Lowenstein Sandler

Client Alert

Data Privacy, Security, Safety & Risk Management

February 24, 2026

Financial Services AI Risk Management Framework: Operationalizing the 230 Control Objectives Before the Market Wakes Up

By Amy S. Mushahwar and Chloe Rippe

At the end of 2025, we wrote that the defining question for 2026 would no longer be whether artificial intelligence (AI) is being used responsibly but whether organizations truly understand and govern the systems they have already built, now that AI has removed their margin for error.

In January 2026, we described the accumulated governance debt inside financial institutions: information sprawl, identity fragmentation, opaque model lifecycles, and infrastructure never designed for AI velocity.
This is the financial services sector's '1999 problem.' Legacy technology stacks, siloed data governance, and inconsistent identity resolution, once manageable, are now tasked with supporting AI-driven decisioning in highly regulated environments.

On February 19, the U.S. Department of the Treasury responded with a sector-specific operational framework.

The Financial Services AI Risk Management Framework (FS AI RMF) is not another governance checklist. It is a remediation blueprint, designed for a highly regulated sector, subject to the process of examination. And while federal government oversight may be less aggressive, state regulators look to guidance like this to understand emerging best practices.

Developed in coordination with more than 100 financial institutions, the Financial Services Sector Coordinating Council (FSSCC), and the Cyber Risk Institute (CRI), the framework introduces 230 control objectives across governance, data, model development, validation, monitoring, third-party risk, and consumer protection. These are not abstract principles. Many controls map to specific system behaviors, ownership assignments, and evidence artifacts expected to withstand audit and supervisory review. Treasury simultaneously released the AI Lexicon to standardize terminology across institutions, regulators, and vendors.

Early commentary has treated the framework as compliance guidance. That interpretation understates its impact. The FS AI RMF functions as an operational architecture standard, and it forces institutions to confront what we have called the '1999 problem' before supervisory expectations harden, as an information governance engineering exercise.

This framework will matter most to boards, risk committees, CISOs, chief data officers, and legal teams responsible for examination readiness, not to just AI policy owners.

What the FS AI RMF[1] Actually Is

The framework includes four primary components:

- AI Adoption Stage Questionnaire – A maturity-based self-assessment aligning control expectations to institutional AI deployment levels.
- Risk and Control Matrix (RCM) – The core engine containing 230 mapped control objectives translating National Institute of Standards and Technology (NIST) AI RMF principles into operational controls.
- Guidebook – Implementation guidance addressing how controls should be operationalized.
- Control Objective Reference Guide – A 400+ page document of evidence examples supporting audit and supervisory review. Audit professionals should pay close attention to this.

A Quick Start Guide supports early-stage adopters. All materials are hosted by the Cyber Risk Institute, and links to the materials are provided above.

Importantly, the framework is designed to integrate with existing governance programs, including NIST Cybersecurity Framework (CSF), enterprise risk management, and SOC 2, rather than fragment them, seemingly taking a page from security standard harmonization over time to reinforce and support complimentary standards. This interoperability addresses one of governance's most persistent failure modes: overlapping frameworks that dilute accountability and obscure control ownership.

I. Operationalizing the 230 Control Objectives

The central question is not whether institutions understand the controls. It is whether they can map them to their technology stack. Operationalization must occur across four infrastructure layers, and it must occur upstream, embedding controls directly into CI/CD[2] and MLOps[3] processes rather than retrofitting them after deployment.

Data Layer

Controls address lineage tracking, feature store governance, training data documentation, de-identification, and cross-border data flows. Effective implementation requires automated lineage logging, dataset version control, and rights-signal propagation at ingestion. Data sensitivity tagging must occur at intake, so training sets are privacy-aware by design.

Model Lifecycle Layer

Control objectives require documented development, bias testing, validation independence, drift detection, and explainability thresholds. These controls must be embedded in MLOps pipelines. Models should undergo automated re-identification regression testing before deployment to ensure sensitive client data has not been memorized.

Identity and Access Layer

The framework addresses human and nonhuman identity management, role-based access controls, decision-path auditability, and cross-system logging. Identity graphs must align with authorization architecture. Recognition capability without consistent rights enforcement creates governance exposure.

Third-Party and API Layer

Vendor transparency, documentation exchange, audit rights, and incident triggers are explicit. Institutions leveraging foundation models must treat vendor artifacts as machine-readable compliance inputs, not static PDFs. Operationalization is pipeline integration, not policy drafting.

Examination Reality: Control Mapping and Evidence Expectations

The FS AI RMF will not remain theoretical guidance. In supervisory environments, frameworks become examination scaffolding, just like the Federal Financial Institutions Examination Council (FFIEC) standards are to security and IT environments in financial services.

Financial institutions are already familiar with spreadsheet-driven examinations structured around mapped control taxonomies. Examiners request control owners, test evidence, assess maturity, and track remediation. AI governance should be expected to follow the same structured mapping approach.

In that environment, narrative explanations will not suffice.

Examiners will ask:

- Where is the control implemented?
- Who owns it?

- What system enforces it?
- What evidence demonstrates effectiveness?
- How is drift monitored?
- How are third-party dependencies tested?

If the answer resides in a policy binder rather than in a system log, defensibility weakens.

The 230 control objectives are structured in a format conducive to supervisory mapping. Institutions that proactively align architecture, ownership, and evidence artifacts to this framework will be materially better positioned than those encountering the mapping exercise for the first time during an examination. AI governance maturity will be assessed through traceability. Architectural gaps are easier to remediate before they appear in a supervisory spreadsheet.

II. Bridging the Privacy Engineering Gap – These Standards Are FS AI Privacy by Design

Traditional approaches frame privacy as documentation: Conduct a Data Protection Impact Assessment (DPIA), update policies, revise notices. The FS AI RMF reframes privacy as infrastructure. Lifecycle accountability must be embedded at ingestion, feature engineering, training, validation, deployment, monitoring, and decommissioning. Recognition architecture must align with rights architecture.

In financial services, this directly affects underwriting, fraud detection, anti-money laundering monitoring, and personalization systems. Controls bolted on post-deployment will not meet operational transparency expectations. Regulators will not ask whether a policy exists – and will ask for logs and dashboarding, like they are now doing with security enforcement. Bear in mind the 400+ Control Objective Reference Guide that provides examples of effective evidence of compliance.

The 2026 update to the NIST Privacy Framework reinforces this operational posture, emphasizing Privacy-Enhancing Technologies (PETs) such as differential privacy, synthetic data, homomorphic encryption, and federated learning. These are production-level tools that make scalable compliance possible, though institutions must balance latency, accuracy, and infrastructure cost implications when deploying them.

The AI Lexicon complements this effort by standardizing terminology across legal, risk, engineering, and product teams, reducing translation friction that often undermines governance execution. Again, this is like the definitional harmonization that has occurred in the security industry as well.

The Machine Unlearning Imperative: Right-to-Be-Forgotten Compliance as Architecture

The framework also surfaces a rarely priced risk: model destruction. Regulators may order algorithmic disgorgement where models are trained on improperly sourced data. For institutions relying on AI-driven underwriting or fraud detection, loss of a core model presents material operational disruption.

Machine unlearning, the ability to remove specific data from trained systems without retraining from scratch, becomes an architectural safeguard. This requires modular or sharded model design from inception. Retrofitting such architecture post-deployment is often infeasible. Training data documentation, lineage tracking, and development artifacts are not academic exercises. They determine whether remediation is survivable.

Disgorgement risk is an architecture question. By the time it becomes a legal response issue, design options are limited, requiring expensive, time-consuming, and perhaps feature-limiting re-architecture.

III. Consumer Protection Implications

The FS AI RMF is fundamentally a consumer protection framework that must be embedded within architectural controls, consisting of considerations for:

- Fairness and Bias Mitigation – Structured bias testing, drift monitoring, and adverse impact analysis intersect directly with Equal Credit Opportunity Act (ECOA), Fair Credit Reporting Act (FCRA), Unfair, Deceptive, or Abusive Acts or Practices (UDAAP), and supervisory model risk expectations.
- Explainability and Transparency – Institutions must define explainability thresholds proportional to risk and establish defensible documentation and communication protocols.
- Accountability and Remediation – Clear ownership, escalation pathways, and traceability determine whether consumer remediation is operationally viable.

Without system-level logging and documentation, policy commitments cannot be operationalized.

IV. Solving the '1999 Problem'

AI governance failures are infrastructure failures running at accelerated speed.

The FS AI RMF forces institutions to confront legacy data sprawl, shadow model development, fragmented ownership, and inconsistent logging practices. It provides boards with an external benchmark for assessing AI governance posture.
The framework does not create new risk. It exposes deferred risk.

Institutions treating this as a documentation refresh will struggle. Those viewing it as architectural modernization will lead.

Conclusion

The Financial Services AI Risk Management Framework is not a checklist. It is a systems blueprint.

The 230 control objectives demand that privacy, fairness, transparency, and accountability move from policy binders into infrastructure. Institutions should map controls to architecture, identify where controls exist only on paper, align identity and rights enforcement, embed lifecycle accountability, inventory third-party model dependencies, and prepare for examination-based control mapping now, before supervisory expectations mature and more regulators and industry practitioners process this guidance.

This is operational risk management in an AI-native financial ecosystem. Welcome to privacy, security, and data design for the age of AI.

---

[1] These components are on the Cyber Risk Institute website in the "Resources and Downloads" section at Financial Services AI Risk Management Framework – Cyber Risk Institute.

[2] Continuous Integration/Continuous Delivery (or Continuous Deployment), commonly referred to as CI/CD, is an automated pipeline that continuously builds, tests, validates, and deploys code so organizations can release updates quickly, safely, and predictably, with appropriate rollback procedures when necessary.

[3] Machine Learning Operations, commonly referred to as MLOps, is the discipline of managing the end-to-end lifecycle of machine learning models (including development, training, validation, deployment, monitoring, retraining, and retirement) through automated, version-controlled, and reproducible processes. MLOps extends CI/CD principles to machine learning environments, incorporating controls for data dependency, model drift, performance degradation, and continuous monitoring.

Contacts

Please contact the listed attorneys for further information on the matters discussed herein.

AMY S. MUSHAHWAR
Partner
Chair, Data Privacy, Security, Safety & Risk Management
T: 202.753.3825
amushahwar@lowenstein.com

CHLOE RIPPE
Associate
T: 212.419.5895
crippe@lowenstein.com

This Alert has been prepared by Lowenstein Sandler LLP to provide information on recent legal developments of interest to our readers. It is not intended to provide legal advice for a specific situation or to create an attorney-client relationship. Lowenstein Sandler assumes no responsibility to update the Alert based upon events subsequent to the date of its publication, such as new legislation, regulations and judicial decisions. You should consult with counsel to determine applicable legal requirements in a specific fact situation. Attorney Advertising.