Just in Time

GenAI Model Risk Management and Governance in Financial Services - From Principles to Practice

February 2026

iason ESSENTIAL SERVICES FOR FINANCIAL INSTITUTIONS

Executive Summary

Generative Artificial Intelligence (**GenAI**) is rapidly being adopted by Financial Institutions to support **analytical**, **operational**, and **decision-support activities.** While these systems deliver significant efficiency gains, they introduce **new and material sources of model risk** that challenge traditional **Model Risk Management** (**MRM**) **frameworks** - namely the U.S. Federal Reserve's SR 11-7 and the UK PRA's SS1/23.

The study[1] explores how **existing model risk management** (MRM) **practices can be adapted** to govern GenAI systems, which rely on **large foundation models**, **dynamic pipelines**, and **Retrieval-Augmented Generation architectures**. Key risk drivers are identified across **data quality issues**, **vendor dependencies**, **behavioural drift**, **reproducibility gaps**, and **automation bias.**

The two **case studies** illustrate **practical adaptations** through **robust validation**, **real-time monitoring**, and **cross-functional governance**. **Recommendations** include **extending model inventories** to full GenAI workflows, **refining risk tiering**, **formalizing lifecycle monitoring**, and **embedding vendor oversight**. Overall, GenAI does not require a wholly new governance regime, **but targeted extensions** to existing MRM practices.

[1] *The Alan Turing Institute: GenAI Model Risk Management and Governance in financial services - from principles to practice.*

At a Glance

Keywords: GenAI, Model Risk, Model Risk Management framework, Retrieval-Augmented Generation, LLMs

01

Introduction

Overview

Introduction

Overview

The report "**GenAI Model Risk Management and Governance in Financial Services: From Principles to Practice**", published by the Alan Turing Institute examines how Financial Institutions (FIs) can adapt existing **Model Risk Management** (**MRM**) frameworks - namely the U.S. Federal Reserve's SR 11-7 and the UK PRA's SS1/23 - to govern **Generative AI** (**GenAI**) systems.

The FIs emphasis the systemic importance of resilience, transparency and control in the use of quantitative models. Traditional financial models underpin essential activities such as credit assessment, capital allocation, stress testing, trading and are therefore governed by mature MRM frameworks. These frameworks establish expectations around **model identification**, **development**, **validation**, **governance**, **and ongoing monitoring**, with a strong emphasis on conceptual soundness and accountability.

The **GenAI represents a qualitative shift** in model usage rather than a simple extension of existing machine-learning techniques. GenAI systems - particularly those based on large language models (LLMs) - are capable of automating complex analytical and cognitive tasks, generating narrative outputs, and supporting a wide range of use cases including document summarization, compliance monitoring, code generation and customer interaction.

However, GenAI systems differ materially from traditional models in several respects:

- They rely heavily on foundation models provided by third-party vendors.
- They operate on dynamic, unstructured and often real-time data sources.
- They produce qualitative outputs for which objective ground truth may be absent.
- Their behaviour may evolve over time due to vendor updates, data changes, or architectural reconfiguration.

Supervisory authorities have responded by intensifying scrutiny of AI adoption in financial services. The Financial Stability Board, the Bank for International Settlements, and the European Central Bank have all highlighted GenAI-specific risks, including hallucination, opacity, concentration risk, and systemic interdependence. While SR 11-7 and SS1/23 provide a robust foundation, their underlying assumptions - stability, interpretability, and bounded data - are increasingly strained by GenAI's scale and dynamism.

The **GenAI does not require an entirely new governance regime**, but rather **targeted refinements and extensions** to existing MRM practices.

02

The RAG Risk Landscape

Data Risks

Vendor Risks

Architecture Risks & Human Factors

The RAG Risk Landscape 1/3

Data Risks

This section provides a structured, non-exhaustive analysis **of the risk dimensions introduced by Retrieval-Augmented Generation** (**RAG**) architectures, which are widely adopted in financial services. RAG systems combine LLMs with external retrieval pipelines, enabling models to ground outputs in institution-specific document bases. GenAI risk extends beyond model accuracy, encompassing **data**, **vendor, architecture and human-factor risks**. These risks are interconnected and often emergent, particularly in modular and vendor-dependent systems.

Data Risks

RAG systems depend on **dynamic, heterogeneous, and semi-structured data**, such as internal reports, regulatory filings, and market commentary. Unlike traditional models trained on curated datasets, RAG systems retrieve information at runtime, increasing exposure to data quality and governance failures.

Document Base Quality

The quality of a RAG system's outputs is directly determined by the quality of its document corpus. Traditional notions of data cleanliness are insufficient; instead, institutions must assess:

- **Semantic fidelity and contextual completeness**
- **Topical coverage and authoritative sourcing**
- **Freshness, versioning, and provenance metadata**
- **Extraction accuracy from PDFs and unstructured formats**
- **Appropriate chunking and representation**
- **Identification of sensitive or regulated information**

Data quality becomes a **governance responsibility**, requiring demonstrable traceability between input quality and output behaviour. Continuous monitoring is necessary, as some data issues only emerge post-deployment. Effective practice links data governance controls with RAG-specific performance metrics such as groundedness and relevance.

Legal and Compliance Burdens

RAG pipelines introduce legal and compliance risks across ingestion, retrieval, generation, and logging. These include:

- Personal and confidential data handling
- Intellectual property and licensing constraints
- Records retention and deletion
- Data residency and cross-border transfer
- Prompt injection and rephrasing attacks

While FIs already operate mature compliance frameworks, GenAI introduces **new vectors** that challenge existing controls. Heavy reliance on vendor APIs raises concerns around **vendor dataflows, information leakage, and log governance**. The vendor transparency and utilisation monitoring must become integral components of model validation and ongoing monitoring.

Ground Truth Availability

GenAI systems often lack a single, objective ground truth. Outputs such as summaries or recommendations are inherently qualitative, influenced by tone, framing, and user expectations. This undermines traditional validation metrics and necessitates alternative approaches:

- Structured Subject Matter Experts (SME) review
- Scenario-based testing
- Longitudinal monitoring of consistency and factual grounding

Supervisory guidance increasingly emphasises **ongoing monitoring over static validation**, recognising that GenAI behaviour evolves over time.

The RAG Risk Landscape 2/3

Vendor Risks

Vendor dependency is a defining feature of GenAI systems. Unlike traditional in-house models, GenAI capabilities are typically accessed via third-party APIs, introducing significant governance challenges.

Vendor Risks

Artefact Re-versioning

Foundation models are updated frequently, often without detailed change logs. Even minor updates can alter model behaviour, creating **behavioural drift between validation cycles**. Snapshot versioning offers only temporary stability, and deprecations may occur with limited notice.

This dynamic conflicts with traditional MRM processes designed for relatively static models. To mitigate this, institutions are embedding:

- Explicit version inventories
- Vendor attestations
- Structured change-impact assessments
- Enhanced validation triggers aligned to vendor updates

Availability and Cost

Reliance on external APIs exposes FIs to:

- Availability and latency risks arising from outages or throttling
- Cost volatility driven by token pricing, retrieval volume, and safety layers

The materiality of these risks varies by use case. High-impact applications require tighter controls, while support tools may tolerate greater variability. Integrating cost and availability oversight into MRM frameworks is therefore essential.

Open-Weight and On-Premises Models

Open-weight models and on-premises hosting can reduce certain vendor risks, such as data disclosure and provider lock-in, but shift other risks in-house. These include:

- Operational complexity
- Security and observability challenges
- Capacity planning and cost management
- Legal and licensing constraints

The open-weight models are not a universal solution and should be deployed as part of a **diversification strategy**, consistent with supervisory guidance on concentration risk.

The RAG Risk Landscape 3/3

Architecture Risks & Human Factors

Architecture Risks

GenAI systems are typically composed of multiple interacting components: retrievers, embeddings, LLMs, orchestration logic, and evaluators. Small specification gaps can propagate and produce **emergent failure modes**. Key architectural risks include:

- Interface coupling and undocumented dependencies
- Behavioural drift across components
- Fragility of LLM-based evaluators
- Configuration complexity undermining reproducibility

The end-to-end system governance is essential. Validation must encompass the full pipeline, not individual components in isolation. This may require revisiting how models are defined and scoped within MRM inventories.

Human Factors

Human interaction with GenAI systems introduces additional risk dimensions that cannot be mitigated through technical controls alone.
Four patterns are highlighted:

- **Automation bias**: over-reliance on model outputs
- **Algorithmic aversion**: undue resistance to model recommendations
- **Cognitive offloading**: erosion of human expertise
- **Mode confusion**: misunderstanding system authority and role

Effective risk management therefore requires **human-factor controls**, including training, interface design, role clarity, and operational guardrails.

03

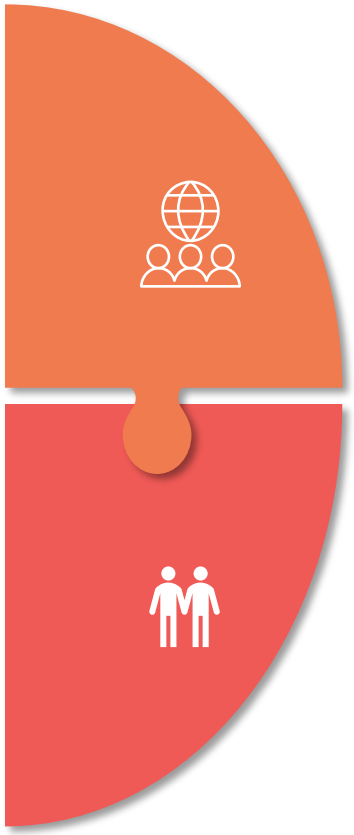
Case Studies

Digital Credit Platform (DCP) & Lead Recommendation Engine (LRE)

Case Studies

Digital Credit Platform (DCP) & Lead Recommendation Engine (LRE)

This section presents two detailed case studies illustrating how FIs are adapting MRM principles for GenAI systems in practice.



Digital Credit Platform (DCP)

The DCP case study describes a RAG-based system used to draft sections of wholesale credit applications. The system operates within a federated GenAI platform and generates candidate text grounded in uploaded documents. Key features include:

- Human-in-the-loop design
- Clear operating boundaries
- Shared ownership between business and technical teams
- Low-impact use case with strong accountability

Validation and monitoring practices extend traditional MRM approaches with GenAI-specific tests, including hallucination and groundedness metrics. Daily monitoring via LLM-as-a-Judge enables near-real-time detection of drift and vendor-related changes.

Lead Recommendation Engine (LRE)

The LRE case study examines a multi-agent system supporting client relationship management. The architecture includes retrieval agents, persona agents, and critic agents orchestrated within an internal platform.

- Architecture-agnostic risk taxonomy
- Interdisciplinary AI Release Board
- SME-driven validation
- Continuous monitoring of vendor and architectural risks

The case demonstrates how **qualitative evaluation and cross-functional governance** can substitute for traditional challenger-model testing where quantitative benchmarks are unavailable.

04

Operationalising GenAI Governance in Financial Institutions

Organisation-Level Challenges & Implications for MRM Practice

Operationalising GenAI Governance in Financial Institutions

Organisation-Level Challenges & Implications for MRM Practice

This section translates the earlier lessons **into practical governance recommendations**, structured around organisation-level challenges and implications for SR 11-7 and SS1/23-style frameworks.

Organisation-level Challenges

Adoption at Scale

GenAI enables rapid development of numerous lightweight tools, challenging the scalability of traditional MRM processes. Institutions are responding by:

- Standardising onboarding for foundation models
- Validating reusable components centrally
- Grouping low-materiality tools for shared review

AI Centres of Excellence

Many institutions have established AI Centres of Excellence to coordinate expertise across technical, risk, and business functions. These bodies support consistent practices, early issue identification, and accountability across the GenAI lifecycle.

Implications for MRM Practice

Model Identification and Scope

The end-to-end GenAI workflow is the relevant unit of risk. Model inventories should capture:

- System boundaries
- Key components and prompts
- Vendor dependencies
- Affected business processes

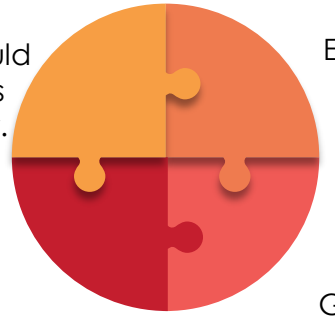
Prompts and orchestration logic should be treated as formal model artefacts subject to version control and review.

Risk Tiering

Traditional tiering frameworks remain applicable but should be extended with GenAI-specific factors such as:

- Vendor dependency
- Change velocity
- Degree of autonomy

Explicit re-tiering triggers enhance transparency and auditability.



Organisation, Processes, and Skills

GenAI governance requires broader functional involvement, including data protection, third-party risk, information security, and compliance. Clear role definitions, training, and documentation standards support consistent application of MRM principles.

Monitoring and Change Management

Given the frequency of change in GenAI systems, institutions are adopting a **lifecycle view of assurance**, with defined materiality thresholds, controlled rollouts, and detailed logging to support reproducibility and supervisory confidence.

05

Conclusions and Key Takeaways

Conclusions and Key Takeaways

GenAI can be governed effectively within existing MRM frameworks, provided those frameworks are adapted to reflect the technology's dynamism, modularity and vendor dependence.

a. **The end-to-end GenAI system is the true unit of risk**, not individual models or artefacts.

b. **Static validation must be complemented by continuous monitoring**, particularly in vendor-dependent environments.

c. **Vendor oversight is integral to model risk management**, requiring explicit documentation and change-management integration.

d. **Human-factor controls are essential**, alongside technical safeguards.

e. **Cross-functional governance structures** enable effective operationalisation of SR 11-7 and SS1/23 in GenAI contexts.

Practical Priorities for Financial Institutions

- **Extend model inventories** to capture full GenAI workflows and dependencies.
- **Refine risk tiering** with GenAI-specific dimensions and explicit review triggers.
- **Embed lifecycle monitoring and change management** as core MRM practices.
- **Integrate vendor governance** directly into model documentation and validation.
- **Formalise cross-functional AI governance and training** to maintain meaningful human oversight.

ESSENTIAL SERVICES FOR FINANCIAL INSTITUTIONS

iason is an international consulting firm that has been supporting both financial institutions and regulators in topics related to Risk Management, Finance and ICT since 2008

Strategy

Strategic advisory on the **design** of **advanced frameworks** and **solutions** to fulfil both **business** and **regulatory needs** in Risk Management and IT departments

Methodology & Governance

Implementation of the designed **solutions** in bank departments **Methodological support** to both **systemically important financial institutions** and **supervisory entities**

Solution

Advanced **software solutions** for **modelling, forecasting, calculating** metrics and **integrating** risks, all on cloud and distributed in Software-as-a-Service (**SaaS**)

KEEP IN TOUCH

iason

Company Profile

iason is an international firm that consults Financial Institutions on Risk Management. **iason** integrates deep industry knowledge with specialised expertise in Market, Liquidity, Funding, Credit and Counterparty Risk, in Organisational Set-Up and in Strategic Planning.

Dario Esposito

Tommaso Travenzoli

www.iasonltd.com