

REPORT

2026 Fraud and AML trends

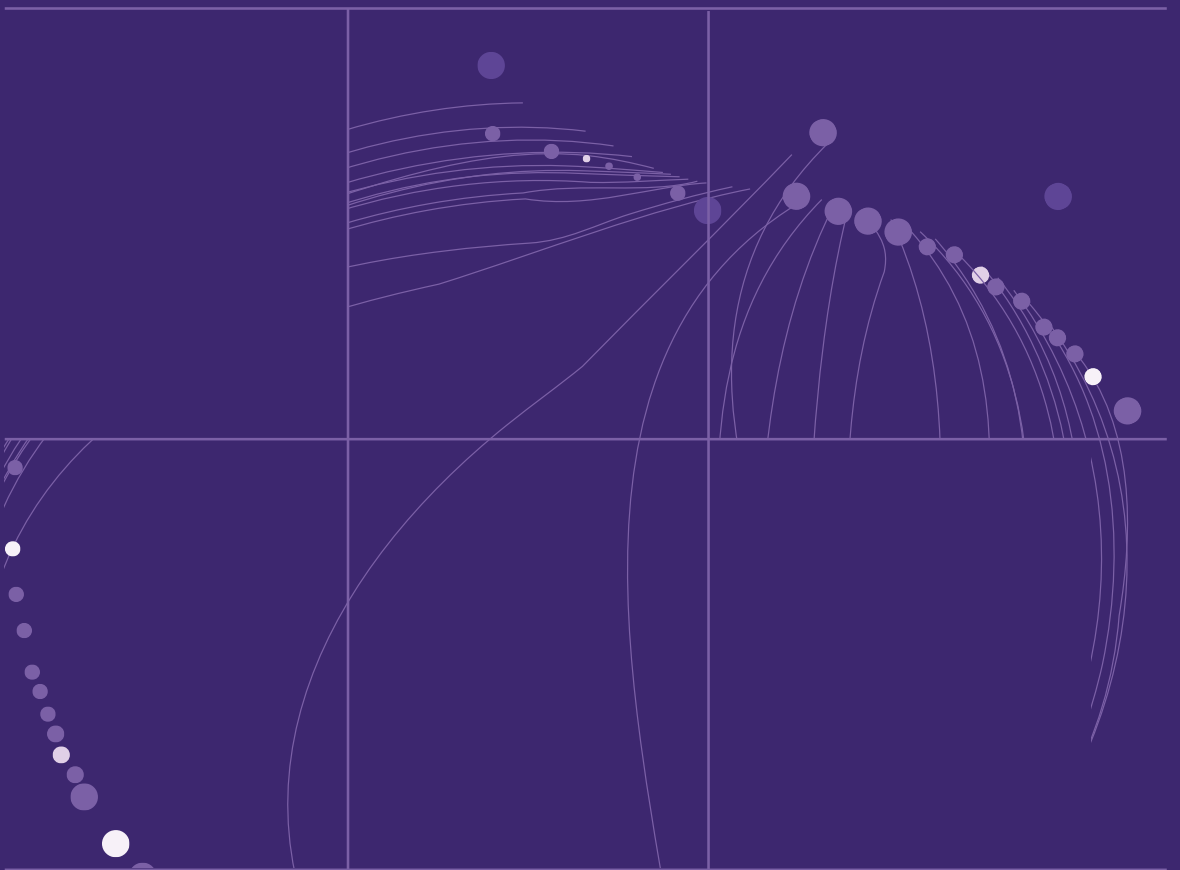


Table of Contents

Executive summary	4
What's change	5
Why it matters	5
How to plan for 2026	5
What's going on with fraudsters	6
What happened, and what we saw	7
What this means	9
What's working right now	10
Where to put your resources	10
Do this next	11
What's going on in the industry	12
What happened, and what we saw	13
What it means	13
What's working right now	14
Where to put your resources	15
Do this next	15
What's going on with regulators	16
What happened, and what we saw	17
What it means	17
What's working right now	18
Where to put your resources	18
Do this next	19
Your planning checklist	20
Can we interrupt scams in minutes?	21
Do identity and payments talk in real-time?	22
Can we slow fraud without breaking the customer experience?	22
Are agents, entities, and delegation explicitly controlled?	23
Are our decisions explainable to an examiner?	23
Are we learning faster than attackers?	24
The takeaway	24
Building for interruption	25
Detection alone isn't enough anymore	26
Treating trust as infrastructure	27
Where AI agents can change outcomes	27
Systems adapt faster than rules	27
What comes next	28
References	29

This isn't another trend report. In fact, it's an anti-trend report.

Instead of guessing what might happen next, we're grounding this report in what we actually saw across fraud, the industry, and regulators in 2025, what it means for 2026 planning, and what fraud leaders should do about it. No hype, no predictions for prediction's sake. Just prescriptive insights you can use.

As we wrapped 2025, one thing became clear. There was no shortage of noise, from scary new tactics to bold claims about AI changing everything overnight. But the real shifts didn't happen where most headlines focused. They happened in how fraud operates, how defenses fail, and how slowly systems respond once something goes wrong.

As you plan for 2026 and beyond, this report focuses on where to **invest**, what to **keep**, and what to **stop** doing. We'll share what peers are doing that's working right now, and how to apply those lessons in your own environment.

It's time to plan fraud defenses as a system, not a set of controls.

01



Executive summary

From incidents to systems

What changed

In 2025, fraud stopped behaving like a system of isolated incidents. Teams could no longer expect fraud to ebb and flow at the pace of humans, or to show up as obvious spikes that triggered investigations. Instead, scams, identity abuse, payments fraud, and AML activity increasingly operated as one continuous system. These threats reinforced each other, moved across channels, and exploited gaps between teams and tools. AI certainly didn't invent fraud, but it did fundamentally change its operating model. It compressed timelines, scaled coordination, and removed much of the friction that historically slowed attackers down.

Why it matters

Most fraud programs are still built for discrete events. Rules fire, cases are reviewed, losses are hopefully recovered, and teams move on. But that model doesn't hold up against continuous, distributed abuse.

Losses now surface later, often after multiple successful approvals across identity, payments, and account changes. In many cases, teams don't even realize they're under attack until the damage is already done.

This is just one reason why fraud costs continue to climb. According to one industry report, U.S. financial services organizations lose \$3.99 for every \$1 of fraud loss.¹

How to plan for 2026

Fraud moves too fast for static controls, this much we know. Planning for 2026 and beyond means funding systems, not just rules.

Detection alone is no longer enough. Teams need to detect fraud early enough to interrupt its flow, slow it down, or stop it entirely before value leaves the system.

This also means treating trust as infrastructure. Like any critical infrastructure, trust systems need to hold up under sustained pressure, not just pass compliance checks on a good day.

02



What's going on with fraudsters

The big idea: Fraudsters are no longer running attacks. They're operating systems.



Karen Boyer, SVP of Financial Crimes, M&T Bank



Fraud is the new friction, and fraud prevention is the new customer service.”

What happened, and what we saw

Scams and impersonation

Scams surpassed “classic fraud” in both volume and losses in 2025, led by impersonation and social engineering. This shift is not about customers being careless. In fact, many experienced professionals have fallen for the same tactics.

AI scaled scams in practical ways: better scripts, more precise victim targeting, deepfake audio and video, and faster iteration when something didn’t work.

Forward-looking fraud leaders are already warning about new coercion formats, including intimidation-style impersonation scams such as “digital arrest” scenarios, where victims are pressured into immediate action through fear and urgency.²

According to the FBI’s Internet Crime Complaint Center (IC3) 2024 report, victims reported \$16.6 billion in losses across more than 859,000 complaints. Investment fraud, often involving crypto, drove the largest losses, and phishing, spoofing, and extortion led by volume.

Older adults continued to be disproportionately harmed. The FBI reported nearly \$5 billion in losses among victims aged 60 or older.³ FTC Consumer Sentinel data shows the 60+ age group lost \$3.1 billion, with a median loss of \$18,000.⁴

Crypto as a value-extraction rail

Crypto remained a major value-extraction rail for sophisticated scam ecosystems. Instant payments and crypto exchanges continued to be preferred cash-out paths. Nacha, citing FBI IC3 data, highlighted crypto’s growing role in scam-related losses.⁵ Chainalysis reported \$10.5 billion in illicit crypto volume, with investment scams as the largest component.⁶ The U.S. Treasury’s Financial Crimes Risk Assessment also noted faster deception-based scams and crypto cash-outs as growing risks.⁷

Synthetic identities and deepfakes

Synthetic identity fraud became cheaper and faster to execute. What once required time, coordination, and stolen identity fragments can now be generated, matured, and reused with the help of AI. Deepfake creation and KYC bypass services have been observed for as little as \$150, according to reporting by NativeRisk. Europol estimates that more than 60 percent of detected AI forgery incidents relate to identity proofing or KYC bypass.⁸

The U.S. Federal Reserve has estimated that synthetic identity fraud accounts for 15-20 percent of credit losses in unsecured lending.⁹

Business email compromise (BEC)

BEC remained one of the most financially damaging fraud categories. IC3 data shows nearly \$8.5 billion in BEC losses over the last three years, consistently placing it among the top-dollar fraud types.¹⁰

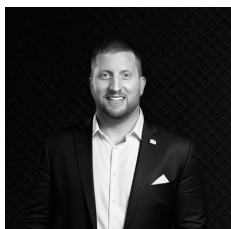
Fraud on paper: Checks and mail theft

While AI is powering new forms of fraud, we've also seen a resurgence in the old school methods. Mail theft and check fraud didn't go away. Instead, they adapted by continuing to exploit something attackers understand well: operational delay. These schemes rely less on technical sophistication and more on the inherent slowness of paper-based processes. Float time, manual review, and fragmented ownership across teams create windows where fraud can succeed before controls or investigations catch up.

This persistence shows up clearly in the data. The FBI and U.S. Postal Inspection Service reported that SARs tied to check fraud nearly doubled between 2021 and 2023, with a "significant volume" enabled by mail theft. Even as digital fraud accelerated, paper-based channels remained an attractive target precisely because they move at human speed.¹¹

Agent-assisted fraud

AI agents are mostly used for good. But when they are abused, the impact escalates quickly. We are seeing early signs of semi-autonomous and agent-assisted fraud workflows. As Matt Vega described in commentary shared via the *Frank on Fraud* blog, these systems can learn from declines and approvals, autonomously adjusting tactics in real-time rather than following predefined paths.¹²



Matt Vega, Chief of Staff,
Sardine



The attacks will be fast, flexible, self-replicating, and will no longer follow pre-defined adaptive pathways associated with polymorphic malware.

Instead, these agents will take in the response, feedback, and the trial and errors from declines or approvals to autonomously shape-shift into a new attack based on what they learn.”

What this means

Fraud is increasingly indistinguishable from legitimate activity at the moment it occurs. Many of today’s highest-loss events are authorized, coerced, or behaviorally consistent with normal customer actions. That makes them harder to flag, harder to explain, and harder to interrupt once they’re underway.

Identity compromise and payment abuse can no longer be treated as separate problems. Fraudsters routinely pass onboarding, establish tenure, and then exploit that trust during high-risk moments. As Steve Lenderman, Head of Fraud Prevention at iSolved said on Sardine’s *Fraud Forward* podcast, **“Every fraud that takes place passes KYC.”**

When identity and payments are evaluated in isolation, each individual event looks clean. The risk only becomes visible when those events are stitched together across time.

Attacks are also more distributed than most controls assume. Fraud no longer arrives as a single spike or campaign. It unfolds gradually, across channels, accounts, and institutions, often below traditional alert thresholds. By the time losses surface, multiple systems have already approved activity.

The most dangerous blind spot for many teams is how normal modern fraud looks when viewed through a single control, queue, or domain.

Don’t underestimate this:

How clean fraud looks when identity and payments aren’t evaluated together.

What's working right now

The institutions seeing real progress are not trying to predict every new scam variant. Instead, they're investing in interruption. They're focusing on slowing fraud down, inserting friction at the right moment, and creating opportunities for human intervention before funds leave the system.

Risk-based payment friction continues to outperform static warnings or blanket limits. When friction is tied to scam markers, such as first-time payees, unusual channels, urgency, or out-of-pattern amounts, it disrupts scams without broadly degrading the customer experience. This approach acknowledges that many scam payments look legitimate in isolation and only become risky in context.

More teams are also paying closer attention to the receiver, not just the sender. **Verification of payee** and receiver intelligence have historically been underemphasized, particularly for non-corporate recipients. As payments increasingly flow to gig workers, casual sellers, and individuals operating across platforms, understanding who is on the other side of a transaction matters more than ever. Institutions that incorporate payee verification, seller intelligence, and consortium signals are better positioned to identify risky destinations before money moves.

Real-time holds and step-ups matter because scams rely on momentum. Once a victim is under pressure, even a short delay can be enough to break the spell. Teams that can pause, escalate, and engage customers in the moment are preventing losses that would otherwise be unrecoverable.

Behavioral and device intelligence is also playing a growing role. In a world of deepfakes and synthetic identities, signals that persist beyond onboarding are becoming critical for monitoring trust that change over time, especially when both senders and receivers appear legitimate on the surface.

Where to put resources

For 2026 planning, **scams should be treated as a payments and customer safety program**, not just a fraud operations problem. That means funding cross-functional ownership, clear escalation authority, and tooling that supports real-time decisions.

Resources are better spent on **interruption and coordination**, rather than adding more static checks at onboarding. Fraudsters increasingly pass those checks anyway, and the damage happens later when money moves.

Teams should assume that attackers will look legitimate at first. Planning should prioritize **what happens after authentication succeeds**, when trust is most likely to be abused, not just how to prevent access in the first place.

While scams should be treated as a payments and customer safety problem, that doesn't mean payments teams need an entirely new set of defenses. What's missing isn't new payment rails and controls, it's a tighter coupling between identity, behavior, and money movement. The priority for 2026 should be securing how value moves once trust has been granted, and making sure identity and payments are evaluated together in real-time.

Do this next

- 1 **Build an end-to-end scam disrupt playbook that includes:**
 - Alert triggers
 - Friction options
 - Escalation paths
 - “Safe exit” scripts

- 2 **Shift how you measure success. Focus on:**
 - Scam saves (prevented loss)
 - Time-to-interrupt
 - Repeat victim rate

- 3 **Keep what still works for core payments, such as:**
 - Strong MFA and device intelligence to prevent unauthorized access
 - Unusual-activity detection for routine money movement
 - Call-back verification for high-risk account or payment changes

- 4 **Add what's missing, like:**
 - Real-time holds and step-ups tied to scam markers
 - First-time payee verification
 - Risk-based escalation authority

- 5 **Rebalance your identity strategy to:**
 - Reduce reliance on static “hard” verification alone
 - Increase behavioral monitoring, risk segmentation, and signal layering

03



What's going on in the industry

The big idea: Leading institutions
are reorganizing around real-time
trust, not isolated controls.



Hailey Windham, Community Lead, Banking at Sardine

“

As a former practitioner in credit unions and community banks, I see a clear shift happening. The institutions pulling ahead aren't adding more siloed controls, they're reorganizing around real-time trust.

With lean teams and limited resources, trust has to be assessed in the moment, across every transaction, not after the fact.”

What happened, and what we saw

Banks and credit unions increasingly implemented risk-based friction for P2P and instant payments, rather than relying solely on consumer warnings. Network-level guidance reinforced expectations for participant risk management.¹³

Instant payments matured through FedNow and RTP expansion, forcing teams to manage 24/7 risk with less time to react.

The Federal Reserve reported that staffing load and attrition grew **22 percent year** over year following the introduction of instant payments.¹⁴

Identity and digital intelligence emerged as the shared control plane for onboarding and transaction security.

Institutions began using AI in more targeted ways to augment fraud operations, particularly in triage, enrichment, and investigation support.

Approaches to scam prevention diverged. For example, Chase implemented delays and blocks on certain Zelle payments tied to social media-origin scams.¹⁵ PayPal introduced AI-driven scam detection and customer alerts.¹⁶

What it means

The industry is slowly but decisively **moving away from siloed fraud programs**. Real-time payments, always-on customer experiences, and scam-driven losses have exposed the limits of channel-specific controls and queue-based review models.

Instant payments have accelerated this shift. When funds move in seconds and operations run 24/7, there is no room for handoffs between disconnected teams.

Fraud, disputes, call centers, BSA/AML, and security are all touching the same events, often without shared visibility or authority.

Identity and digital intelligence partnering up is a natural response to this pressure. Institutions are realizing that onboarding decisions, account changes, and payment approvals cannot be evaluated independently. Trust must be assessed continuously, across the customer lifecycle.

AI has started to play a meaningful role, but not as a replacement for human judgment. Where it's working, AI is being used to compress time: surfacing risk faster, enriching context automatically, and helping teams act before losses occur. Without clear workflows and authority, however, AI adds noise rather than value. The best practice remains having a human-in-the-loop that acts on the AI's data synthesis and recommendations.

Don't underestimate this:

Coordination is now as important as detection performance. A highly accurate alert that arrives too late, or lands with the wrong team, is functionally useless.

What's working right now

Institutions making progress have invested in operating models, not just tooling. Payments fraud "fusion" cadences bring together fraud, disputes, call centers, AML, and security to review risk holistically and align on intervention strategies. These forums reduce lag, clarify ownership, and surface patterns that would otherwise remain fragmented.

Risk-based payment friction has also proven effective when applied dynamically. Rather than blanket warnings or static limits, teams are tailoring step-ups based on context: who the payee is, how the payment was initiated, and whether the behavior aligns with historical patterns.

For instant payments, institutions are leaning more heavily on network-level controls and reporting. FedNow guidance explicitly calls out participant responsibilities for managing instant-payment risk, and teams that integrate those controls early are better positioned to operate at speed.

Consortium and network intelligence are helping surface coordinated activity that no single institution can see alone. This is particularly important for scams and synthetic identity abuse, where reuse across platforms is common.

Finally, AI-assisted triage and investigation support are reducing analyst burden where they are tightly scoped and well-governed. The goal is not automation for its own sake, but faster, more consistent decision-making.

Where to put your resources

For 2026, the biggest returns are coming from changes to **how teams work together**, not from adding another standalone tool. Funding should prioritize **coordination**, **shared visibility**, and **clear authority** to intervene in real-time.

AI investments should be evaluated through a single lens: does this reduce time-to-action? If it only improves analyst efficiency after a loss has occurred, it's solving the wrong problem.

Build versus buy decisions are also shifting. AI lowers the barrier to building certain capabilities in-house, but only when teams have the domain expertise, data quality, and operational maturity to support them. In many cases, **buying remains the faster and safer path**, especially for controls that need to operate reliably at scale.

Do this next

- 1 **Stand up a payments fraud “fusion” cadence that includes:**
 - Fraud, disputes, call center, BSA/AML, security
 - Shared metrics and escalation paths
- 2 **Implement risk-based payment friction such as:**
 - Dynamic limits
 - Payee verification and step-ups
 - Customer prompts that capture scam indicators
- 3 **For instant payments, be sure to:**
 - Use network controls and reporting explicitly recommended by FedNow
 - Validate 24x7 operational coverage and escalation
- 4 **Adopt metrics leadership understands:**
 - Scam saves and time-to-interrupt
 - Percentage of high-risk payments with step-ups applied
 - Repeat victim rate
 - Check fraud detection lead time versus funds availability
- 5 **Train and empower staff with:**
 - Regular training on scam behaviors and coercion patterns
 - Clear policy cover to slow or stop transactions

04



What's going on with regulators

The big idea: Regulators are circling payments fraud, but the rules haven't fully landed.



Nyla Cortes, Director of Compliance, Earthmover Credit Union



Regulators are paying close attention to payments fraud risk, especially around authentication and consumer protection, but key frameworks are still missing. Until the U.S. takes a clearer stance on APP scams and delegated decision-making, institutions are left managing risk in a policy gray zone.”

What happened, and what we saw

U.S. banking agencies increased attention on payments fraud across rails and requested industry input on potential actions.¹⁷

Regulators continued emphasizing strong authentication as baseline hygiene, reinforcing layered and risk-based access controls.¹⁸

What did not happen is equally important. There is still no clear U.S. stance on APP scam reimbursement. In the U.K, new rules resulted in a **46 percent reimbursement rate and £459 million in APP fraud losses**.¹⁹ Similar clarity has not yet emerged in the U.S. Regulators have also largely avoided addressing agentic commerce and delegated decision-making, despite increasing adoption and investment.

What it means

Regulators are signaling concern about payments fraud across rails, but they are still in information-gathering mode. Requests for industry feedback and reinforced expectations around authentication suggest increased scrutiny, even without immediate new rules.

The absence of a clear U.S. stance on APP scam reimbursement is notable, especially given developments in the U.K. and Australia. As scam losses continue to concentrate, it becomes harder for regulators to remain on the sidelines indefinitely. The lack of regulatory attention on agentic commerce and delegated decision-making is equally important. As AI systems increasingly initiate or influence transactions, questions around liability, control, and accountability remain unresolved.

The direction we're going is clear, even if the destination is not. Liability is increasingly tied to who had control and whether that control was reasonable, effective, and continuously monitored.

Don't underestimate this:

How quickly regulatory posture can change once losses reach a tipping point or public pressure mounts.

What's working right now

Institutions that treat **trust systems as regulated infrastructure** are better positioned for future scrutiny. This means designing controls that are explainable, monitored, and adaptable, rather than narrowly optimized for today's guidance.

Documented decision logic and escalation paths are becoming just as important as detection itself. Teams that can clearly articulate why a transaction was allowed, delayed, or blocked are better prepared for examiner questions. Consistency in decision-making is what regulators ultimately look for, especially when losses occur.

Monitoring international developments, particularly in the UK and Australia, is also proving valuable. These markets often serve as early indicators of where regulatory thinking may go next, even if timelines differ.

Where to put your resources

Where to put your resources

Planning for 2026 should assume regulatory attention will accelerate, not fade. Resources are best spent on **building flexibility** into systems and roadmaps, rather than betting on regulatory inertia.

Teams should **avoid designing controls that only satisfy current rules**. When new guidance arrives, narrowly built systems are more expensive and disruptive to rework.

Agent-driven transactions and delegated authority deserve **early internal attention**, even without explicit regulation. Waiting for formal rules increases the risk of rushed, reactive changes later.

Do this next:

- 1 Stress-test your roadmap, making sure it includes:**
 - APP scam reimbursement scenarios
 - Agent-driven transaction liability
 - Cross-rail fraud obligations

- 2 Pair detection with defensibility through:**
 - Monitoring and documentation
 - Clear ownership and escalation paths
 - Explainable and consistent decision logic

- 3 Track leading indicators such as:**
 - U.K and E.U regulatory developments
 - Industry enforcement actions
 - Public-loss concentration signals

- 4 Plan for change like:**
 - Assume draft guidance will emerge before final rules
 - Build adaptability into 2026 initiatives

05



Your planning checklist

A practical gut check for fraud
leaders

Rather than worrying about maturity scoring or benchmarking, here are practical ways to pressure-test whether your fraud program is built for how fraud actually operates now.

Each question reflects a failure mode we repeatedly see when fraud moves faster than systems can respond.

Can we interrupt scams in minutes?

Why this matters

Most scam losses are not caused by a failure to detect risk. Instead they're caused by delays in acting on it. Once a victim is under pressure, momentum works against you. Every additional minute increases the likelihood that funds will move and become unrecoverable.

Sanity check

- How long does it take today from the first scam signal to intervention?
- Can front-line teams slow or stop a transaction without manager approval?
- Are you monitoring from the date of the first scam signal detection?
- Are you performing root cause analysis after scam loss?

What good looks like

- Real-time or near-real-time holds tied to scam markers
- Clear escalation paths when risk is detected
- Authority for frontline to pause transactions without waiting for post-event review

Do identity and payments talk in real-time?

Why this matters

Fraud increasingly passes onboarding and shows up later, when money moves. When identity risk and payment risk are evaluated separately, each decision looks clean on its own. The fraud only becomes visible when those decisions are stitched together across time.

Sanity check

- Could a high-risk payment be approved even if identity confidence is actively degrading?
- Are identity signals reused downstream, or trapped in onboarding tools?

What good looks like

- Identity confidence that rises and falls with behavior
- Payment decisions informed by recent identity, device, and behavioral signals
- Shared visibility across onboarding, account changes, and transactions

Can we slow fraud without breaking the customer experience?

Why this matters

Scams rely on urgency. Strategic friction, applied at the right moment, is one of the most effective ways to disrupt them. When you apply blanket friction everywhere, customers will just become frustrated.

Sanity check

- Are most of your controls binary (allow/deny), or do you have graduated responses?
- Can you (or your frontline) explain to a customer why a payment was slowed in plain language?

What good looks like

- Risk-based friction triggered by context, not channel
- Step-ups that are explainable to customers
- Short delays that create space for intervention

Are agents, entities, and delegation explicitly controlled?

Why this matters

As AI systems begin to initiate or influence transactions, traditional human-first assumptions break down. If authority is delegated without clear limits, mistakes and abuse become harder to unwind.

Sanity check

- Do you know where automated systems can initiate or accelerate money movement?
- Who is accountable when an agent makes a bad decision?

What good looks like

- Clear rules around what agents can and cannot do
- Scoped permissions, limits, and reversibility
- Monitoring and override paths for agent-driven actions

Are our decisions explainable to an examiner?

Why this matters

Regulatory expectations are moving toward control effectiveness and governance, not just outcomes. When losses occur, the question increasingly becomes whether controls were reasonable, monitored, and adaptive.

What good looks like

- Documented decision logic and escalation paths, consistently applied.
- Clear ownership across fraud, payments, and compliance
- The ability to explain why an action was taken, not just what happened

Sanity check

- Could you explain a blocked or delayed transaction months later with confidence?
- Are key decisions documented, or locked inside individual tools?

Are we learning faster than attackers?

Why this matters

Fraud systems that don't learn quickly fall behind quickly. Attackers iterate in real-time, adjusting based on what succeeds and what fails.

Sanity check

- How often do insights from losses translate into control changes?
- Are lessons shared across teams, or stuck in postmortems?

What good looks like

- Feedback loops from declines, interventions, and losses
- Regular cross-team reviews of scam and fraud patterns
- The ability to adjust controls without long release cycles

The takeaway

This checklist isn't about perfection. It's about readiness.

If you can interrupt scams quickly, evaluate identity and payments together, apply friction surgically, control delegation, explain decisions, and learn faster than attackers, you're planning for fraud as a system.

If not, 2026 will feel reactive no matter how many tools you add.

06



Building for interruption

At Sardine, we've been planning for the same shifts outlined in this report not because they're emerging, but because they're already here.

Our approach starts with a simple premise: fraud defense has to operate as a system. That means identity, payments, fraud, and AML decisions cannot live in separate silos, and they cannot rely on static checks that assume fraud will arrive slowly or look obvious.

Our point of view is straightforward: **fraud defense must be designed to interrupt risk, not just document it.**



Soups Ranjan CEO, Sardine



What banks are increasingly realizing is that they have a lot of vendor sprawl. They've had several point solutions, one for fraud at onboarding, one for fraud at login for account takeovers, another for just payment fraud, and then a fourth system for AML. Even the cybersecurity team has its own system for detecting denial of service attacks.

And none of these things talk to each other. And fraudsters essentially live in the gaps or the shadows between these point solutions.”

Detection alone isn't enough anymore

Most institutions don't struggle to surface fraud signals. They struggle to act on them in time. Modern fraud often looks legitimate in isolation. Credentials are real, devices are familiar, and payments are authorized. The risk only becomes visible when identity, behavior, and money movement are evaluated together in the same view, across time.

That's why Sardine's approach treats fraud defense as a connected system. Identity signals, behavioral risk, transaction context, and network intelligence are evaluated continuously, rather than handed off between tools or teams. In practice, this means moving decision-making closer to the point of action, during account changes, payment initiation, and other high-risk moments, instead of relying on post-event cleanup.

Treating trust as infrastructure

Trust is no longer something you establish once and rely on forever. It degrades, recovers, and shifts as behavior changes.

Systems built for static trust inevitably fail when attackers establish tenure and wait for the right time to strike. But systems built for continuous trust can slow fraud down, even when it shows up disguised as normal behavior.

This is the difference between blocking everything, and interrupting selectively.

The goal isn't more friction, but better timed friction. Low-risk activity continues uninterrupted. High-risk moments trigger holds, steps-ups, or escalation for humans to intervene.

Where AI agents can change outcomes

AI doesn't replace human judgement in fraud defense, but it can compress time in ways that humans can't.

At Sardine, AI agents are used to continuously monitor risk, correlate signals, and resolve routine decisions automatically so humans can step in where it matters most. Agents gather context, surface explainable signals, and escalate only when risk thresholds are crossed.

By compressing time between signal and action, we can reduce dependency on manual triage. This also allows fraud teams to operate at the same speed at modern payments, without sacrificing control.

It's crucial that every one of these decisions remain explainable and auditable. Agents don't operate in a black box. Instead they follow explicit workflows, thresholds, and escalation paths that can be reviewed by operators and examiners alike.

Systems adapt faster than rules

Fraud evolves too fast for static controls to keep up. Chasing new scam variants with new rules is expensive and reactive.

What actually scales adaptability? Systems that learn from outcomes, adjust without long release cycles, and allow teams to change how they respond to risk, not just what they detect.

This is where Sardine continues to invest: in platforms that support interruption, coordination, and defensibility as design goals, not afterthoughts bolted on to legacy workflows.

What comes next

Fraud defense in 2026 isn't about better rules. It's about systems that can reason, adapt, and act at machine speed, while keeping humans firmly in control where judgment matters most.

As fraud becomes more continuous, more coordinated, and more behaviorally convincing, the institutions that succeed won't be the ones with the most alerts. They'll be the ones that can interrupt risk early, explain their decisions clearly, apply consistently, and adapt faster than attackers can iterate.

That's the standard we're building toward, and the lens we'll continue using as this year unfolds.

References

- ¹ **2025 LexisNexis® True Cost of Fraud Report:** <https://risk.lexisnexis.com/insights-resources/research/us-ca-true-cost-of-fraud-study>
- ² **Frank on Fraud: No Brakes, No Limits. Our Fraud Predictions For 2026**
<https://frankonfraud.com/no-brakes-no-limits-our-fraud-predictions-for-2026/>
- ³ **FBI Internet Crime Report 2024** https://www.ic3.gov/AnnualReport/Reports/2024_IC3Report.pdf
- ⁴ **FTC Issues Annual Report to Congress on Agency's Actions to Protect Older Adults**
www.ftc.gov/news-events/news/press-releases/2025/12/ftc-issues-annual-report-congress-agencys-actions-protect-older-adults
- ⁵ **FBI's IC3 Finds Almost \$8.5B Lost to Business Email Compromise in Last Three Years:** www.nacha.org/news/fbis-ic3-finds-almost-85-billion-lost-business-email-compromise-last-three-years
- ⁶ **The Chainalysis 2025 Crypto Crime Report:** www.chainalysis.com/blog/2025-crypto-crime-report-introduction/
- ⁷ **U.S. Department of the Treasury Illicit Risk Assessment of Decentralized Finance:**
<https://home.treasury.gov/system/files/136/DeFi-Risk-Full-Review.pdf>
- ⁸ **Europol: Facing Reality? Law enforcement and the challenge of deepfakes**
<https://www.europol.europa.eu/publications-events/publications/facing-reality-law-enforcement-and-challenge-of-deepfakes>
- ⁹ **The Federal Reserve: New survey of risk management officers reveals top concerns in 2024:** <https://www.frbservices.org/news/fed360/issues/040125/risk-management-survey-top-concerns-2024>
- ¹⁰ **FBI Internet Crime Report 2024** https://www.ic3.gov/AnnualReport/Reports/2024_IC3Report.pdf
- ¹¹ **FBI and USPIS PSA, Jan. 27, 2025: Mail Theft-Related Check Fraud is on the Rise**
<https://www.ic3.gov/PSA/2025/PSA250127>
- ¹² **Frank on Fraud: No Brakes, No Limits. Our Fraud Predictions For 2026**
<https://frankonfraud.com/no-brakes-no-limits-our-fraud-predictions-for-2026/>

- ¹³ **Federal Reserve:** FedNow Fraud controls and instant payments guide
<https://explore.fednow.org/explore-the-city?id=4&building=features-gallery&resource=87&role=fi&resourceTitle=fraud-controls-and-instant-payments%3A-what-you-need-to-know>
- ¹⁴ **The Federal Reserve: New survey: Innovative use cases drive businesses to instant payments** <https://www.frbservices.org/news/fed360/issues/050125/industry-perspective-faster-payments-survey-business>
- ¹⁵ **NBC San Diego: Why your bank may block your next Zelle payment**
<https://www.nbcsandiego.com/nbc-7-responds-2/zelle-banks-security-measures/3959840/>
- ¹⁶ **PayPal Newsroom: Introducing AI-Powered Scam Alerts for Friends and Family Payments**
[google.com/url?q=https://newsroom.paypal-corp.com/2025-7-21-Introducing-AI-Powered-Scam-Alerts-for-Friends-and-Family-Payments&sa=D&source=docs&ust=1770907822012838&usg=AOvVaw0Qe_fpxG59ve2DiuUzH-SC](https://newsroom.paypal-corp.com/2025-7-21-Introducing-AI-Powered-Scam-Alerts-for-Friends-and-Family-Payments&sa=D&source=docs&ust=1770907822012838&usg=AOvVaw0Qe_fpxG59ve2DiuUzH-SC)
- ¹⁷ **OCC, Federal Reserve, FDIC RFI: Request for Information on Potential Actions To Address Payments Fraud** <https://unblock.federalregister.gov/>
- ¹⁸ **FFIEC Authentication Guidance, 2021** <https://www.ffiec.gov/sites/default/files/media/press-releases/2021/authentication-and-access-to-financial-institution-services-and-systems.pdf>
- ¹⁹ **UK Payment Systems Regulator: Authorised push payment (APP) scams performance report 2024** <https://www.psr.org.uk/media/uaag25pp/app-fraud-publication-jul-2024-v6.pdf>