Rocket® software

The Trust Gap: How Regulated Enterprises are Setting the Bar for Agentic AI

What 250 global IT leaders reveal about agentic AI in regulated enterprises

Findings from the June 2026 Mainframe AI Study, conducted by Hanover Research for Rocket Software

Contents

Executive summary

IT leaders in financial services and banking already know AI. Familiarity isn't the barrier anymore. In a June 2026 study of 250 IT infrastructure and operations leaders, 93% reported familiarity with conversational AI and 92% with agentic AI. These leaders are directors and above at organizations with at least $500 million in annual revenue, all of them running mainframes in mission-critical environments. They understand the technology. What they're still working out is how to trust it in production.

That gap between understanding and trust is the real story. Nearly all of these leaders (94%) rank enhancing IT operations with AI as a high or top priority. Yet their initiatives stall most often when scaling across the organization (40%) or during initial production rollout (32%). Pilots work. Production is different. The harder question is whether it can act safely inside systems that process claims, settle transactions, and run overnight batch cycles, where a single mistake carries regulatory and financial weight.

The study points to a clear conclusion. The next phase of enterprise AI won't be decided by capability. It'll be decided by governance. Leaders are moving from conversational tools toward agentic ones, but they're setting a high bar first: strong governance controls, proven accuracy, and security reviews before anything touches production.

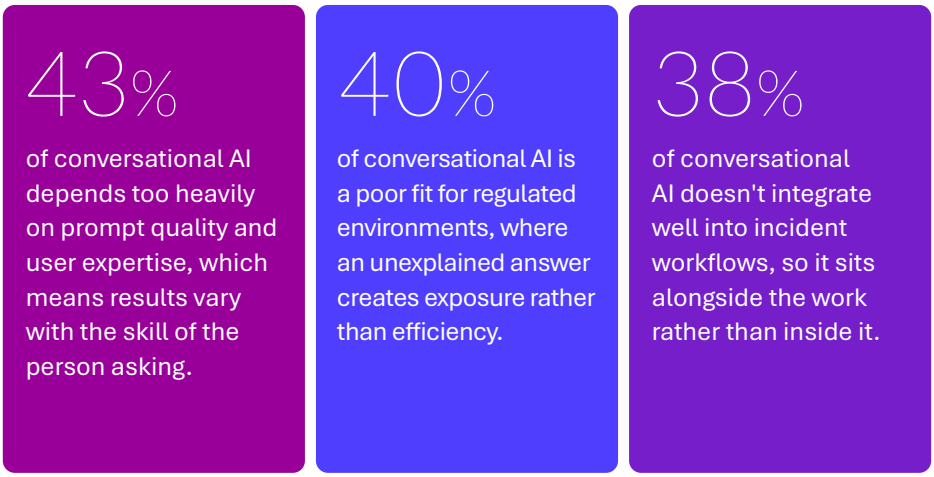
This paper walks through five findings from the study and what they mean for anyone planning to scale AI in a regulated environment. Where it helps, we use Rocket EVA™, Rocket Software's governed agentic AI platform, as an example of what building for these conditions looks like in practice. **The takeaway is simple: trust will determine which enterprises scale agentic AI, and which stay stuck in the pilot phase.**

Finding 01: The shift from conversational to agentic AI is already underway

Conversational AI opened the door. It made large language models feel accessible, and it proved that natural language could sit on top of complex systems. But the study suggests its role in core operations is limited, and financial services leaders are already looking past it.

The numbers describe a transition in motion. While 43% of organizations use a mix of conversational and agentic tools, 34% now primarily use agentic tools. That's a meaningful share moving toward AI that can take action, not just answer questions.

The reasons conversational AI falls short in operational settings are practical, not theoretical. Leaders point to three limitations. First, it depends too heavily on prompt quality and user expertise (43%), which means results vary with the skill of the person asking. Second, it's a poor fit for regulated environments (40%), where an unexplained answer creates exposure rather than efficiency. Third, it doesn't integrate well into incident workflows (38%), so it sits alongside the work rather than inside it.



Agentic AI is emerging to close those gaps, and the study shows where leaders are most willing to deploy it. Half (50%) want it to detect anomalies and correlate signals. Nearly as many want it to proactively identify security risks (48%) or operational risks (42%). These are diagnostic and monitoring tasks at the heart of running core systems, and they reward speed, pattern recognition, and the ability to see across a fragmented environment.

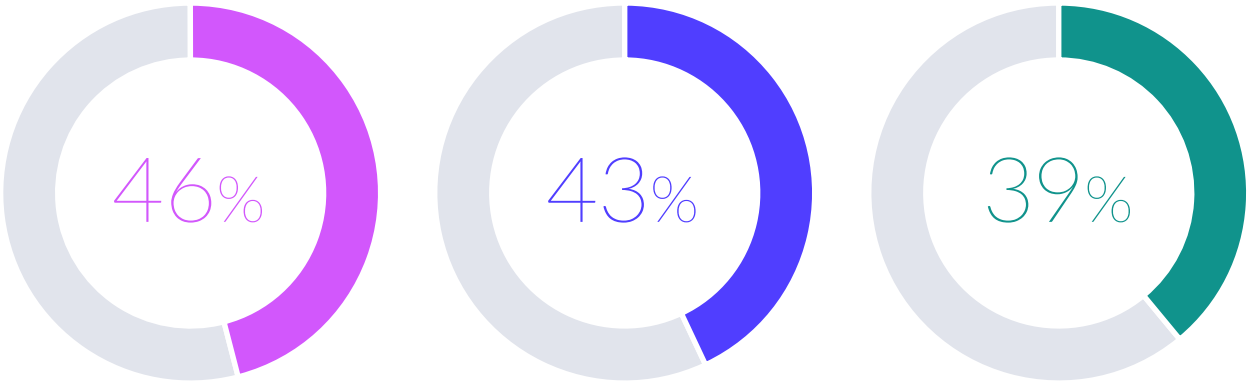
This is the space governed agentic platforms are built for Rocket EVA, for example, is designed to correlate signals across mainframe, distributed, and cloud systems, trace an issue from symptom toward root cause, and support action inside existing incident workflows rather than beside them. The shift the study describes isn't a leap into autonomy. It's a move toward AI that can do specific operational work reliably, and prove it did so within the rules.

Finding 02: Trust, governance, and regulation set the bar for production

If capability were the deciding factor, more of these initiatives would already be in production. They aren't, and the study is direct about why. Regulation shapes nearly everything.

For 78% of respondents, regulatory considerations limit their ability to deploy AI significantly or extremely. That's not a minority concern to manage around. It's the operating reality for more than three-quarters of the leaders surveyed. In financial services and banking, automated decisions face scrutiny on explainability, fairness, and defensibility, and any AI that touches a system of record has to hold up to that scrutiny later.

Because of this, leaders don’t trust AI in production without clear conditions. The top requirement is strong governance controls and approval workflows (46%). Close behind is proven accuracy validated against historical incidents (43%), followed by security reviews (39%). Read together, these paint a precise picture of what "production-ready" means to a regulated enterprise. The AI has to operate within enforced policy, prove it works against real past cases, and pass the same security bar as any other system touching sensitive data.



The study is equally clear about who stays in charge. Nearly half of respondents (48%) permit AI to take limited actions while humans remain accountable overall. Another 34% allow AI to provide insights while humans stay fully accountable. In other words, the vast majority keep a person responsible for outcomes. Autonomy for its own sake isn't what these leaders are buying.

This is where governed agentic AI earns its place. A platform like Rocket EVA is built to enforce policy as code, tie every action to identity and role, and route high-risk steps to human approval before anything runs. Routine diagnostics can proceed within defined limits, but actions that change system state or affect a customer stop and escalate. That model maps almost directly onto what the study says leaders require. The point isn't to remove people. It's to give them faster, better-supported decisions while keeping accountability where it belongs.

Finding 03: Scaling is where ambition meets reality

In the financial services industry, the appetite for AI in IT operations is close to universal. The study found that 94% of leaders view enhancing IT operations with AI as a high or top priority. That level of consensus is rare. The challenge isn't convincing anyone that AI matters. It's getting past the pilot.

The stalling points are specific. Initiatives most often break down when scaling across the organization (40%) or during initial production rollout (32%). These are the two moments when an experiment has to become an operation, when the AI moves from a controlled sandbox into systems that can't afford surprises. That's precisely where the earlier findings come home to roost, because a pilot rarely has to answer for governance, and a production deployment always does.

The barriers behind these stalls are consistent with the rest of the study. The leading obstacle is security, privacy, or regulatory concerns (31%). Next come integration difficulties (29%) and data quality or access limitations (28%). None of these are about whether the model is smart enough. They're about whether the surrounding environment (its controls, its connections, and its data) is ready for the model to act.

That distinction matters for how leaders choose a path forward. An AI tool that requires a rip-and-replace project, or that can't connect to existing identity, service management, and observability systems, runs straight into the integration barrier the study names. A platform designed to work in place, alongside the tools a team already trusts, sidesteps it. Rocket EVA, for instance, is built to integrate with existing IAM, ITSM, observability, and automation systems, and to process data where it already lives, without forcing migration into a new repository.

There's also a clear reward for getting this right. When leaders were asked what justifies increased AI investment, the top answers were reduced compliance risk and easier audits (44%), faster incident detection and response (40%), and reduced outages (39%). Those outcomes describe the payoff of getting past the scaling wall: start with one focused use case, prove it stayed within policy, and expand as confidence grows.

Finding 04: Mainframes are a strategic AI asset, not a liability

There's a persistent assumption that mainframes are obstacles to AI adoption. The financial services leaders closest to these systems see it differently, and the study makes that plain.

More than half of respondents (52%) view mainframes as an AI advantage, citing their reliability, data, and transaction integrity. Only 22% see them as a limitation. That's better than a two-to-one margin in favor of the mainframe as an asset. It makes sense when you consider what these systems hold: decades of high-integrity transaction data and the workloads a financial institution can't afford to break. That data is exactly what makes AI diagnostics valuable, and the mainframe's reliability is exactly the quality regulated operations depend on.

Financial services leaders are acting on that view. The study found that 42% partially include mainframes in their AI strategy for select use cases, and 32% fully include them as primary targets for AI value. Combined, that's nearly three-quarters treating the mainframe as part of their AI future rather than a system to work around.

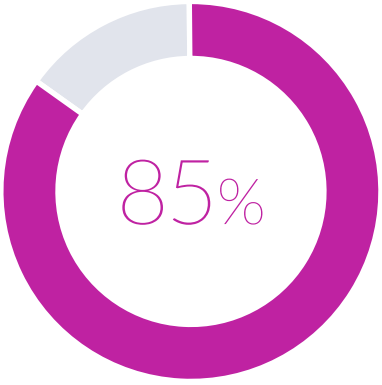
There's also a workforce dimension worth noting. A significant 81% of organizations report a very or extremely significant mainframe skills gap, and 87% believe AI will help address it over the next two years. That's a real and pressing concern. But the skills gap alone doesn't capture the full opportunity. Even experienced teams working at full capacity can't hold the entire operational context of a complex, multi-platform environment in view at once. When a transaction spans a mainframe job, a distributed application, a cloud service, and an integration layer, the signals that matter are scattered across systems that don't naturally talk to each other. No amount of expertise makes that correlation fast or complete without the right tools.

Rocket EVA is designed with mainframe environments in mind, using a context graph to link applications, jobs, transactions, infrastructure, and services so an issue can be traced from symptom toward root cause across the full environment. That correlation produces diagnostic visibility that wouldn't exist otherwise. When a transaction spans a mainframe job, a distributed application, a cloud service, and an integration layer, the signals that matter are scattered in ways that no single team, however experienced, can connect quickly or completely without the right tools. EVA makes those connections, surfacing patterns and insights that are genuinely new, not just faster versions of what a skilled person would find manually. As a secondary benefit, that same capability helps extend the reach of less experienced staff, guiding them through diagnostic paths that once required a senior expert. Every action stays governed and traceable throughout.

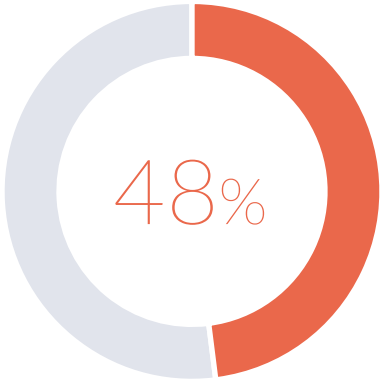
Finding 05: Architecture and deployment reflect a preference for control

How financial services leaders want to run AI reveals as much as whether they want to. On both deployment environment and model strategy, the study shows a consistent instinct: keep control, keep options open, and don't concentrate risk.

Deployment preferences are notably split, and the reason isn't hard to read. Private cloud leads at 24%, followed closely by public cloud (20%), SaaS (20%), and on-premises on the mainframe (18%). No single environment dominates because no single environment describes how these organizations actually run. That reality shows up even more directly in a separate finding: 85% of respondents cited the ability to operate in hybrid environments as one of their top two considerations when adopting AI for IT and security. That's not a preference. It's a requirement. Leaders want AI that can meet their data and workloads wherever they sit, across mainframe, distributed, and cloud systems at once, rather than a tool that works well in one layer and ignores the rest.



Respondents cited the ability to operate in hybrid environments as one of their top two considerations when adopting AI for IT and security.



Respondents favor a hybrid approach that combines cloud-based large language models with internally managed or hosted models.

Model strategy points the same direction. Nearly half (48%) favor a hybrid approach that combines cloud-based large language models with internally managed or hosted models. That's a deliberate balance, pairing the reach of external models with the control of internal ones. It reflects a refusal to hand core operations entirely to a black box.

The concerns behind these choices are, once again, about trust. The top worries about agentic AI in regulated environments are regulatory non-compliance risk (36%), data privacy and residency (35%), and a lack of explainability or auditability (31%). Every one of these is a governance concern, not a capability concern. Leaders aren't asking whether the AI is powerful. They're asking whether they can prove it behaved.

A platform built for these conditions has to respect all three. That means processing data in place so privacy and residency stay intact, enforcing policy so compliance holds, and logging every meaningful action so explainability and auditability are available on demand. Rocket EVA is designed to work across hybrid environments, keep data where it lives, and produce a clear record of what happened, when, why, and who approved it. In an estate where no two systems sit in the same place, that combination of reach and traceability is what makes agentic AI safe to adopt.

Conclusion

Taken together, the five findings tell a coherent story. These leaders know AI well. They want it badly, with 94% making it a priority. They see their mainframes as an advantage, not an obstacle. And they're already shifting from conversational tools toward agentic ones. By every measure of intent, the industry is ready to move.

What holds it back is the same thread running through every finding: trust. Regulation limits deployment for 78% of respondents. Initiatives stall at scaling and rollout. The top production requirements are governance, proven accuracy, and security. The leading concerns about agentic AI are compliance, data residency, and auditability. Read across the whole study, the message is unmistakable. The organizations that scale AI won't be the ones with the most advanced models. They'll be the ones who can prove their AI acted safely, stayed within policy, and left a record to show it.

That reframes what a successful AI investment looks like. The goal isn't autonomy. The study shows leaders don't even want it, with 82% keeping humans accountable for outcomes. The goal is trusted execution: AI that supports faster diagnosis and safer action inside the systems an enterprise can't afford to break, without loosening the controls that keep those systems defensible.

This is the standard governed agentic AI is built to meet, and it's the standard Rocket EVA was designed for. Not to move fastest, but to help regulated enterprises move safely, repeatedly, and at scale, with proof at every step. That means operating across mainframe, distributed, and cloud environments as a single, coherent diagnostic layer — the kind of hybrid reach that 85% of leaders identified as a top-two requirement. And it means surfacing the connections and patterns that exist within those environments, but that no team, however skilled, could consistently see and act on without the right tools. The leaders in this study have already set the bar. The next phase of enterprise AI belongs to whoever can clear it.

Rocket EVA is built for the reality this study describes: regulated environments that need AI to operate across complex systems with governance, visibility, and control. For teams looking to turn these findings into a practical next step, EVA offers a grounded example of what trusted agentic AI can look like in production.

Learn more about EVA

This whitepaper draws on the June 2026 Mainframe AI Study, conducted by Hanover Research for Rocket Software, surveying 250 IT infrastructure and operations leaders (directors and above) in financial services and banking across the US and Europe (UK, Germany, France, and the Netherlands), at organizations with at least $500 million in annual revenue that operate mainframes. It's intended to help IT, operations, and compliance leaders evaluate what trusted agentic AI requires and how a governed platform like Rocket EVA fits within their operating model. Specific outcomes depend on each organization's systems, controls, and deployment approach.

About Rocket Software

Rocket Software is a global technology leader in modernization and a partner of choice that empowers the world's leading businesses on their modernization journeys, spanning core systems to the cloud. Trusted by over 12,500 customers and 750 partners, and with more than 3,200 global employees, Rocket Software enables customers to maximize their data, applications, and infrastructure to deliver critical services that power our modern world. Rocket Software is a privately held U.S. corporation headquartered in the Boston area with centers of excellence strategically located throughout North America, Europe, Asia and Australia. Rocket Software is a portfolio company of Bain Capital Private Equity.

Rocket® software

Modernization. **Without Disruption.**™

Visit RocketSoftware.com

MAR-18636_WP_EVAHanoverResearch_V3