IDC

SPOTLIGHT

Sponsored by: Cisco

Agentic AI is shifting regulated risk, fraud, and compliance activities from analytic insight to automated execution, requiring governed infrastructure to ensure consistent, controlled, and defensible outcomes.

From Autonomous AI to Controlled Outcomes in Financial Services

March 2026

Written by: Sam Abadir, Research Director, Risk, Financial Crime, and Compliance

AT A GLANCE

WHAT'S IMPORTANT

As agentic AI moves into execution, financial institutions must reassess whether their AI environments can govern authority, accountability, and cost at scale. Performance alone no longer defines readiness.

KEY TAKEAWAYS

- Agentic AI shifts from insight to execution, requiring control and visibility required to scale.
- Fragmented AI increases execution risk, forcing convergence across functions.
- Secure operating models determine whether autonomy scales control or scales exposure.

Introduction: AI is mature, but functionally fragmented

AI has been embedded in financial services for decades. Risk, compliance, and financial crime functions adopted AI at different times and for different reasons, resulting in distinct models, data pipelines, controls, and decision frameworks across the enterprise.

This fragmentation was manageable as long as AI primarily supported analysis and advisory tasks. Decisions remained human led, and inconsistencies across systems carried limited operational consequences.

As AI begins to execute actions, that tolerance disappears. Fragmented execution environments introduce material operational risk, creating governance gaps, unclear ownership, and inconsistent controls that are increasingly difficult to manage and audit at scale.

Generative AI exposes control and interpretation gaps

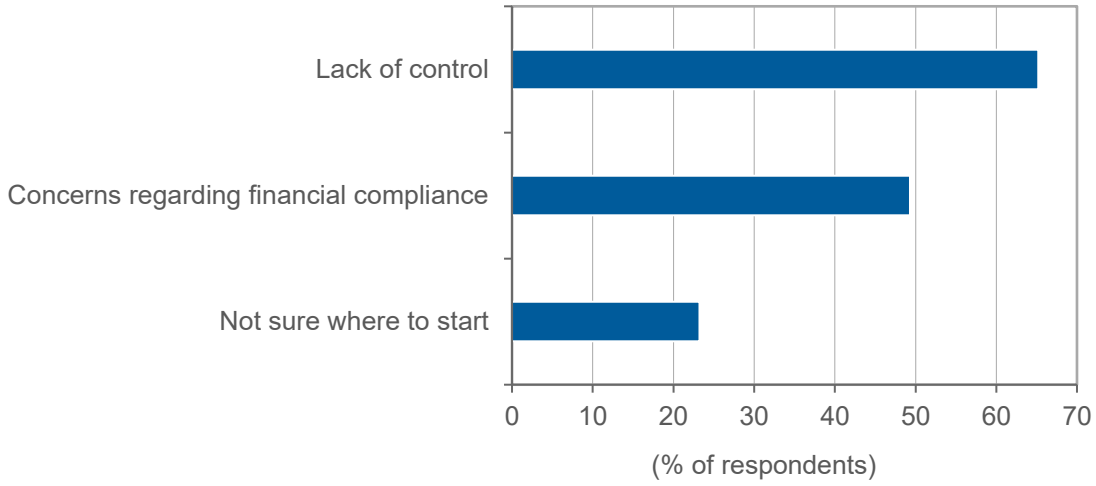
Generative AI (GenAI) expands AI usage beyond prediction into interpretation, explanation, investigation support, and regulatory response. These activities span risk, compliance, and financial crime rather than residing cleanly within a single function.

According to IDC's research, the primary barriers to AI adoption in financial services are not model availability or technical readiness but concerns about control, compliance, and the defensibility of outcomes (see Figure 1). With the independent deployment of generative models, these concerns quickly become operational realities.

FIGURE 1: ***Primary barriers to implementing AI in financial services***

Control uncertainty and compliance risk outweigh technology readiness concerns

Q *What are the biggest barriers to implementing AI technology?*



n = 69

Source: IDC's Office of the CFO Survey, August 2024

The same regulatory text may be interpreted differently across systems. Identical risk or fraud signals may drive inconsistent actions. Control rationales and audit explanations diverge across teams, directly reinforcing buyer concerns around financial compliance, accuracy, and auditability, which IDC's findings highlighted.

Generative AI does not create these gaps; it exposes them and accelerates their impact, making fragmented governance, unclear ownership, and inconsistent controls materially harder to manage as AI scales.

Agentic AI as the convergence catalyst

Agentic AI introduces systems that can plan, execute, and adapt across multiple steps with limited human intervention. In financial services, this shift directly affects the execution of risk, compliance, and financial crime activities, not just how businesses analyze them.

As autonomy increases, execution changes in function-specific ways. Risk teams move from periodic exposure assessment toward continuous limit management and automated responses. Compliance teams shift from retrospective testing to real-time validation of controls and obligations. Financial crime operations increasingly automate investigation flow, escalation, and disposition.

As these activities become autonomous, inconsistencies in authority, data access, and execution logic translate directly into operational and regulatory exposure. What were once localized decisions now produce enterprisewide consequences.

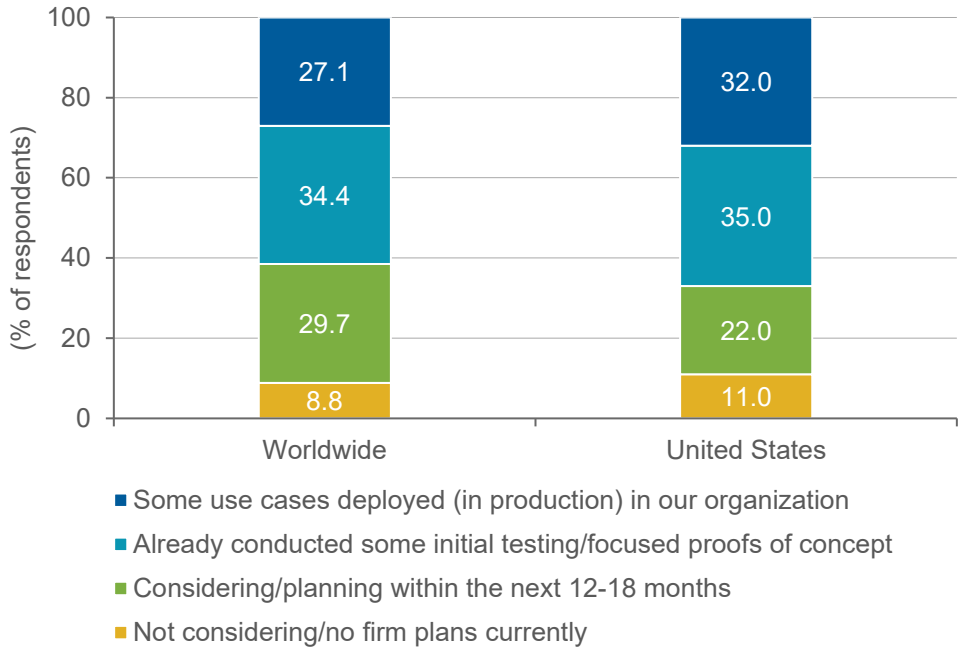
These functions remain distinct lines of business. Convergence does not require organizational consolidation; it requires shared execution logic, shared control frameworks, and aligned outcomes. Agentic AI creates the opportunity to reduce operational risk, improve control adherence, and accelerate execution without increasing ambiguity across functions.

This shift is already underway. IDC's survey data in Figure 2 shows that a majority of financial services institutions are either actively testing agentic AI or have already deployed some use cases in production. As adoption moves beyond experimentation, the risks associated with fragmented execution, inconsistent controls, and unclear authority move from theoretical to operational.

FIGURE 2: ***Financial services institutions advancing agentic AI beyond experimentation***

Agentic AI adoption is moving into execution

Q *Which of the following best describes your organization's plans for agentic AI?*



n = 917

Source: IDC's Industry Insights Survey: Financial Services, 2025

Why autonomy without security scales failure

As autonomy increases, the dominant failure mode shifts away from model accuracy toward loss of visibility. When AI systems execute actions across functions, institutions must be able to determine how decisions are formed, what authority is exercised, and how actions propagate across systems and teams.

Visibility in this context extends beyond explainability or retrospective audit trails. It determines whether institutions retain operational and financial control as AI scales. When execution becomes opaque, minor inefficiencies, duplicated

actions, or misaligned authority compound into material exposure. Institutions may remain compliant at a policy level while losing control of day-to-day operations.

Effective visibility enables consistent outcomes across distributed functions. It supports defensible explanations to regulators and auditors, aligns execution logic even when ownership is fragmented, and provides early indicators when scale begins to erode economic discipline or control effectiveness.

Institutions that use AI strictly for decision support operate under a materially different risk profile than those that permit AI to initiate actions, escalate cases, or enforce controls. As AI moves into execution paths, governance gaps shift from design considerations to operational exposures.

This distinction is increasingly reflected in supervisory expectations across both the United States and Europe. Model risk guidance, such as SR 11- 7, establishes expectations for understanding, validation, and ongoing monitoring of models that influence or drive decisions. In the United States, related supervisory frameworks, including OCC Heightened Standards and interagency operational resilience guidance, extend these principles beyond models to the systems and processes that execute decisions, emphasizing management accountability, effective challenge, and demonstrable control under stress. DORA formalizes similar expectations within the European Union by requiring financial institutions to evidence traceability, resilience, incident response, and recoverability across ICT systems that support critical functions. The EU AI Act further reinforces this trajectory by imposing defined governance, documentation, human oversight, and risk management obligations on high-risk AI systems whose outputs materially affect individuals or markets, particularly where autonomy operates at scale.

Without sufficient visibility, agentic AI scales both productive and defective behavior at the same rate. As execution expands, speed, confidence, and financial impact increase faster than oversight or intervention capacity. In this context, autonomy without security and security without visibility do not counterbalance risk. They compound it.

AI factories as an enterprise operating model decision

As AI systems move from isolated use cases to enterprisewide execution, financial institutions require a repeatable operating model for building, deploying, and governing AI at scale. This need has driven the emergence of AI factories as standardized environments that replace fragmented experimentation and sandboxes with production-grade platforms and infrastructure.

IDC characterizes AI factories as a structural shift from traditional MLOps toward industrial-scale automation. They integrate AI compute, software frameworks, orchestration, and infrastructure into a unified operating model designed to continuously transform data into actionable intelligence. Importantly, AI factories are best understood as an operating model decision rather than a discrete product choice.

For financial services, this foundation is necessary but insufficient. As AI factories scale execution across regulated domains, they also concentrate accountability and amplify control gaps. Performance and efficiency alone do not address access control, auditability, explainability, or operational resilience. Without embedded governance, AI factories do not simply scale intelligence; they scale exposure.

When this operating model decision is deferred or treated as purely technical, institutions often encounter delayed deployments, supervisory resistance, and the growth of unmanaged AI activity outside formal governance or "shadow AI." These outcomes reflect operating model misalignment rather than technology failure.

This gap creates the requirement for a secure AI factory as a control and assurance layer over enterprise AI execution.

Considering Cisco

Cisco Secure AI Factory extends the AI factory operating model by introducing a consistent control and assurance layer over AI execution. Rather than replacing underlying AI platforms, it governs how AI systems are accessed, monitored, and permitted to act across their life cycle. Under these conditions, AI shifts from an innovation asset to regulated infrastructure, subject to enterprise expectations for control, auditability, and operational resilience.

At the foundation, NVIDIA defines the AI factory as an industrialized stack for large-scale AI training and inference, spanning accelerated compute, high-performance networking, and AI life-cycle tooling. This layer is optimized for performance, throughput, and scalability. Its design emphasis is on execution efficiency rather than governance, leaving authorization, oversight, and accountability to surrounding enterprise control systems as AI models and agents transition into production.

Within this context, Cisco positions its Secure AI Factory as an operational control plane that spans AI execution environments. It focuses on secure connectivity, segmented execution domains, identity-aware access, and continuous telemetry across data pipelines, model runtimes, and agent-driven workflows. The Secure AI Factory is positioned to operate across heterogeneous AI stacks, enforcing consistent execution boundaries and policy application without displacing AI development, training, or orchestration platforms.

Together, these layers create an interaction effect that is greater than simple functional alignment. By coupling AI performance infrastructure with a pervasive execution control plane, the operating model shifts from parallel capability stacks to a unified system of governed intelligence production. NVIDIA enables AI at an industrial scale, while Cisco influences how that scale is exercised through consistent enforcement, visibility, and policy coherence across environments. The resulting effect is not merely additive capacity, but a change in how AI-driven activity is introduced, supervised, and defended in production. For risk, compliance, and financial crime functions, this model enables AI execution to be governed as a continuous enterprise process rather than as a collection of isolated deployments.

Use case focus: Agentic AI–enabled enterprise-level control convergence

Risk, compliance, and financial crime remain distinct lines of business at most financial institutions, each with its own mandate, ownership model, and success metrics. Under analytic AI, this separation allowed functions to optimize locally with limited downstream impact. Under agentic AI, that same separation becomes a source of operational tension.

When execution is fragmented, autonomous systems act within function-specific boundaries but produce enterprisewide consequences. A fraud action intended to reduce losses may trigger compliance obligations or have a significant impact on customers. A compliance control designed to satisfy regulatory testing may constrain risk responses. A risk-driven limit adjustment may alter downstream investigation behavior.

Without shared execution foundations, organizations manage these interactions informally, after the fact, or not at all.

The Secure AI Factory enables convergence at the execution layer rather than the organizational layer. It provides consistent guardrails across AI-driven processes, regardless of functional ownership. It uniformly applies authority, data access, model behavior, and enforcement logic, even as accountability remains distributed across teams.

This shifts how institutions operate. Trade-offs become explicit and governed rather than implicit and reactive. Functions retain independence while aligning shared execution standards, enabling consistent outcomes without sacrificing oversight.

Challenges

Cisco faces two material challenges in advancing this position within regulated financial services environments.

The first is how responsibility and dependency are defined across execution and control layers. Because the Secure AI Factory is explicitly layered over an existing AI factory foundation, integration boundaries must be clearly articulated. Without clear ownership of runtime governance, policy enforcement, and exception handling, accountability for AI behavior can become distributed across infrastructure, security, and application teams.

The second challenge is translating technical control capabilities into outcomes that align with regulatory and supervisory expectations. Financial services organizations increasingly look for explicit connections between controls and outcomes such as audit readiness, transparency of automated actions, and operational resilience. Where these linkages are not clearly articulated, the Secure AI Factory risks being interpreted as generalized infrastructure security rather than a control architecture tailored for regulated AI execution.

If Cisco clearly operationalizes its role as the control and assurance layer governing AI execution, these challenges become addressable through design and execution rather than structural constraints. Cisco's historical role as a horizontal control plane across complex, multivendor enterprise environments provides a relevant precedent. Extending this model into AI execution environments positions the Secure AI Factory as a mechanism for maintaining oversight, resilience, and accountability as AI systems become embedded in regulated financial operations.

Conclusion

AI in financial services is moving decisively from insight to execution. Agentic systems amplify both opportunity and risk, not because they are new, but because they operate continuously and at scale. In this environment, autonomy without security accelerates failure as efficiently as it accelerates productivity.

The central question for regulated institutions is no longer whether to adopt agentic AI; it is whether existing operating models can govern execution, accountability, and cost as autonomy expands. Performance-focused AI environments answer only part of the question.

As AI moves from insight to execution, control and visibility determine whether autonomy scales intelligence or scales exposure.

A secure AI factory reframes AI as regulated infrastructure. It provides the shared visibility, enforcement, and control foundations required to support convergence across risk, compliance, and financial crime without forcing organizational consolidation. For institutions seeking to scale agentic AI responsibly, secure execution is not an enhancement; it is a prerequisite.

About the analyst

Sam Abadir, *Research Director, Risk, Financial Crime, and Compliance*

Sam Abadir is research director for IDC Financial Insights, responsible for the risk, financial crime, and compliance practice. Mr. Abadir's core research coverage includes compliance regulators around "bad actors," including AML, eCDD, and KYC, and other regulatory programs such as GDPR and compliance watchlists (e.g., OFAC and PEP). In addition, Mr. Abadir's research will cover different aspects of risk and compliance in financial services and how technology can support mandated and best practices in areas such as DORA, GDPR, and other regulations that impact financial services.

MESSAGE FROM THE SPONSOR

The Cisco Secure AI Factory with NVIDIA is a purpose-built industrial scale environment where compute, networking, storage, security, and software operate as a unified production line for AI use cases and workloads. It's a security-first architecture, embedding protection at every single layer of the stack, going far beyond traditional AI factory approaches.

With joint engineering and reference architectures from both Cisco and NVIDIA, you get a secure, proven, scalable AI foundation that accelerates delivery of real AI value to the business, and new AI experiences to your customers.

Learn more about Cisco Secure AI Factory with NVIDIA here: http://secureaifactory.com

IDC Custom Solutions

The content in this paper was adapted from existing IDC research published on www.idc.com.

IDC Research, Inc.
One Beacon Street
Suite 33100
Boston, MA 02108, USA
T 508.872.8200
F 508.935.4015
blogs.idc.com
www.idc.com