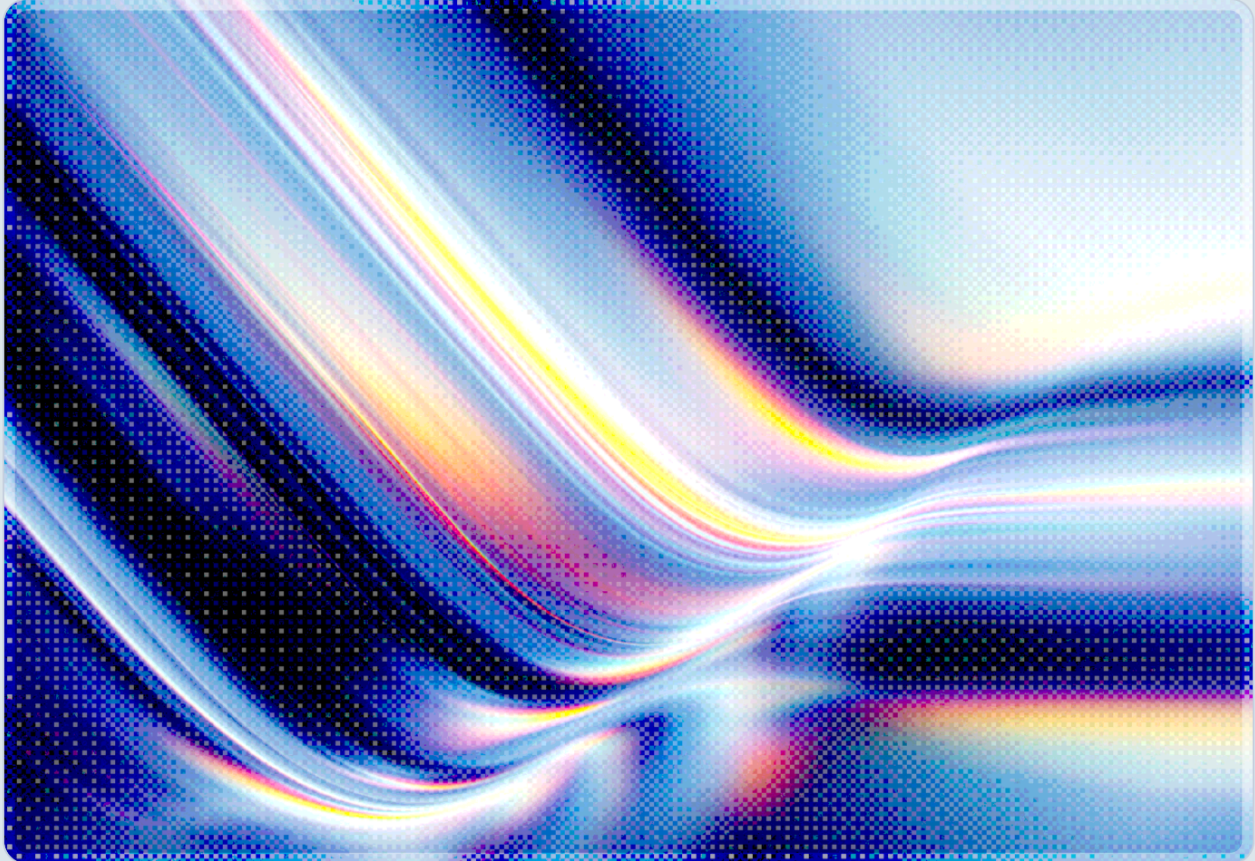


The Agentic Bank Run: A New Financial-Stability Risk

2026-09-28

 AI-assisted Rapid Research



CSAI

cloud
security
alliance®

© 2026 Cloud Security Alliance. Some rights reserved.

You may download, store, display, view, print, redistribute, and link to this document in its original, unmodified form, provided that attribution to the Cloud Security Alliance is maintained and all trademark and copyright notices remain intact.

This document may not be modified or altered. You may quote portions of the document as permitted by the Fair Use provisions of the United States Copyright Act, provided that attribution is given to the Cloud Security Alliance.

This document may be shared on professional and social media platforms in its original form with attribution.

This document was generated with AI assistance and has not undergone official CSA review and approval processes.

Key Takeaways

Apollo Global Management's chief economist, Torsten Slok, warned on September 27-28, 2026 that personal AI agents such as Meta's newly launched Muse could trigger what he termed an "agentic bank run": a coordinated, machine-speed migration of household cash out of low-yield checking accounts into higher-yielding fintech alternatives, executed simultaneously across millions of households without any single triggering event [1][2][3]. The mechanism does not depend on panic or rumor, the classic drivers of historical bank runs; it depends only on agents doing exactly what they are designed to do, which is optimize for the best available rate, at a moment when the gap between the national average checking yield of roughly 0.1% and fintech deposit products paying 3.3% to 5.0% gives every rational agent the same instruction [2][3]. This is a distinct risk category from the security vulnerabilities CSA typically documents in individual agent deployments: the danger here is not that an agent misbehaves or is compromised, but that thousands of correctly functioning agents, all optimizing independently toward the same signal, produce a collectively destabilizing outcome that no single institution intended or can unilaterally prevent [4][5]. Financial regulators have begun treating correlated AI behavior as a systemic-risk concern distinct from single-institution oversight: the Financial Stability Board's (FSB) June 2026 consultation report on AI adoption in finance identified agentic AI's speed and autonomy as risks that existing supervisory practices were not designed around [6][7], a concern that policy researchers have since framed explicitly in macroprudential terms [5], while the Bank for International Settlements (BIS) has separately flagged AI-industry financing concentration as its own pressure point on global financial stability [12]. For security and risk leaders, the agentic bank run scenario illustrates a broader point this note develops in detail: agentic AI risk is increasingly a question of aggregate system behavior, not just individual agent security, and organizations that only assess their own agents' controls will miss the exposure created by everyone else's agents acting in parallel.

Background

The proximate trigger for this discussion was Meta's September 8, 2026 launch of Muse, described by the company as the first personal AI agent built for everyday consumers, which included financial account connectivity from day one through a partnership with Plaid giving the agent read-only access to more than 12,000 U.S. financial institutions [8][16]. Muse can connect to a user's bank and credit card accounts, track spending, build budgets, and recommend where to cut costs, and Meta has said that

users who granted it this access found it an effective budgeting coach that prompted them to cancel unnecessary subscriptions [9]. Both Meta and Plaid have stated that the integration is read-only for account data and that Muse requires the user's explicit approval before it can authorize any spending [9] [16]. The product's reception was immediate: news of Muse's rise to the top of app-store charts and analyst upgrades on its monetization potential drove an 11% single-day jump in Meta's stock price in late September 2026 [10]. Muse is not alone in this category; a broader wave of agentic personal-finance tools, alongside crypto-native payment rails such as Coinbase's x402 protocol, which has already processed between 188 million and 205 million transactions across roughly 69,000 active autonomous agents, is normalizing the idea of software making real-time financial decisions on a user's behalf without per-transaction human approval [2].

Slok's argument, published through Apollo's "Daily Spark" research note and amplified by CNBC and CoinDesk, does not claim a bank run is imminent or that any institution is currently at risk; current reporting on the piece explicitly notes there are no signs a run on any bank is happening or is likely to happen soon [1][2][3]. His point is structural: the entire commercial banking model depends on a base of low-cost, "sticky" deposits that customers leave in checking accounts despite near-zero yield, because switching banks or actively chasing rates has historically required more effort than most depositors are willing to expend. An AI agent is designed to remove that friction. If millions of households delegate cash management to an agent instructed to maximize yield, and if many of those agents draw on similar training, similar rate-comparison data, and similar optimization logic, the result is not thousands of independent decisions but one large, correlated decision arrived at by parallel computation rather than coordination [2][3].

This concern did not originate with Slok, and the financial industry has direct experience with how quickly deposit flight can move once frictions are removed. The 2023 collapse of Silicon Valley Bank demonstrated that digital banking and social media could compress a bank run that once took days into a matter of hours, with customers attempting to withdraw nearly \$42 billion in a single day after depositors coordinated informally, in that case through Twitter posts and venture-capital group chats, rather than any centralized signal [11]. Commentators at the time described SVB's collapse as the first Twitter-fueled bank run precisely because the coordination mechanism, not the underlying financial weakness, explained why the outflow happened in hours rather than weeks [11]. An agentic bank run would replace informal human coordination with an even faster and more literal form of correlation: agents that receive the same rate data and are built on the same or similar underlying models may converge on the same action without any communication between them at all, a dynamic that has come to be called algorithmic herding.

Security Analysis

The core analytical challenge posed by the agentic bank run scenario is that it does not fit the vulnerability-and-patch model that dominates most agentic AI security discussion. There is no flaw in Muse or any comparable agent to disclose or fix; each individual agent may be functioning exactly as designed, correctly comparing rates and correctly moving funds to maximize the user's return. The risk instead emerges from aggregation: what is rational and even beneficial for one household, in isolation, becomes destabilizing when replicated across a large enough population acting on the same signal at the same time. CSA's own research on AI infrastructure concentration has documented a structurally similar dynamic in a different context, finding that when institutions rely on the same or highly similar models to make similar decisions, algorithmic herding can produce measurable amplification of losses, with modeled tail-loss amplification of 18% to 54% in financial-market scenarios once correlated model-driven decisions are introduced [4]. The agentic bank run scenario appears to be a structurally similar dynamic applied to a different domain: where that research models correlated trading decisions amplifying market losses, correlated cash-management decisions could analogously amplify deposit outflows, though the magnitude of that amplification in a retail-deposit context has not yet been separately modeled [4].

Financial regulators reached a similar diagnosis: the FSB's June 2026 consultation report on sound practices for AI adoption in finance was explicit that the high levels of autonomy AI agents may exercise "can create or amplify certain risks, which can materialise at great speed," and identified agents' capacity to execute unauthorized or unintended actions faster than institutions can detect or correct them as a defining characteristic of agentic risk, distinct from the risks posed by earlier generations of decision-support AI [6][7]. Writing in ProMarket, policy researcher Piergiuseppe Fortunato framed the underlying regulatory gap directly: financial-system stability requires macroprudential tools that examine system-wide interaction effects, not just microprudential tools that examine whether any single institution or product is individually sound. He illustrates the point directly: "Consider financial agents programmed to reduce risk when volatility rises. Each can be perfectly aligned with its mandate. Yet if thousands react to the same signals at machine speed, they can amplify precisely the volatility they were designed to escape" [5]. The BIS has separately flagged the opacity of circular financing arrangements within the AI industry itself, including chipmakers and hyperscalers taking equity stakes in AI firms that in turn commit to purchasing their compute, as its own pressure point on global financial stability – a related but distinct concentration concern from the deposit-flight dynamic described here [12].

What distinguishes the agentic bank run from a conventional bank run is the removal of two natural circuit breakers that have historically slowed deposit flight: human deliberation time and observable social signaling. A human depositor watching a bank's stock price fall or reading alarming social media posts still has to notice the signal, decide to act, and execute a transfer, a process that, even at its fastest

during the SVB episode, took hours. Today's agents, including Muse, still require the user to approve each transfer, which preserves a checkpoint the SVB-era bank run lacked; but that checkpoint depends on the human continuing to review and approve rate-driven prompts quickly and without much scrutiny, since the agent itself does not need to notice anything unusual to generate the prompt, only the rate-comparison logic it already runs continuously. That checkpoint could narrow further if agents progress toward standing, pre-authorized execution: in that scenario, a transfer could occur as soon as a more attractive account is identified, with no pause for reflection and, absent specific design choices, no visibility to the bank whose deposits are leaving until the outflow has already occurred. Either way, this compresses the response window available to bank treasury and regulatory functions from the days or hours available in prior episodes toward something closer to the frequency of agent-prompted approvals today, and potentially, if agents move toward greater autonomy, the settlement time of the underlying payment rail itself.

Recommendations

Immediate Actions

Financial institutions should begin modeling deposit-concentration exposure to agentic optimization behavior specifically, distinct from existing liquidity stress tests built around historical human withdrawal patterns, since the assumptions underlying those models, including withdrawal speed limits and the role of depositor inertia, no longer hold once a meaningful share of depositors delegate cash management to an agent. Treasury and asset-liability management teams should inventory which of their deposit products are priced closest to the checking-account end of the yield curve, since these are likely to be among the balances an optimization agent would target first and among those most likely to move as a bloc rather than gradually. Banks should also evaluate whether their own digital channels expose real-time balance and rate data through APIs in a form that third-party agents can already query, since that exposure, not any flaw in the agent, is what makes automated rate arbitrage technically possible at scale today.

Short-Term Mitigations

Institutions should engage directly with agent platform providers such as Meta, and with fintech rate-comparison and account-aggregation services, to understand what throttling, batching, or staged-execution mechanisms exist or could be built into agent-driven transfer logic, since smoothing the timing of correlated transfers, even without changing their ultimate volume, would restore some of the response window that human-speed bank runs historically afforded regulators and bank treasuries. Regulators and

industry bodies should prioritize the population-level stress testing approach that policy researchers have proposed, in which regulators simulate how a large simultaneous population of agents, rather than a single representative agent, would respond to a plausible rate-differential or news scenario, since single-agent testing cannot reveal emergent correlation effects that only appear at scale [5]. Deposit insurance and liquidity-backstop frameworks designed around the SVB-era assumption of a days-to-hours withdrawal timeline should be reassessed against a machine-speed scenario in which a comparable share of deposits could move within the settlement window of instant-payment rails.

Strategic Considerations

The agentic bank run is best understood as one instance of a broader pattern this note argues security and risk professionals should expect to recur: as agentic AI adoption scales, individually well-governed agents can still produce collectively destabilizing outcomes, and the unit of risk analysis needs to expand from "is this agent secure and well-behaved" to "what happens when a large population of similarly instructed agents acts at once." This reframing has direct implications beyond banking, including in cybersecurity response automation, where defensive agents reacting to the same threat signal could similarly synchronize into unintended collective action, and in any other domain where large numbers of autonomous agents draw on shared underlying models or shared external data feeds. Financial institutions and their security teams should treat the emerging macroprudential regulatory conversation, exemplified by the FSB's 2026 consultation process and expected final guidance, as directly relevant to enterprise AI governance programs, not as a separate compliance track owned solely by treasury or macro-risk functions, since the underlying technical question, namely how to detect and bound correlated agent behavior before it produces a system-level event, is fundamentally a security and monitoring problem as much as an economic one.

CSA Resource Alignment

CSA's own research has already documented the specific mechanism underlying the agentic bank run scenario in a related context. [AI Compute Concentration and Systemic Risk](#) [4] analyzes how reliance on a small number of shared foundation-model and infrastructure providers can produce algorithmic herding, and models tail-loss amplification of 18% to 54% in financial markets once correlated, model-driven decisions replace independent human ones. The deposit-flight dynamic Slok describes is the same correlation mechanism applied to consumer cash management rather than institutional trading, and institutions evaluating their exposure to an agentic bank run should apply that paper's

concentration-risk framework, specifically its guidance on maintaining multi-provider redundancy and monitoring for shared-dependency exposure, to their own assessment of how many of their depositors' financial agents likely rely on the same handful of underlying models.

CSA's [State of Cloud and AI for Financial Services 2026](#) [13] survey report provides the adoption context that makes this scenario plausible rather than theoretical: it finds that 62% of financial services firms have already deployed AI agents, that 93% of those firms have granted their agents some level of operating autonomy, and that 85% of institutions anticipate autonomous AI-driven financial transactions becoming routine, while governance maturity has not kept pace with that adoption curve. That gap between autonomy and governance is consistent with the conditions under which correlated agent behavior of the kind described in this note could emerge at meaningful scale, though the survey does not measure correlation risk directly.

More broadly, the agentic bank run scenario is an instance of the multi-agent interaction risks CSA's [MAESTRO](#) [14] threat-modeling framework was built to analyze, particularly the framework's treatment of ecosystem-level risks that emerge only when many agents interact with shared external systems rather than risks contained within any single agent's architecture. Institutions building governance programs around this class of risk should map their controls to the AI-agent-relevant domains of CSA's [AI Controls Matrix \(AICM\) v1.1](#) [15], which provides a baseline for evaluating third-party and supply-chain concentration exposure, an increasingly central concern as agentic finance depends on a small number of shared model providers and payment rails.

References

- [1] Investing.com / CNBC. "[Apollo raises specter of an AI agentic 'bank run' hitting financial industry.](#)" CNBC, September 28, 2026.
- [2] CoinDesk. "[AI agents could drain cheap bank deposits, Apollo's Torsten Slok warns.](#)" CoinDesk, September 28, 2026.
- [3] Apollo Global Management. "[Is an Agentic Bank Run Coming?](#)." The Daily Spark, September 2026.
- [4] Cloud Security Alliance AI Safety Initiative. "[AI Compute Concentration and Systemic Risk.](#)" CSA Lab Space, 2026.
- [5] Piergiuseppe Fortunato. "[AI Governance Needs a Macroprudential Turn.](#)" ProMarket, September 18, 2026.
- [6] Financial Stability Board. "[Sound Practices for Responsible Adoption of Artificial Intelligence \(AI\): Consultation Report.](#)" FSB, June 10, 2026.
- [7] PYMNTS. "[Agentic AI Risk Catches Eye of Financial Stability Board.](#)" PYMNTS, 2026.
- [8] about.fb.com. "[Introducing Muse: The World's First Personal AI Agent Built for Everyone.](#)" Meta Newsroom, September 8, 2026.
- [9] Yahoo Finance. "[Meta's Muse says it can manage your money. Should you let it?](#)" Yahoo Finance, September 2026.
- [10] The Motley Fool. "[News About Muse Drove a 1-Day 11% Jump in Meta Stock.](#)" The Motley Fool, September 24, 2026.
- [11] CNN Business. "[SVB collapse was driven by 'the first Twitter-fueled bank run'.](#)" CNN, March 14, 2023.
- [12] Insurance Journal. "[AI 'Arms Race' Still Poses Global Stability Risk, BIS Chief Says.](#)" Insurance Journal, September 11, 2026.
- [13] Cloud Security Alliance. "[State of Cloud and AI for Financial Services 2026.](#)" Cloud Security Alliance, 2026.
- [14] Cloud Security Alliance. "[MAESTRO: Agentic AI Threat Modeling Framework.](#)" CSA Lab Space, 2025.

[15] Cloud Security Alliance. "[AI Controls Matrix \(AICM\) v1.1](#)." Cloud Security Alliance, 2026.

[16] Plaid. "[Plaid powers Meta's new AI agent, Muse](#)." Plaid, September 8, 2026.