AI Update

CAHILLNXT

AI Model Validation in Regulated Financial Firms: Supervisory Expectations and Practical Considerations

I. Introduction

Over the past several years, broker-dealers and investment advisers have begun incorporating artificial intelligence ("AI") tools into an expanding range of regulated and operational workflows, including customer communications, trading strategies, investment research, compliance monitoring, and internal administrative processes.[1] While the adoption of these technologies continues to accelerate, U.S. securities regulators have consistently emphasized that existing legal and regulatory obligations remain fully applicable as they apply to this new technology.[2]

The SEC, FINRA, and other financial regulators have repeatedly underscored that their rules are technology neutral, meaning that the same supervisory, recordkeeping, disclosure, and investor-protection requirements apply whether a task is performed by a human employee, traditional software, or an AI system.[3] This principle has important practical implications. For example, FINRA has reminded member firms that AI tools used in business processes must be subject to supervisory systems and governance structures in the same manner as other technologies used in regulated workflows.[4] The consequence of this is straightforward: when firms deploy AI in supervised workflows, regulators expect supervisory systems to account for the integrity, reliability, and accuracy of the models, the sufficiency of policies and procedures surrounding those models, and the protection of client information.[5]

Against this backdrop, it is clear that the challenge for regulated firms is not implementing new regulatory obligations, but translating existing requirements into governance frameworks appropriate for AI-enabled systems. In practice, this involves determining how to adapt in a reasonable fashion long-standing compliance concepts—such as

---

[1] *See, e.g.*, 2026 Annual Regulatory Oversight Report (Dec. 2025), at 24–26. AI, as defined by FINRA, "broadly refers to applications of technology to perform tasks that resemble human cognitive function and is generally defined as the capability of a machine to imitate intelligent human behavior." FINRA, Artificial Intelligence (AI) in the Securities Industry (June 2020).

[2] *See* FINRA Regulatory Notice 24-09 (June 27, 2024).

[3] FINRA Regulatory Notice 24-09 (June 27, 2024) ("FINRA's rules–which are intended to be technology neutral and the securities laws more generally, continue to apply when member firms use Gen AI or similar technologies in the course of their businesses, just as they apply when member firms use any other technology or tool"); *see also* FINRA, 2026 Annual Regulatory Oversight Report (Dec. 2025), at 24; Erik Gerding, The State of Disclosure Review, June 24, 2024, available at https://www.sec.gov/newsroom/speeches-statements/gerding-statement-state-disclosure-review-062424#_ftnref10.

[4] *See* FINRA Regulatory Notice 24-09 (June 27, 2024).

[5] *See* SEC Department of Examinations, Fiscal Year 2025 Examination Priorities, at 13-14, available at https://www.sec.gov/files/2025-exam-priorities.pdf.

supervision, model validation, recordkeeping, privacy controls, and vendor oversight—to technologies that function quite differently from legacy systems.

This memorandum examines issues that may become increasingly important for broker-dealers and investment advisers as AI adoption continues to expand. And considers how traditional model-validation concepts may need to evolve when applied to generative AI systems that produce nondeterministic outputs.

II. Why Traditional Validation Concepts Strain under Generative AI

Traditional model validation frameworks in financial services were developed in the context of deterministic systems—models whose outputs can be evaluated against stable benchmarks using repeatable inputs.[6] In many established risk and quantitative modeling environments, validation methodologies assume that identical inputs should yield consistent outputs. This assumption underlies widely used validation techniques such as back testing, benchmarking, sensitivity analysis, and rule-based testing.[7] The core assumptions consist of a bounded input space and a bounded output space with a repeatable true answer. Validation under these conditions is well understood: compare outputs to expectations, test at defined intervals, and document results. However, large language models ("LLMs") invert these assumptions. They generate output probabilistically, an identical prompt executed at different times can yield different outputs,[8] and there is often no single correct answer against which to test. Conventional back testing and challenger-model comparisons become unreliable under such circumstances.[9] The essential task of validating a model where variance is expected does not align with existing validation playbooks.

From a compliance and supervisory perspective, this distinction can create challenges that extend beyond technical LLM performance. Where firms deploy generative AI tools within regulated workflows—such as compliance monitoring, client communications, or investment research—variability in model outputs can complicate the ability of validation teams, internal audit, and regulators to apply conventional testing methodologies. Scripts, checklists, and sampling frameworks that assume stable outputs may prove less effective when language-based systems may produce a broad range of open-ended responses rather than deterministic calculations.

Regulators have increasingly highlighted these governance challenges. For example, FINRA's 2026 Annual Regulatory Oversight Report identifies a range of governance risks associated with AI deployment, including AI agents operating without human validation or approval, agents acting beyond the user's intended scope, difficulty auditing or explaining automated decisions, the storage or exposure of sensitive data, and misaligned reward functions that may produce unintended outcomes.[10] The report also emphasizes well-known model risks such as bias, hallucinations, and related reliability concerns.[11] In this environment, the central supervisory concern may not be the probabilistic nature of AI systems themselves, but whether firms can demonstrate that appropriate controls exist around how these tools are used. One practical implication is the growing importance of governance mechanisms

---

[6] FINRA, in its 2020 AI in the Securities Industry Report, endorsed the Federal Reserve's definition of Model Risk Management which describes it as "the set of processes and activities intended to verify that models are performing as expected, in line with their design objectives and business uses. Effective validation helps to ensure that models are sound, identifying potential limitations and assumptions and assessing their possible impact." Board of Governors of the Federal Reserve System, Supervisory Letter (SR 11-7) on Guidance on Model Risk Management (Apr. 4, 2011), available at https://www.federalreserve.gov/supervisionreg/srletters/sr1107.htm.

[7] *See* Krishan Sharma, *Rethinking Model Validation for AI Governance*, Risk.Net (Feb. 2, 2026), available at https://www.risk.net/comment/7963013/rethinking-model-validation-for-genai-governance?.

[8] *See generally* IBM, What is an AI model, available at https://www.ibm.com/think/topics/ai-model.

[9] *See* Krishan Sharma, *Rethinking Model Validation for AI Governance*, Risk.Net (Feb. 2, 2026).

[10] FINRA, 2026 Annual Regulatory Oversight Report (Dec. 2025), at 27.

[11] *Id*.

designed to manage known limitations of generative models. For example, LLMs may produce so-called "hallucinations," generating information that appears plausible but is inaccurate or unsupported by source material.[12] FINRA has highlighted the risk that such outputs could affect compliance reviews, client communications, or supervisory processes if not properly monitored.[13]

Moreover, there is a human-factors dimension that compliance professionals should take seriously. Automation bias—the tendency for humans to accept machine-generated outputs uncritically, especially when they appear authoritative or technically sophisticated—can exacerbate these risks.[14] If an associated person accepts an AI-drafted communication without meaningful review, or a compliance analyst relies on an AI-generated surveillance summary without independent verification, a human checkpoint may become a formality rather than a control.

In sum, the validation problem presents as a governance problem. Legacy testing techniques—scripted tests, checklists tied to stable outputs, narrow sampling—may be poorly equipped in the era of LLMs.

III. Emerging Validation Concepts

Though traditional validation methods may not work well to validate probabilistic, language-based systems, the answer is not to abandon validation but to rethink how it can be achieved. To address these gaps, new validation approaches have begun to emerge. These range from basic guidance on responsible supervision of AI systems to more detailed governance frameworks. This section details some of these concepts.

A. Active Monitoring and Human-In-the-Loop Oversight

One fundamental aspect of emerging AI governance frameworks is an emphasis on human-in-the-loop oversight and controls designed to ensure that AI-generated information is independently reviewed before it is relied upon in regulated workflows.[15] In the broker-dealer context, human-in-the-loop oversight generally refers to supervisory structures that require a qualified person to review, validate, or approve AI-assisted outputs before those outputs are used in customer communications,[16] compliance processes and conclusions, trading decisions, or other regulated activities. This approach is consistent with the long-standing supervisory requirements well understood in the financial services industry.[17]

---

[12] *See* OpenAI, Why Language Models Hallucinate (Sep. 5, 2025), *available* at https://openai.com/index/why-language-models-hallucinate/.

[13] FINRA, 2026 Annual Regulatory Oversight Report (Dec. 2025), at 25.

[14] Lauren Kahn, Emelia S. Probasco, and Ronnie Kinoshita, *AI Safety and Automation Bias: The Downside of Human-In-The-Loop*, Center for Security and Emerging Technology (November 2024), available at https://doi.org/10.51593/20230057.

[15] FINRA, 2026 Annual Regulatory Oversight Report (Dec. 2025), at 26.

[16] Use of LLMs in customer communication is further complicated by the use in chatbots or similar features, in which the response is immediate. FINRA has affirmed that "firms using AI technology to create chatbot communications that are used with investors" still have an obligation to "supervise the chatbot communications in accordance with applicable FINRA rules." FINRA, Rule 2210 Frequently Asked Questions About Advertising Regulation, Questions B.4 and D.8 (May 10, 2024), available at https://www.finra.org/rules-guidance/guidance/faqs/advertising-regulation. This clarifies that FINRA expects, among other things, "firms to establish, maintain, and enforce written procedures for the review of incoming and outgoing written [ ] correspondence relating to the firm's investment banking or securities business."

[17] FINRA, Rule 3110; *see also* FINRA, Artificial Intelligence in the Securities Industry (June 2020), available at https://www.finra.org/rules-guidance/key-topics/fintech/report/artificial-intelligence-in-the-securities-industry/key-challenges (affirming that FINRA Rules 3110 and 3120 includes "having reasonable procedures and control systems in place for supervision" of AI-based tools across "applicable functions of a broker dealer").

While regulators have not mandated any particular governance structure, FINRA and the SEC's technology-neutral approach to regulation suggests that the key question for firms will be whether their supervisory systems clearly allocate responsibility for reviewing and approving outputs generated by AI-enabled processes.[18] FINRA Rule 3110 requires broker-dealers to maintain a supervisory system reasonably designed to achieve compliance with applicable securities laws and FINRA rules, and that obligation applies regardless of whether a task is performed by a human employee or an automated system. As such, one supervisory approach would be to treat AI systems as participants within the supervisory chain. Under this model, a designated supervisory principal or other responsible person retains accountability for reviewing or approving AI-generated outputs used in regulated activities, analogous to the oversight applied to human employees.

In practice, human-in-the-loop oversight often operates through monitoring and validation controls embedded in a firm's AI governance framework. These controls may include ongoing monitoring of prompts, responses, and outputs to confirm that a generative AI system continues to perform as expected and produces compliant results and implementing validation and human review of model outputs, including periodic checks for errors or bias.

B. Limiting Model Drift

A related challenge for firms deploying generative AI systems is the phenomenon commonly described as model drift—the gradual deterioration or alteration of model behavior over time as inputs, data environments, or usage patterns evolve. Traditional financial models are also subject to drift, typically arising from changes in market conditions or shifts in the statistical relationships underlying model assumptions. However, generative AI systems introduce additional dimensions of drift because their outputs are shaped not only by training data, but also by prompt design, system instructions, and evolving user behavior.

Moreover, generative AI tools are often deployed within shifting software ecosystems that include retrieval systems, document repositories, application programming interfaces, and third-party model providers. Changes to any component of this ecosystem, or the addition of new data sources may alter system behavior. In some cases, drift may arise not from changes in the model itself but from how users interact with the system, as employees experiment with new prompts or apply the tool to use cases beyond its original design.

Regulators have begun to emphasize the importance of monitoring these dynamics as part of firms' broader AI governance programs.[19] FINRA has noted that firms adopting AI technologies should incorporate life-cycle testing and monitoring into their supervisory frameworks, including maintaining inventories of AI models and evaluating their reliability and accuracy over time.[20] The use of AI in compliance, surveillance, or trading processes may therefore require ongoing review to ensure that model outputs remain consistent with the firm's policies, procedures, and risk appetite.

---

[18] *See* FINRA Regulatory Notice 24-09 (June 27, 2024) (stating that FINRA's rules are "intended to be technology neutral" and "continue to apply when member firms use Gen AI or similar technologies in the course of their businesses, just as they apply when member firms use any other technology or tool"); Mark T. Uyeda, SEC Comm'r, Remarks at the SEC Roundtable on Artificial Intelligence in the Financial Industry (Mar. 27, 2025) ("[F]inancial regulators should take a technology-neutral approach to regulation."), available at https://www.sec.gov/newsroom/speeches-statements/uyeda-ai-roundtable-032725.

[19] FINRA, 2026 Annual Regulatory Oversight Report (Dec. 2025), at 26 (highlighting the risk of models that are limited, inaccurate, or trained on outdated training data as potentially "leading to concept drifts").

[20] FINRA, Artificial Intelligence (AI) in the Securities Industry (June 2020); *see also* FINRA, 2026 Annual Regulatory Oversight Report (Dec. 2025), at 26

One way to test for model drift is to employ techniques that focus on semantic consistency rather than exact textual replication.[21] This approach involves comparing the meaning of outputs generated in production environments with outputs generated during testing or validation phases. Rather than asking whether two outputs are identical, this technique evaluates whether the outputs convey materially similar conclusions or recommendations.[22] While these methods remain relatively novel and have not been formally endorsed by regulators, they reflect an effort to adapt traditional monitoring concepts to the realities of language-based systems.

Ultimately, limiting model drift is less about eliminating variability than about ensuring that variability occurs within defined governance boundaries. Firms experimenting with AI-enabled workflows may therefore wish to consider whether their monitoring processes can detect meaningful changes in system behavior, whether supervisory personnel retain sufficient visibility into how AI tools are used across the organization, and whether model outputs remain aligned with the firm's written policies and procedures.

IV. Conclusion

In sum, the regulatory treatment of artificial intelligence in the securities industry remains grounded in existing, technology-neutral principles. Broker-dealers and investment advisers are therefore not confronted with an entirely new regulatory regime, but rather with the challenge of adapting long-standing supervisory, fiduciary, recordkeeping, and governance frameworks to technologies that produce probabilistic outputs and evolve through use. As firms experiment with AI-enabled tools, regulators are likely to focus on familiar questions: whether firms can explain how AI systems operate, demonstrate that their outputs are subject to appropriate human supervision and review, and evidence that those tools are producing results that are consistent with written policies and procedures. Firms that approach AI deployment through the lens of existing compliance frameworks, while remaining attentive to emerging validation and governance challenges, may be best-positioned to integrate these technologies responsibly within regulated financial markets.

* * *

If you have any questions about the issues addressed in this alert, or if you would like a copy of any of the materials mentioned in it, please do not hesitate to call or email Frank J. Weigand (partner) at 212.701.3890 or fweigand@cahill.com and Louis Capizzi (associate) at 212.701.3482 or lcapizzi@cahill.com.

[21] *See* Krishan Sharma, *Rethinking Model Validation for AI Governance*, Risk.Net (Feb. 2, 2026), available at https://www.risk.net/comment/7963013/rethinking-model-validation-for-genai-governance?.

[22] See FINRA, 2026 Annual Regulatory Oversight Report (Dec. 2025), at 26 (recommending "[o]ngoing monitoring of prompts, responses and outputs to confirm the GenAI solution continues to perform as expected"); *cf*. Sharma, supra note 20 (proposing embedding-based similarity metrics as a means of operationalizing this principle).