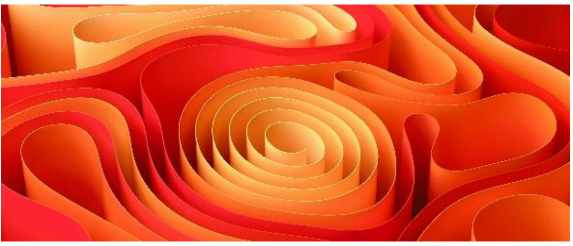
Pills

Regulatory/Supervisory Pills | N.9 January 2026



TOPICS:

Model Risk

SOURCE:

The Alan Turing Institute

# ATI: Model Risk Management of GenAI Workflows

- The rise of **generative AI** (GenAI) in financial services introduces transformative potential and **significant model risk**. This the Alan Turing Institute (ATI)' paper outlines a structured approach to model risk management (MRM) for GenAI workflows, adapting principles from existing supervisory guidance such as the U.S. SR 11-7 and UK SS1/23 to the unique complexities of GenAI systems.
- GenAI systems are characterized by **composability, dynamism** and **open-ended outputs**, requiring enhanced governance, design and monitoring practices.
- The authors propose a **three-pillar** framework: 1) **Governance and Tiering** - Institutions must update their model inventories to reflect GenAI-specific components - LLMs, RAG systems, prompt templates, and toolchains. Risk tiering should incorporate factors such as model autonomy, content sensitivity and configuration volatility. Change management is crucial due to the fluid nature of GenAI components, including vendor APIs and third-party tools;

2) **Design Standards** - GenAI development must emphasize rigorous documentation, alignment with intended use and transparent justification of heuristic design choices. Fitness metrics must extend beyond standard ML loss functions to include domain-specific risk metrics (e.g. factual accuracy, compliance). Key areas include sensitivity analysis, data quality evaluation and uncertainty quantification. Effective model design requires clear fallback strategies and the ability to detect, measure, and respond to uncertainty; 3) **Testing and Monitoring** - Traditional benchmarks must be supplemented by dynamic, task-specific evaluations including adversarial probes, stress testing and red teaming. Continuous monitoring is recommended, using business-aligned metrics, to ensure resilience and mitigate emergent failure modes in production. Uncertainty quantification and process transparency are highlighted as pivotal for risk mitigation.