KPMG

How AI is changing model risk management

kpmg.com

Executive Summary

This paper is the first in a multi-part series. Upcoming papers will dive deeper into classification, risk tiering, independent validation, continuous monitoring, change management for fine-tuned models, retrieval-augmented generation (RAG), and third-party AI (Artificial Intelligence) oversight—providing practical tools to help you scale AI responsibly and cost-effectively.

As AI adoption accelerates, Model Risk must adapt to AI's unique factors: opaque mechanisms, unstructured inputs, dynamic behavior, and continual change. Institutions must move from periodic, checklist reviews to dynamic, risk-based, outcome-focused controls that right-size validation and monitoring, manage frequent retraining updates, and integrate third-party oversight while enabling AI adoption.

Not all AI tools are models. Treating every AI use case as a model over-governs, inflates cost, and creates bottlenecks. Institutions must classify AI functions, separate decisioning models from assistive tools, and tier risk so heavy governance is reserved for high-impact, high-risk decisions. This tiered approach cuts time-to-value, reduces control cost, and enables safe scale—while meeting evolving regulatory expectations.

Traditional MRM Frameworks are Insufficient to Address AI Risks

Applying traditional Model Risk principles to AI systems becomes significantly more complex—and in some cases infeasible—because these systems process unstructured data, adapt behavior over time, rely on mechanisms that resist traditional explanation, and can have a nearly infinite number of uses. Traditional frameworks were built for statistical and econometric models: mostly parametric, interpretable via coefficients and clear causal structures, enabling human challenge. Related regulatory guidance has emphasized conceptual soundness, data quality, methodological rigor, diagnostic testing, outcomes evaluation, and ongoing performance monitoring. For AI, those assumptions break. Post-hoc explainability replaces direct interpretability; continuous drift detection replaces static monitoring; and frequent retraining and data refreshes make change management central. Utilizing traditional frameworks and timelines for AI can be costly and slow, turning Model Risk into a bottleneck to AI development and use.

Transforming Model Risk Management Practices

Model Risk for AI requires a practical shift from predominantly point-in-time controls to continuous, risk-based practices across four pillars: governance, development, validation, and monitoring. Financial institutions need to update these pillars to address AI-specific risks (e.g., bias, explainability, and model drift) and to strengthen oversight of third-party AI so models remain accurate, reliable, and compliant as they evolve. The goal is outcome-focused control that right-sizes effort by risk tier, reduces manual overhead, and shortens cycle time—managing risk without over-engineering.

Risk Based Governance

AI models necessitate governance that is transparent, accountable, and calibrated to risk. An effective approach starts with clear definitions and taxonomy: distinguish AI models from traditional models, categorize by functionality (e.g., tabular ML, NLP/LLMs, RAG, computer vision), potential impact, data types, and associated risks. Not all models are AI—and not all AI is high risk—so classification drives efficiency.

Risk tiering then prioritizes effort by evaluating criteria such as operational significance, customer impact, data sensitivity, opacity, retraining frequency, and vendor dependency. Tiering enables institutions to allocate resources, avoid box-checking, and cut time and cost.

Context matters. Use cases like credit underwriting or fraud detection have distinct risk profiles. Governance should align to business objectives while safeguarding against domain-specific risks. Given AI's dynamic nature, governance must include continuous monitoring and periodic re-validation to detect and address behavior changes quickly, mitigating drift through real-time data assessments and performance checks.[1]

As institutions rely more on vendor-provided AI, robust third-party governance becomes critical. Strengthen the interface between Model Risk and Third-Party Risk Management (TPRM)—with TPRM as the tip of the spear. Capture AI use via vendor attestations, upgrade due diligence, refine contractual provisions (transparency, audit rights, incident SLAs, change controls), and establish continuous capability monitoring. Done well, this reduces duplication, speeds onboarding, and lowers oversight costs.[2]

Development

AI development is inherently different than traditional statistical modeling. It often leverages advanced ML and deep learning, uses unstructured data (text, images, audio), and depends on multiple abstraction layers and large datasets for pattern recognition. Development should confront three realities: complexity and opacity, data quality and bias, and accuracy in dynamic environments.[2]

A disciplined, efficient development process emphasizes reproducibility, gated promotion, and secure operational practices: model registries with versioning; reproducible pipelines; canary releases and rollback criteria; immutable audit trails; and change classification (material vs. immaterial). Data governance must enforce lineage, consent, usage restrictions, proxy detection for protected classes, and policy for synthetic data. Robustness and safety testing—covering adversarial resilience, prompt-injection resistance for LLMs, and content safety filters—should be built into the pipeline. The aim is to reduce rework and cycle time while improving model quality.

Validation Frameworks

Validation must evolve from static, checklist-driven reviews to outcome-focused, right-sized testing tailored to each model's risk tier. Enhanced frameworks include:

- **Explainability:** For high-stakes decisions, implement layered explainability—global (feature importance, partial dependence), local (SHAP/LIME), and decision narratives that support disclosures (e.g., adverse action). Aim for meaningful insight sufficient for customer and regulator understanding.
- **Drift and dynamic retraining:** AI systems retrain and evolve. Establish ongoing validation processes with drift detection that identifies deviations from the validated baseline. Define retraining triggers, materiality thresholds, and validation gates before promotion; avoid accumulating "validation debt" that slows delivery and increases risk.[4]
- **LLM/RAG-specific validation:** Evaluate faithfulness/groundedness (answers supported by retrieved sources), hallucination rates, toxicity and PII leakage, jailbreak/prompt-injection resilience. For retrieval, measure precision/recall, MRR/nDCG, corpus coverage, and freshness. Right-size these tests to the model's impact to control cost and time.
- **Bias and fairness testing:** Assess outputs across demographic groups and protected characteristics. Select metrics that fit the use case (e.g., equalized odds for fraud, adverse impact ratio and demographic parity considerations for credit), document thresholds and trade-offs, and reassess periodically.[3] Focus on demonstrable fairness outcomes, not exhaustive metric catalogues that add cost without insight.

Ongoing Monitoring

AI models can degrade quickly; quarterly or annual reviews won't suffice. Monitoring should be real-time or near-real-time, automated, and event-driven.[5] Track performance (calibration, AUC/PR-AUC), data quality and drift (population stability, concept drift), and prediction distributions. Monitor fairness by segment over time and set alert thresholds. For LLM/RAG, add groundedness scores, retrieval health, hallucination/toxicity rates, PII detection, latency/SLOs, and prompt/template change audit. Automation reduces manual effort, shortens detection-to-action time, and lowers operating cost.

Regulatory Expectations are Evolving

The regulatory landscape is shifting toward principle based, risk proportional expectations rather than blanket prescriptive rules. In many jurisdictions and use cases, requirements are clarifying—and in some cases relaxing—with greater emphasis on demonstrable outcomes such as fairness, transparency, appropriate human oversight, and sound documentation.[6] While states continue to advance diverse regimes, a federal approach is evolving alongside them, and timing, scope, and harmonization remain uncertain.[7]

Against this backdrop, the priority for financial institutions is to manage AI model risk effectively and cost efficiently. That means right sizing controls to each model's risk tier, streamlining validation and documentation, automating continuous monitoring, and integrating third party oversight to avoid duplication and unnecessary cost. Designing flexible, outcome driven controls positions institutions to meet current obligations today and adapt as the federal framework takes shape—managing risk pragmatically without over engineering.

Strategic Implications and Future Outlook

Transforming Model Risk for AI is an organizational change, not a process tweak. Institutions should recruit or outsource specialized talent (data science, ML engineering, explainability, AI ethics), and implement monitoring platforms, automated testing frameworks, and governance workflows built for dynamic AI. The objective is clear: reduce cycle time, cut the cost of control, and avoid bottlenecks that slow AI delivery.

Conduct targeted assessments against AI requirements to identify gaps in governance and validation. Uplift documentation standards with standardized artifacts (model cards, data cards, change logs). Deploy real-time monitoring tailored to AI, and integrate cloud-based AIOps to manage scale. Clarify accountabilities and decision rights; set risk appetite for AI; and prioritize transparency in development. Institutions that move quickly will take cost out, accelerate safe deployment, grow revenue, and attract more customers.

Conclusion

AI is reshaping decisioning—and Model Risk must keep pace. The path forward is not choosing between traditional and AI-enabled methods, but integrating the strengths of both into risk-based, continuous, and outcome-focused practices. Financial institutions that modernize Model Risk will manage AI risk more effectively and cost-efficiently, speed innovation, and strengthen trust with customers and regulators. Those that cling to periodic, checklist-driven control will face growing bottlenecks and missed opportunities.

KPMG helps financial institutions operationalize this evolution. We assess current practices; strengthen governance and independent validation to address bias, explainability, and drift; establish continuous monitoring; and integrate oversight of third-party AI. Drawing on deep financial services experience and proven tools, we build regulator-ready operating models, documentation, and training so organizations can scale AI responsibly while maintaining confidence with regulators, boards, and customers. We serve GSIBs, super-regional, regional, and local banks—and we help clients reduce control costs and time-to-value while improving risk outcomes.

Sources

[1] Precisely, "Opening the Black Box: Building Transparent AI Governance Frameworks" by Sue Pawlak (August 12 2025)

[2] MineOS, "AI Governance Framework: Key Principles & Best Practices" by Gal Golan (May 20, 2025)

[3] Stanford Institute for Human-Centered Artificial Intelligence, *The 2025 AI Index Report* (2025)

[4] Databricks, "Introducing the Databricks AI Governance Framework (DAGF v1.0)" (2025)

[5] National Institute of Standards and Technology (NIST), *Artificial Intelligence Risk Management Framework (AI RMF 1.0)* (Report No. NIST AI 100-1, January 2023)

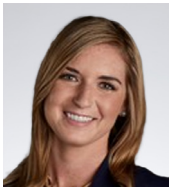
[6] Colorado General Assembly, "SB24-205: Consumer Protections for Artificial Intelligence" (May 17, 2024)

[7] Reuters, "US Senate strikes AI regulation ban from Trump megabill" by David Morgan & David Shepardson (July 1, 2025)

Contacts

Adam Levy
Principal
FS Risk, Regulatory, and Compliance Services
KPMG LLP
E: adamlevy@kpmg.com

Ben Harden
Managing Director
FS Risk, Regulatory, and Compliance Services
KPMG LLP
E: bharden@kpmg.com



Kelly Combs
Managing Director
FS Risk, Regulatory, and Compliance Services
KPMG LLP

Jacob Armstrong
Director
FS Risk, Regulatory, and Compliance Services
KPMG LLP
E: jacobarmstrong1@kpmg.com

Kevin Lowery
Managing Director
FS Risk, Regulatory, and Compliance Services
KPMG LLP
E: klowery@kpmg.com

Abigail Holden
Managing Director
FS Risk, Regulatory, and Compliance Services
KPMG LLP

Liming Brotcke
Director
FS Risk, Regulatory, and Compliance Services
KPMG LLP
E: lbrotcke@kpmg.com

Sakineh Tavakkoli
Director
FS Risk, Regulatory, and Compliance Services
KPMG LLP
E: stavakkoli@kpmg.com

Some or all of the services described herein may not be permissible for KPMG audit clients and their affiliates or related entities.

Learn about us: **kpmg.com**