Sovereign AI in Financial Services: Taking Back Control of the Future

Owning the Intelligence Layer: Why Control, Alignment and Assurance Matter

aveni

01 Executive Summary

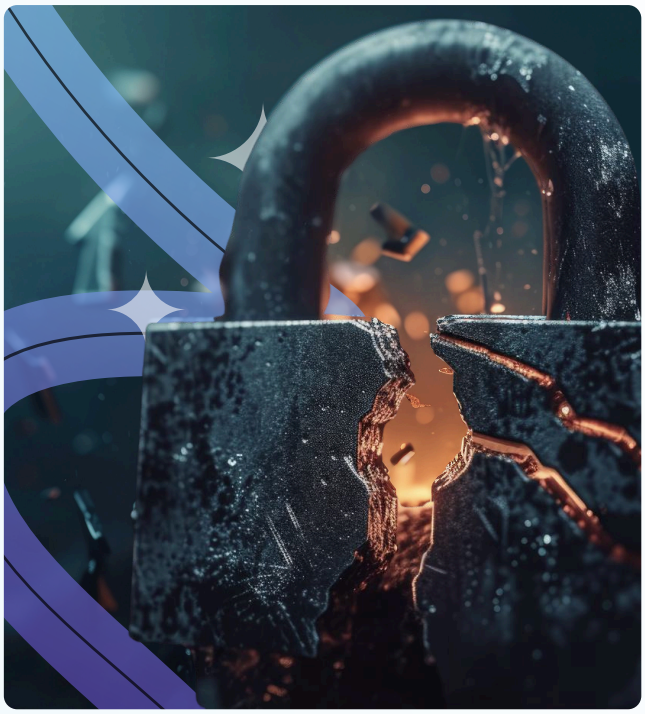
Over the course of 2025, the business world has watched deeply unsettling headlines of AI-driven CEO deepfake scams[1] tricking employees into transferring funds, increasingly sophisticated phishing campaigns, and supply-chain attacks impacting core systems. These are real-world wake-up calls costing firms millions and supply-chain breaches affecting nearly 60% of UK financial institutions[2].

The State of AI in Business 2025 Report[3] shows a stark truth: **while 80% of companies have piloted generative AI, only 5% have achieved measurable business impact**, a 'GenAI Divide' driven not by lack of infrastructure but by tools that can't embed into real workflows. In financial services, copilots and productivity add-ons may draft content, but they don't shift outcomes where it matters: affordability checks, complaint triage, conduct monitoring, regulated documents. When AI moves from drafting to *doing*, governance must rise to meet the oversight this demands. This is why **sovereign AI is critical**: without control over how AI is built, where it runs, and how its actions are assured, firms risk repeating the cycle of failed pilots, only this time with consequences for customers, compliance, and systemic trust.

Sovereign AI is not a buzzword; it is a practical, imperative shift in how financial services businesses use AI.

Aveni's whitepaper asserts that sovereignty means more than localisation. It means responsible ownership. It's about aligning models to financial services-specific language, regulation and risk; deploying them in environments within an organisation's control; and embedding assurance into every part of the system so every decision is explainable and auditable.

Recent industry research backs this urgency. IBM's latest breach report reveals that **13% of organisations suffered AI-related security incidents**[4], many due to lack of access controls. Shadow AI, which is the unauthorised tools used by employees accounted for **20% of AI-related breaches**[5], and these incidents cost firms hundreds of thousands more per event. Meanwhile, regulators and compliance leaders are demanding visibility, accountability, and human oversight, especially amid surging AI adoption with scant governance in place.



[1] https://www.wsj.com/articles/ai-drives-rise-in-ceo-impersonator-scams-2bd675c4?
[2] https://www.fnlondon.com/articles/the-rush-to-ai-in-the-financial-sector-risks-more-data-breaches-7dd577d6?
[3] https://fortune.com/2025/08/18/mit-report-95-percent-generative-ai-pilots-at-companies-failing-cfo/
[4] https://technologymagazine.com/news/ibms-cost-of-a-data-breach-report-on-risks-of-ai-breaches?
[5] https://www.fm-magazine.com/news/2025/aug/shadow-ai-emerges-as-significant-cybersecurity-threat/

aveni

What this paper sets out isn't another blueprint for digital transformation, it's a challenge to the industry. We can no longer afford to treat AI as a shiny add-on or a black box we hope will behave itself. In financial services, every model decision has a consequence for a customer, a balance sheet, and the system's stability. That means sovereignty and assurance are not nice-to-haves, they are the difference between AI that strengthens our institutions and AI that undermines them.

The firms that act now will shape the standards others follow. Those that don't will find themselves explaining to regulators, shareholders and clients why they ceded control of the very intelligence that runs their business. Sovereign AI is not simply a safer path, it is the only path to long-term resilience, competitiveness, and trust in this industry.

Joseph Twigg
CEO, Aveni

02 The State of AI in Financial Services

The last 18 months have seen financial services experiment with general-purpose AI tools, from copilots in productivity software to off-the-shelf chat interfaces. These pilots have largely centred on **content generation**, summarising notes, drafting emails, or pulling insights from datasets. While useful, these applications sit at the edges of the business.

The real inflection point now emerging is the move from **copilots to agentic AI**. Instead of simply suggesting or drafting, **AI agents execute tasks end-to-end:** checking affordability, collating evidence, monitoring conduct, or even drafting and sending regulated communications. This shift brings opportunity: faster processes, lower costs, and new customer experiences. But it also changes the risk profile fundamentally.

With copilots, weak governance meant poor drafts or hallucinated summaries. With agents, weak governance could mean mis-sold products, flawed compliance checks, or exposure to fraud. In other words, when AI starts doing the doing, the stakes for control, assurance, and explainability rise dramatically.

The regulators are already responding. The **FCA and Bank of England's joint survey (2024)**[6] found that while **over half of UK financial firms are already using AI**, many lack robust governance frameworks and only a minority have defined accountability at senior levels (FCA report). In 2025, the **FCA partnered with NVIDIA to establish an AI sandbox**[7], explicitly recognising the need for controlled experimentation in safety-critical use cases. The FCA's **AI update** reinforced the Government's five regulatory principles: safety, transparency, fairness, accountability, and contestability, and called for industry-wide alignment.

The FCA has also launched an **AI Lab**[8] to engage firms, technologists, and regulators in testing real-world AI

[6] https://www.fca.org.uk/publications/research-notes/ai-uk-financial-services
[7] https://www.reuters.com/business/finance/uk-financial-regulator-partners-with-nvidia-ai-sandbox-2025-06-09/
[8] https://www.fca.org.uk/firms/innovation/ai-lab

deployments. These steps show a clear regulatory direction: experimentation is encouraged, but only under frameworks that prove safety and accountability.

For financial services, this means the next phase of AI adoption cannot rely on generic copilots or unchecked third-party models. As AI systems begin to act as agents inside critical customer and control workflows, firms will need sovereignty over models and a higher bar of assurance: every task registered, every decision explainable, every risk actively monitored.

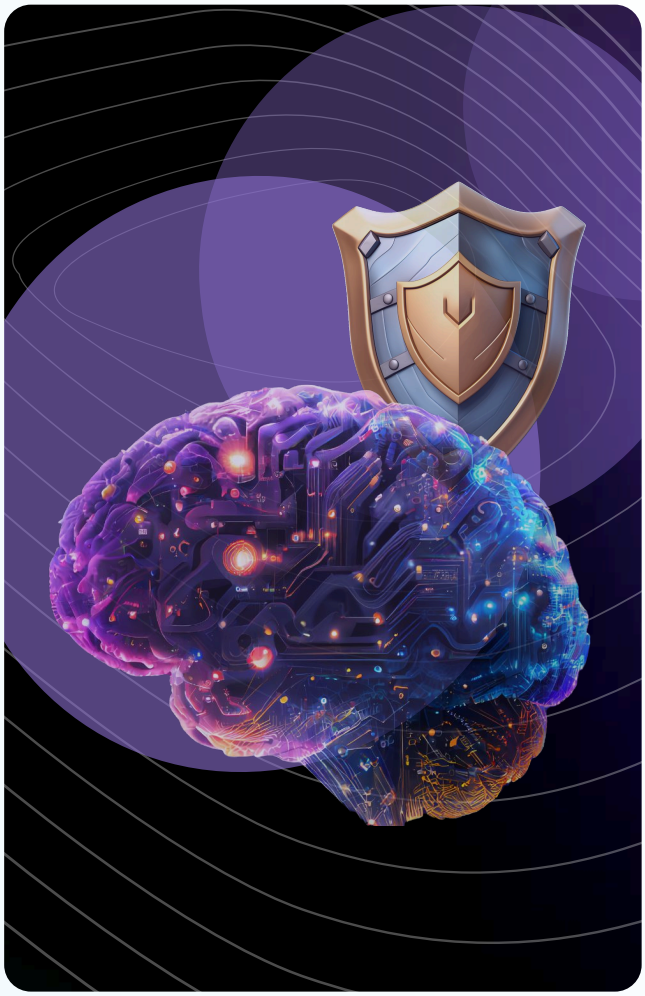
03 Defining Sovereign AI for FS

When most people hear the term 'sovereign AI', they picture national strategies or geopolitical independence from Silicon Valley and Shenzhen. For financial services, sovereignty is something more practical, more immediate, and arguably more urgent. It is about control, control over the models that make decisions, the data that fuels them, the assurance that governs their behaviour, and the environments in which they run.

Why does this matter? Because in finance, AI is not just a productivity tool. It is becoming part of the infrastructure of decision-making. Whether it is assessing affordability, detecting fraud signals, or monitoring conduct, AI agents are moving into roles that do the work. If firms cannot show ownership of the full lifecycle: model selection, training data, deployment environment, and assurance framework, then they cannot credibly claim to regulators, auditors, or customers that they are in control of their business.

This is what distinguishes **sovereign AI** from generic AI adoption:

Models: General-purpose large language models (LLMs) such as ChatGPT or Bard for example may generate fluent language, but they are not aligned to FS regulation, products, or terminology. Sovereignty means deploying **domain-tuned models** like Aveni's FinLLM[9], or small language models (SLMs) shown by NVIDIA to be better suited for task-specific, efficient, and governable agentic systems[10].



[9] https://labs.aveni.ai/finllm/
[10] https://arxiv.org/abs/2506.02153?

Data: FS firms operate under strict regimes (GDPR, FCA Consumer Duty, Basel requirements). Sovereignty means keeping sensitive data inside controlled environments, with the ability to trace how it was used in training and inference. **The FCA & Bank of England's joint survey (2024)**[11] highlighted that less than 20% of firms have robust monitoring of AI data flows, exposing critical governance gaps.

Assurance: Sovereignty without assurance is illusionary. Owning a model is meaningless if outputs are unaudited or misaligned with regulation. The FCA's **AI Update (2024)**[12] stressed the need for accountability and contestability, firms must be able to prove their AI works as intended, and regulators must be able to interrogate its decisions.

Deployment: Finally, sovereignty means freedom to deploy AI **where and how you choose:** in-house, in-cloud, or in-region, without dependency on opaque black-box APIs whose terms can shift overnight. This principle was underscored when the FCA partnered with NVIDIA to build an AI sandbox in 2025[13], signalling that regulators expect firms to host models in environments they can fully govern.

In short, **Sovereign AI in FS is not about flags or borders, it is about accountability**. It's about ensuring that when AI agents act, they act under rules you own, on data you control, within systems you can audit, and under assurance you can prove. Without that, sovereignty is just a slogan. With it, it becomes the foundation of safe, resilient, and competitive financial services.

04 Why FS Needs Sovereign AI: a new standard

It's tempting to treat 'explainability', 'auditability', and 'fairness' as abstract principles. But regulators are not talking in abstractions anymore, they're tying AI directly into the **obligations firms already know**: Consumer Duty, SM&CR, operational resilience, and risk governance.

What's different, and often missed, is the shift in where responsibility sits. For the FCA, the emphasis was not on novel AI laws but on folding AI into the existing scaffolding of accountability. That means if an agent drafts a suitability report that misrepresents a client's risk appetite, the accountability doesn't evaporate into the cloud provider or the vendor. It lands, directly, with the firm, and potentially with a named Senior Manager under SM&CR.

This is why reliance on third-party AI is particularly fraught. It isn't just a matter of whether the model performs well. It isn't just a matter of whether the model performs well. It's about whether you can **stand behind its decisions when a regulator asks for evidence**. The FCA and Bank of England's joint survey found most firms couldn't identify a senior individual responsible for AI oversight. In any other regulated activity, that would already be unacceptable.

[11] https://www.fca.org.uk/publications/research-notes/ai-uk-financial-services
[12] https://www.fca.org.uk/publication/corporate/ai-update.pdf
[13] https://www.fca.org.uk/news/speeches/supercharging-digital-sandbox-collaborating-nvidia-accelerate-ai-innovation

aveni

The unspoken regulatory message is clear: *if you can't explain it, you can't use it*. Not because regulators are anti-AI, but because they know AI is no longer just doing the low-value admin, it's shaping outcomes that affect customers, capital, and market integrity.

The bar, then, is not innovation versus compliance. It is innovation that is **auditable by design**: every AI-enabled action must leave a trail that can be challenged, defended, and aligned with obligations that already exist.

05 The Rise of Smaller, Vertical Models

The first wave of enterprise AI was built on massive, general-purpose models trained to generate language about anything and everything. Impressive, yes, but not optimised for the realities of financial services. When accuracy is crucial, and explainability is a regulatory requirement, a model that 'sounds fluent' is not the same as a model that understands finance.

A different path is now opening up. Research led by NVIDIA in 2025[14] made the case that small language models (SLMs) are better suited for agentic systems, where AI is not just producing content but executing tasks. They argued that SLMs offer greater efficiency, lower cost to run, and crucially, more control and adaptability. In other words, the future of AI agents is. not bigger, but smarter, narrower, and more governable.

This is exactly the philosophy behind FinLLM[15]. Instead of training models on internet-scale text with no regard for context, FinLLM is tuned specifically for the language, regulation, and products of financial services. That means when it analyses affordability, reviews a call for conduct, or drafts a suitability letter, it draws on a model designed to interpret the nuances of FS, not guess at them.

Smaller, vertical models also create a more sustainable foundation for AI in FS. They can be deployed in private or regional environments, retrained on firm-specific data, and benchmarked against regulatory expectations. They reduce dependency on third-party APIs while giving firms a degree of sovereignty over the intelligence driving their most sensitive operations.

The broader AI industry may continue chasing ever-larger foundation models. But in financial services, where the stakes are higher, the future will belong to vertical AI, compact, domain-specialised, assured models like those in the FinLLM suite that are built for tasks that matter.

[14] https://arxiv.org/abs/2506.02153
[15] https://labs.aveni.ai/finllm/

06 Assurance as the Core of Sovereign AI

Sovereignty without assurance is a hollow promise. Owning a model or running it in a local cloud is not enough if you cannot prove that its behaviour is consistent, fair, and aligned with regulatory expectations. In financial services, where AI agents are increasingly trusted to carry out real tasks, assurance becomes the defining ingredient of trust.

The call for assurance is surfacing across regulatory and academic discourse. The Bank for International Settlements[16] has argued that as AI systems move deeper into financial infrastructure, firms must build *"robust validation, monitoring and governance frameworks"* to ensure models remain aligned to their intended purpose. Similarly, **the Department for Science, Innovation & Technology** has emphasised that assurance mechanisms are essential if AI is to be deployed responsibly in high-stakes industries[17]. Even the **European Banking Authority (EBA)** has highlighted the importance of auditability and independent oversight as core requirements for AI in finance[18].

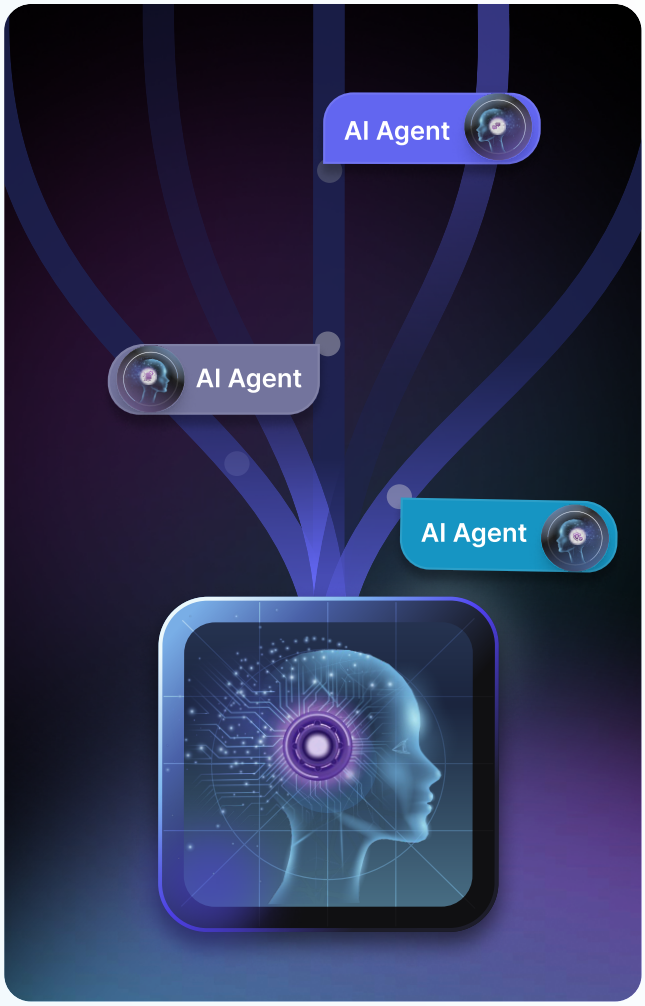
This is the gap assurance fills: turning AI from a promising tool into a governable system. For financial services, assurance means that when an agent takes an action that action is explainable, monitored, and independently verifiable.

This raises an important question: what would an assurance framework for AI in financial services actually look like?

One possibility could be a **Task Registry**, a catalogue defining what each agent is permitted to do, the data it can access, and the policies it must follow, ensuring that no action takes place outside an approved and transparent register.

On top of this, an **AI agent responsible for auditing other AI agents** would provide continuous monitoring and alerts, tracking agent performance against agreed policies and surfacing deviations before they create risk.

To strengthen oversight further, firms could employ an independent LLM-driven system that reviews and scores outputs for fairness, correctness, and compliance, offering a "second opinion" on agent behaviour.



[16] https://www.bis.org/publ/bcbs_nl27.htm
[17] https://www.gov.uk/government/publications/assuring-a-responsible-future-for-ai/assuring-a-responsible-future-for-ai
[18] https://www.eba.europa.eu/publications-and-media/publications/special-topic-artificial-intelligence

aveni

Taken together, these elements point towards a future in which the financial services industry has the equivalent of an **auditor for AI agents.** If AI is to graduate from pilots and experiments to the core of business-critical infrastructure, it will not be enough for systems to be sovereign; they will need to be **assured**, every action traceable, testable, and trusted.

07 From Possibility to Responsibility

AI in financial services is moving fast, from copilots that draft to agents that act. That shift changes the stakes. When AI starts making decisions that touch customers, compliance, and capital, the industry can no longer afford experiments built on generic, opaque systems.

The way forward is clear: **AI in FS must be sovereign and assured**. Sovereign, so firms retain control of the models, data, and environments they depend on. Assured, so every action is monitored, tested, and ready to stand up to regulator and client scrutiny.

Whilst some might be concerned that it's slowing innovation, Aveni believes it's about making it safe to scale. The firms that embrace this approach will set the pace, define the standards, and earn lasting trust. Those that don't will be left explaining why they handed over their core intelligence to systems they could neither see nor govern.

Aveni invites FS firms and regulators to join in shaping sovereign AI standards, through collaboration, research, open assurance frameworks, and shared benchmarks: hello@aveni.ai

aveni

AI that’s safe, ethical and built for financial services. Start the conversation today.

hello@aveni.ai

aveni