given a goal, a browser, and access to basic tools. The cost of attempting sophisticated fraud has fallen dramatically, while the scale at which those attempts can be repeated has risen sharply. Once these agents can reliably navigate verification workflows, fraudsters no longer need to defeat every control manually. They need only create an agent that can learn the workflow and repeat the attack.

This matters not only for fraud teams, but also for product, compliance, and operations leaders. The same user experience improvements that reduce friction for legitimate customers can reduce friction for autonomous adversaries. The risk is not that AI agents behave exactly like humans. The risk is that our systems continue to treat successful navigation of human-oriented tasks as evidence of human presence.

Instant payments comprise a unique use case, and are changing the risk calculus because speed and irrevocability make prevention more urgent and recovery more difficult once funds move.

MARKET DATA

The broader market data points in the same direction. Alloy's 2026 State of Fraud Report found that 67% of financial institutions and fintechs reported an increase in fraud events, 91% reported more financial crimes being committed with AI technology, and 22% lost more than $5 million to fraud in the prior year.[2]

The 2025 AFP Payments Fraud and Control Survey similarly found that 79% of organizations experienced attempted or actual payments fraud in 2024, while only 22% recovered 75% or more of lost funds, down from 41% the year before.[3]

The threat is also omnichannel and increasingly blended. AI amplifies synthetic identity fraud, account takeover, authorized push payment scams, money mule networks, agentic bots, and phantom hacker scams alongside persistent legacy vulnerabilities such as check fraud and business email compromise.

AFP also reported that 63% of organizations identified business email compromise as a leading payments fraud avenue and that 63% experienced attempted or actual check fraud in 2024, even as 91% continued to use checks for payments. Javelin Strategy & Research's 2026 Fraud Management Trends likewise identifies money mules, agentic bots, and phantom hacker scams as major emerging threats.[4]

EPCOR, a national payments association serving more than 1,700 institutions, reinforced that this convergence is already showing up in day-to-day financial institution operations. Participants in EPCOR's 2026 Fraud Roundtables described fraud as a coordinated, multi-rail threat spanning checks, cards, ACH, wires, digital wallets, online banking, and P2P channels.[5]

They also drew a critical planning distinction: debit cards, digital wallets, P2P apps, and disputes create the highest operational volume, while wires, commercial checks, ACH account takeover, and business email compromise create the most severe losses. That distinction is especially important for agentic AI risk because autonomous systems can increase both the volume of low-dollar attempts and the speed at which high-severity events are orchestrated.

SECTION 04

The dispute resolution gap

A related concern is emerging in payments and dispute resolution. Existing rules were written for a world in which a person, not an autonomous agent, made the operative purchase decision. EPCOR has noted that AI-powered tools, including virtual voice assistants, are already part of the payments ecosystem and that they introduce new complexity into authorization, Standard Entry Class code selection, and compliance analysis. At the same time, fundamentals remain unchanged. Transaction amount, timing, frequency, receiver information, account information, authorization and revocation terms, and appropriate fraud and identity controls still must be captured and evaluated when a transaction is investigated or disputed.

That framework becomes more difficult to apply when the disputed act is performed by an agent interpreting delegated instructions. If a consumer authorizes an agent to buy a red sweater under a set price limit, but the agent orders a white sweater, exceeds the limit once shipping and fees are included, or selects a service that technically satisfies the instruction while imposing materially higher all-in costs, the central question is no longer simply whether payment credentials were used. The question becomes whether the agent remained within the consumer's authorization parameters and whether those parameters were sufficiently specific, auditable, and preserved at the time of execution.

This creates a grey area that regulation has not yet fully addressed. Regulation E and the Nacha Operating Rules provide established processes for determining whether an error occurred, but agent-mediated commerce introduces new evidentiary questions.

NEW EVIDENTIARY QUESTIONS

- What instruction set did the consumer approve?
- What constraints did the agent receive?
- Were taxes, shipping, subscriptions, surcharges, or platform fees included in the spending limit?
- Did the merchant or platform present information in a way that an agent could accurately interpret?
- Did the agent's logs preserve enough detail to support dispute investigation?

There is also a market integrity dimension. If an agent is told that a user is willing to spend up to $1,000, that ceiling could become a pricing signal rather than a consumer protection boundary. In poorly governed environments, ranking algorithms, affiliate incentives, sponsored placements, or undisclosed compensation arrangements could steer the agent toward a higher-priced option that benefits an intermediary rather than the user. The risk is not only fraud in the traditional sense, it is misalignment between the consumer's delegated intent and the commercial incentives embedded in the systems the agent relies upon.

For financial institutions, merchants, and payment ecosystem participants, the practical implication is that authorization records will need to become more machine-readable, more granular, and more durable. EPCOR's guidance points to the same baseline discipline that already governs payment authorization and dispute handling: document the authorization method, preserve the material transaction terms, apply required fraud and identity controls, and use the applicable rules framework to determine whether an error occurred. Agentic AI does not eliminate those obligations. It raises the standard for proving that the obligations were satisfied.

The EPCOR Roundtables also pointed to practical evidence gaps that agent-mediated payments will intensify. Institutions emphasized the need to capture transaction amount, timing, frequency, receiver information, account details, authorization and revocation terms, customer warnings, and the identity and fraud controls applied at the time of the transaction. In an agentic environment, those records may need to expand to include the customer's delegated instruction set, the agent's constraints, the sequence of prompts or decisions that led to execution, and whether staff had sufficient authority to delay, escalate, or refuse a transaction when red flags appeared.

SECTION 05

Liveness detection and its shortcomings

To combat sophisticated fraud attacks, including those conducted by AI agents, many organizations employ liveness detection, a tool designed to identify whether or not a human is present at the moment of a transaction. Understanding fraud attacks by AI agents requires understanding of how current liveness systems work, and where each approach is vulnerable to agentic exploitation.

Passive liveness detection

Passive liveness detection analyzes a presented image or video stream for artifacts associated with “realness,” such as skin texture, three-dimensional facial geometry, and micro-movements consistent with breathing or pulse. These checks are designed to defeat presentation attacks from criminals holding up printed photographs or wearing silicone masks. Agentic AI defeats passive biometrics through injection. Rather than presenting a spoofed stimulus to the camera, an agent can intercept the video pipeline between the camera hardware and the verification API, substituting a photorealistic synthetic video stream generated in real time. The liveness system never sees the absence of a camera feed. It sees a feed that exhibits the surface-level characteristics passive systems are trained to accept.

Active liveness detection

Active liveness detection introduces a challenge-response layer. The system instructs the user to blink, turn their head, smile, or speak a phrase, then verifies that the response matches the instruction. The intent is to defeat pre-recorded video loops by requiring real-time responsiveness. Modern agentic systems defeat this approach through two mechanisms. First, vision-language models can read and interpret challenge instructions and generate a synthetic response appropriately calibrated to the challenge. Second, the real-time synthesis capabilities of current deepfake engines are sufficient to produce compliant responses within the timeframe most active liveness systems accept.

Document and selfie matching

Many verification workflows also require the submission of a government-issued identity document alongside a selfie, with a matching check performed to link the two. This approach layers identity verification on top of liveness detection. But its vulnerability to agentic attack may be compounded, not reduced, by this layering. Generative models can now produce synthetic identity documents that defeat template-based forensic checks. The same models can produce photorealistic face images consistent with those documents. An agentic system can combine these capabilities to submit a synthetic document, a synthetic selfie, and a synthetic liveness video in a single coordinated workflow, none of which reflects a real human being.

Behavioral biometrics and CAPTCHA

Finally, behavioral biometrics include tracking tools such as mouse movement patterns, typing cadence, scroll behavior, and click timing, and have been proposed as an additional signal layer. CAPTCHAs evolved to require tasks that were presumed to require human perception and judgment. Agentic AI systems trained on a large body of human interaction data can now mimic human behavioral patterns sufficiently to defeat behavioral biometric systems. CAPTCHA systems face the same evolution that has characterized the arms race between deepfake generation and detection. Each generation of challenge is eventually solved by the next generation of model.

CONTROL	WHAT IT CHECKS	AGENTIC BYPASS
Passive liveness	Skin texture, facial geometry, micro-movement	Video-pipeline injection of real-time synthetic feed
Active liveness	Challenge-response: blink, turn, smile, speak	VLM reads the challenge; deepfake engine renders the response in time
Document + selfie	Template forensics, face-to-document match	Coordinated synthetic document, selfie, and liveness video
Behavioral / CAPTCHA	Mouse paths, typing cadence, perception tasks	Behavior mimicry learned from human interaction data

Figure 1 · Each control tests for the absence of known spoofing signals, not the presence of confirmed human signals.

SECTION 06

The path forward

The argument so far is largely one of failure modes. Each of the aforementioned attack types are, in one way or another, checking for the absence of known fraud signals rather than the presence of confirmed human signals. They ask, “Does this input look like it might be fake?” They do not ask, “Does this input prove it is real?” Asking the latter question can make all the difference. This distinction is not merely philosophical. It is the source of the asymmetry that makes conventional liveness detection ultimately inadequate against sophisticated adversaries. An attacker who understands what a system is looking for can construct inputs that lack those signals. But an attacker who cannot produce what the system is looking for cannot pass it, regardless of how sophisticated their other capabilities become.

First: continuous, multilayered identity assurance

Therefore, two things are essential to successfully preventing or defeating AI-related attacks. A practical first step to combating AI fraud attacks is to stop treating anomalous workflows and unexpected system behavior as routine technical defects. EPCOR CEO Brandie Thacker noted that injection attacks are often misdiagnosed as bugs, prompting workarounds that can unintentionally reinforce the attack path.[6] The path forward therefore requires tighter integration among fraud, cyber, engineering, operations, product, and compliance teams. When these teams operate in silos, organizations may solve the visible symptom while leaving the underlying AI-driven attack intact.

The recommended control model is continuous, multilayered identity assurance rather than one-time verification. No single control, including KYC, liveness detection, device analysis, behavioral biometrics, or transaction monitoring, can serve as a silver bullet. Stronger defenses will weave together behavioral patterns, device integrity, camera and liveness signals, transaction history, image-level signals, and transaction tagging so that each event is evaluated in context. The objective is not merely to identify whether an image, voice, or session appears fake; it is to establish whether the action is anchored to a trusted identity, a known device or environment, and a plausible pattern of human intent. This also requires better recognition and reporting of deepfakes and synthetic identities.

Second: human verification, not bot detection

Secondly, frontline intervention as a core fraud control is paramount, and a shift from bot detection to human verification is a necessary addition to any fraud prevention stack. But rather than continuing participation in a perpetual arms race with bad actors, new solutions, such as Moveris' human verification system that detects actual human physiological signals, are necessary to stay ahead of bad actors and effectively determine whether a real human is personally involved in a transaction, or not.

In addition, it's also clear that technology alone cannot carry this burden, and may especially benefit from human-in-the-loop review before customer-impacting actions such as blocking payments, refusing wires, restricting accounts, or treating an alert as dispositive. A stronger control model therefore pairs automated human verification with defined escalation authority, documented customer contact, well-tuned alerts, and staff procedures that convert signals into defensible decisions.

As agentic AI becomes embedded in everyday workflows, organizations must be able to prove human authorization, not simply automated completion. Agentic AI is an amplifier. It can accelerate legitimate activity or scale abuse. Without a multi-layered approach that includes human verification, organizations will struggle to distinguish human intent from automated execution. The path forward is therefore not to build higher walls around isolated controls, but to build stronger connections across signals, teams, institutions, and rules frameworks.

IN CLOSING

The emergence of agentic AI as a mainstream capability requires a fundamental reassessment of what liveness detection is for and how it should work. The old answer of detecting the absence of known spoofing signals is insufficient against adversaries that can study, model, and defeat any pattern a system is trained to require. The new answer must be grounded in the presence of signals that cannot be spoofed. The threat will continue to evolve. Agentic AI systems will become more capable, more accessible, and more widely deployed, for legitimate and fraudulent purposes alike. The question for organizations is not whether to engage with this evolution, but whether their verification infrastructure will keep pace with it.

ENDNOTES

1. Allouah, A., Besbes, O., Figueroa, J. D., Kanoria, Y., & Kumar, A. (2025). *What is your AI agent buying? Evaluation, biases, model dependence, and emerging implications for agentic e-commerce.* arXiv preprint arXiv:2508.02630.
2. Alloy. (2025). *2026 State of Fraud Report.* Alloy.
3. Association for Financial Professionals. (2025). *2025 AFP Payments Fraud and Control Survey Report.* Underwritten by Truist.
4. Javelin Strategy & Research. (2025). *2026 Fraud Management Trends.*
5. EPCOR. (2026). *2026 EPCOR Fraud Roundtable Report.* EPCOR.
6. Thacker, B. (2026). *EPCOR Payments Insights: A Deep Dive on Houston's Deepfake Summit* [Video]. EPCOR. YouTube.