THE REVIEW OF
BANKING & FINANCIAL SERVICES
A PERIODIC REVIEW OF SPECIAL LEGAL DEVELOPMENTS AFFECTING LENDING AND OTHER FINANCIAL INSTITUTIONS

Vol. 41 No. 9 September 2025

ARTIFICIAL INTELLIGENCE MODELS IN FINANCIAL SERVICES: EMERGING ISSUES AND AREAS OF RISK

This article explores the rapidly evolving landscape of artificial intelligence in the financial services industry, and discusses emerging risks and regulatory challenges. It examines how financial institutions are leveraging predictive and generative AI to enhance operations, and highlights the tension between innovation and compliance, particularly with the patchwork of emerging state laws and guidance that seek to address algorithmic bias and data governance. Practical considerations for data management, model risk, and explainability are discussed to help institutions consider how to responsibly implement AI solutions. As the regulatory environment continues to shift, this article offers timely insights for industry participants seeking to balance innovation with regulatory compliance.

By Sherry-Maria Safchuk, Sasha Leonhardt, Caroline Stapleton, and Samantha Goldberg-Seder *

Artificial Intelligence ("AI") is changing the way we view and analyze data. With predictive AI, financial institutions increasingly began to implement models that analyzed prior consumer information to predict outcomes — an improvement over previous formulas. However, in the past three years, the advent of generative AI[1] has taken human capabilities to an entirely new level of efficiency — AI can draft essays, analyze thousands of rows of data, and summarize multi-volume treatises instantaneously. Furthermore, in the years since, reliance on agentic AI has increased, and AI agents and interactive chatbots that can communicate with consumers without human intervention have

[1] Large language models ("LLMs") are a subset of generative AI systems trained on extensive datasets to generate human-like language and responses. They utilize deep learning techniques, particularly transformer architectures, to perform a range of language-related tasks. Nat'l Inst. of Standards & Tech., *footnote continued from previous column…*

Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile 1, 3–4, July 2024, NCSL Artificial Intelligence 2025 Legislation, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf.

* *SHERRY-MARIA SAFCHUK is a partner at Orrick, Herrington & Sutcliffe LLP's Santa Monica, CA office. SASHA LEONHARDT and CAROLINE STAPLETON are partners at the same firm's Washington, DC office, and SAMANTHA GOLDBERG-SEDER is a senior associate at Orrick's Silicon Valley office. Their e-mail addresses are ssafchuk@orrick.com, sleonhardt@orrick.com, cstapleton@orrick.com, and sgoldbergseder@orrick.com. Special thanks go to SHEILA ZERANG, an associate at the firm's Santa Monica office. The views and opinions expressed in this article are solely those of the authors and do not necessarily reflect the policy or position of Orrick or any of its clients.*

RSCR Publications LLC Published 12 times a year by RSCR Publications LLC. Executive and Editorial Offices, 2628 Broadway, Suite 29A, New York, NY 10025-5055. Subscription rates: $650 per year in U.S., Canada, and Mexico; $695 elsewhere (air mail delivered). A 15% discount is available for qualified academic libraries and full-time teachers. For subscription information and customer service call (866) 425-1171 or visit our Web site at www.rscrpubs.com. General Editor: Michael O. Finkelstein; tel. 212-876-1715; e-mail mofinkelstein@gmail.com. Associate Editor: Sarah Strauss Himmelfarb; tel. 301-294-6233; e-mail shimmelfarb@comcast.net. To submit a manuscript for publication contact Ms. Himmelfarb. ISSN: 1051-1741.

proliferated. Their success, as well as the improving technology, have led senior executives, board members, and rank-and-file employees across industries to explore new use cases for AI.

Unsurprisingly, while financial institutions may be slow to adopt new technologies, such companies are investing in digital tools that utilize AI to improve the accuracy, efficiency, and cost-effectiveness of their operations. While financial services companies have been relying on algorithmic data analysis, including predictive AI, for decades to estimate default rates, identify prepayment risk, combat fraud, and assess asset quality, generative AI and complex machine learning models are also presenting new use cases in areas like complaint analysis, compliance monitoring and testing, and customer communications. As consumers seek greater (and quicker) access to financial services, with less friction and greater interactivity, the financial services industry stands poised to engage AI to meet the market's evolving demands.

Still, industry participants, their regulators, and consumer advocates recognize that although the use of these tools may have significant societal benefits, such uses are not without risk. Given that many of the applicable federal and state requirements were enacted before the widespread adoption of digital AI tools, financial services companies and their service providers often face the challenge of tailoring the use and development of this novel technology to ensure compliance with legal requirements established during paper-based or early computing eras. At the same time, changes to existing laws and regulations to specifically address AI — and, in some cases, the introduction of entirely new legislation aimed at curbing AI-related risks — further complicate entities' aim to innovate responsibly in a rapidly changing regulatory environment.

This article discusses, at a high level, the federal and state obligations that apply to financial institutions that implement AI technology. We also identify several best practices that can be implemented as the industry grapples with ever-developing AI technology and regulators' varying expectations.

UNDER THE REGULATORY MICROSCOPE

Few industries are subject to the same level of constant regulatory vigilance and oversight that exists in the financial services industry. Indeed, many financial services companies are subject to ongoing regulatory oversight throughout their operational lifecycle — from the investigation and de novo licensing application process required to offer certain products or services, to periodic reporting of information to federal and state agencies, to independent reviews and audits from partners, and routine examinations by regulatory bodies. As a result, financial services companies are subject to constant oversight, in many cases, by multiple regulators.

This oversight can come from both federal and state agencies. Federal regulatory bodies, such as the Consumer Financial Protection Bureau ("CFPB"), the Office of the Comptroller of the Currency ("OCC"), the Federal Deposit Insurance Corporation ("FDIC"), the Federal Reserve Board ("FRB"), and the National Credit Union Administration ("NCUA"), have oversight over a whole host of financial institutions and financial services companies including national banks, bank holding companies, state-chartered banks, credit unions, non-depository financial institutions, and other financial services entities. At the state level, brokers, lenders, servicers, debt collectors, and money transmitters, among others, that are required to obtain state licenses, are subject to additional state-level supervision and review. Furthermore, for all institutions, state attorneys general have broad authority under state laws designed to protect their residents from harm. Accordingly, financial institutions and their service providers, and financial services companies, are constantly interacting with regulators, many of whom have their own robust views on the appropriate use of AI.[2]

[2] *See, e.g.,* NCUA, *Artificial Intelligence Compliance Plan*, Nov. 2024, https://ncua.gov/ai/ncua-artificial-intelligence-compliance-plan; CFPB, CFPB Issues Guidance on Credit Denials by Lenders Using Artificial Intelligence, Sept. 19, 2023, https://www.consumerfinance.gov/about-us/newsroom/cfpb-issues-guidance-on-credit-denials-by-lenders-using-artificial-intelligence/ (withdrawn); New York Department of Financial Services, Industry Letter, Cybersecurity Risks Arising from

In addition, the web of federal laws that apply to the financial services industry outside of privacy and AI is vast — and most of these laws were enacted decades before AI and have seen only limited revisions. For example, at the federal level, the Equal Credit Opportunity Act ("ECOA") and, for mortgage lending, the Fair Housing Act ("FHA"), both prohibit discrimination on the basis of various protected characteristics, including race, color, religion, national origin, sex, marital status, age, or use of a public assistance program.[3] Although these statutes are decades old and the existence of AI was not contemplated at their enactment, federal regulators during the prior administration indicated that they have authority to enforce these laws to prevent algorithmic bias, including discriminatory outcomes arising from opaque or "black box" models.[4] Similarly, interagency guidance issued last year emphasized that it was the responsibility of the CFPB, Department of Justice ("DOJ"), Equal Employment Opportunity Commission ("EEOC"), and Federal Trade Commission ("FTC") "to ensure that these rapidly evolving automated systems are developed and used in a manner consistent with federal laws" and noted that each of these agencies had "previously expressed concern about potentially harmful uses of automated systems."[5]

State regulators have also taken up the AI oversight mantle in parallel with — and at times ahead of — their federal counterparts. The National Conference of State Regulators noted that in 2025, "48 states and Puerto Rico have introduced legislation on the topic this year, while 26 states adopted or enacted more than 75 new measures."[6] Some of these provisions expressly prohibit algorithmic discrimination, similar to the broad historical prohibition under federal law. For example, Colorado recently enacted an AI-specific anti-discrimination law, colloquially referred to as the Colorado AI Act ("CAIA"), which is modeled after European Union AI legislation. The CAIA, which is slated to take effect on Feb. 1, 2026, prohibits algorithmic discrimination, which is defined as "any condition in which the use of an artificial intelligence system results in an unlawful differential treatment or impact that disfavors an individual or group of individuals on the basis of their actual or perceived age, color, disability, ethnicity, genetic information, limited proficiency in the English language, national origin, race, religion, reproductive health, sex, veteran status" or any other protected class.[7] The anti-discrimination requirements in the CAIA apply to "high-risk" AI uses, including in housing and financial services.

Once the CAIA goes into effect, financial services companies that develop and/or deploy AI models and do not qualify for an exemption may be required to comply with requirements aimed at preventing or detecting algorithmic discrimination, including risk assessments and consumer notices. Other states may soon follow in Colorado's footsteps.[8] For example, Virginia's legislature recently passed similar legislation — though the governor vetoed it in March 2025 — and a number of other states have pending or enacted legislation prohibiting discrimination in the use of AI.[9] Thus, while the existing federal laws may not expressly address technological advancements in financial services, states

footnote continued from previous page…

Artificial Intelligence and Strategies to Combat Related Risks (Oct. 2024), https://www.dfs.ny.gov/industry-guidance/industry-letters/il20241016-cyber-risks-ai-and-strategies-combat-related-risks.

[3] Equal Credit Opportunity Act, 15 U.S.C. § 1691(a); Fair Housing Act, 42 U.S.C. § 3605(a).

[4] *See, e.g.,* CFPB Circular 2022-03, Adverse Action Notification Requirements in Connection with Credit Decisions Based on Complex Algorithms, May 26, 2022, https://www.consumerfinance.gov/compliance/circulars/circular-2022-03-adverse-action-notification-requirements-in-connection-with-credit-decisions-based-on-complex-algorithms/ (withdrawn).

[5] EEOC, CRT, FTC, and CFPB Joint Statement on Enforcement Efforts Against Discrimination and Bias in Automated Systems, Apr. 25, 2023, https://www.ftc.gov/system/files/ftc_gov/pdf/EEOC-CRT-FTC-CFPB-AI-Joint-Statement%28final%29.pdf.

[6] NCSL Artificial Intelligence 2025 Legislation, https://www.ncsl.org/technology-and-communication/artificial-intelligence-2025-legislation.

[7] Colo. Rev. Stat. § 6-1-1701(1)(a) (effective February 1, 2026).

[8] At the time of drafting, pending federal legislation includes a proposal to prohibit states from regulating AI for a 10-year period upon enactment. While this proposal failed, to the extent this or similar restrictions are at the federal level, there may be a profound chilling effect on states' ability to pursue AI-related legislation or regulatory interpretations.

[9] *See, e.g.*, Cal. S.B. 420, 2025–2026 Leg., Reg. Sess.; Ill. S.B. 2203, 104th Gen. Assemb., Reg. Sess.; Md. S.B. 936, 2025 Gen. Assemb., Reg. Sess.; Mass. H. 94, 194th Gen. Ct.; N.Y. S. B. 822; N.Y. Assemb. B. A3265, 2025 Leg., Reg. Sess.; R.I. S. 627, 2025 Gen. Assemb., Reg. Sess.; Tex. H.B. 149, 89th Leg., Reg. Sess.; Tex. H.B. 1709, 89th Leg., Reg. Sess.; Vt. H. 340, 2025–2026 Leg., Reg. Sess.

appear to have taken the lead on enacting new legislation to regulate the use of AI in this space.

ENSURING AN AI-READY COMPLIANCE PROGRAM

Against this backdrop, financial services companies face a conundrum: how can they implement AI-based solutions and other advanced analytics tools to enhance their products and maintain and grow market share while complying with both existing and pending laws and evolving regulatory expectations? While there is no easy solution — and the answer changes based on the company, the business model, and a company's tolerance for regulatory risk — legal and compliance departments, in consultation with senior management and the board of directors, where appropriate, should consider thoughtfully identifying steps they can take concerning the use of AI that may complement or enhance regulatory compliance.

Data Considerations

Understanding and assessing the data used in an AI model is one of the most important components to determining the federal and state laws that may apply. AI and machine learning have the "ability to continuously learn from and make predictions based on data."[10] The quality and source of the data used in the AI model likely impacts whether the results of such model may be relied upon.[11] Specifically, the data used to train the model likely materially impacts how the model operates, as "models inherit the characteristics of the data they are trained on, which can include bias, inaccuracy, and impropriety."[12]

Furthermore, it is critical that a financial services company has the proper authority to collect and use data to build models. Several cases are currently pending where plaintiffs have asserted that, due to copyright, privacy concerns, or contractual limitations, companies lacked the right to use certain data to build AI models.[13] In addition, in several consent orders, the FTC found that companies lacked the authority to use data and mandated "algorithmic disgorgement" as a remedy, whereby the companies needed to eliminate all prohibited data from their models.[14]

Moreover, some of the data itself may have novel characteristics, particularly if used in connection with offering and providing financial services. For example, some models use "alternative data," which includes information not traditionally included in financial underwriting models, such as cash flow data, employment type, educational information, or online activity metrics. Consumer advocates and regulators have recognized that using alternative data can have the effect of expanding access to credit to previously underserved populations, for whom traditional data may not capture important indicators of creditworthiness.[15] However, the use of alternative data may also create risk where the inputs are, as the CFPB has previously expressed, "not intuitively related to [consumers'] finances or financial capacity."[16] In such cases, there may be challenges in identifying reasons for credit decisions in adverse action notices, which are required by ECOA and the Fair Credit Reporting Act ("FCRA"). Moreover, if an AI model can associate alternative data points with certain protected characteristics (*e.g.*, proxy identification), and use such information to make decisions about eligibility for financial products or services, there is an increased risk that the financial institution may violate anti-discrimination laws such as ECOA, FHA, and similar state laws.

---

[10] Madison Blevins, *When Dirty Data Leads to Dirty Policing*, 29 Rich. J.L. & Tech. 166, 174 (2023) (citations omitted).

[11] *Id*.

[12] GAO, *Artificial Intelligence: Generative AI Technologies and Their Commercial Applications*, GAO-24-106926, at *5 (June 20, 2024).

[13] *Kadrey v. Meta Platforms, Inc.,* No. 3:23 cv 03417 VC (N.D. Cal. filed July 7, 2023) (complaint); *Reddit, Inc. v. Anthropic, PBC, No*. CGC-25-524892 (Cal. Super. Ct. S.F. Cnty. June 4, 2025) (complaint); *Authors Guild v. OpenAI Inc.*, No. 1:23 cv 08292 (S.D.N.Y. Sept. 19, 2023) (complaint).

[14] *In the Matter of Cambridge Analytica, LLC*, Final Order, FTC Docket No. 9383, Nov. 25, 2019, https://www.ftc.gov/legal-library/browse/cases-proceedings/182-3107-cambridge-analytica-llc-matter; *In the Matter of Everalbum, Inc., Final Decision and Order*, FTC File No. 1923172 (May 6, 2021), https://www.ftc.gov/system/files/documents/cases/1923172_-_everalbum_decision_final.pdf; *United States v. Kurbo, Inc. & WW Int'l, Inc.*, Stipulated Order for Permanent Injunction and Civil Penalty Judgment, No. 3:22-cv-00946-TSH (N.D. Cal. Mar. 3, 2022, https://www.ftc.gov/system/files/ftc_gov/pdf/wwkurbostipulatedorder.pdf.

[15] *See, e.g.,* Terri Bradford, *"Give Me Some Credit!": Using Alternative Data to Expand Credit Access*, Payments System Research Briefing, Fed. Rsrv. Bank of Kan. City (June 28, 2023), https://www.kansascityfed.org/research/payments-system-research-briefings/give-me-some-credit-using-alternative-data-to-expand-credit-access/.

[16] CFPB Circular 2023-03, Sept. 19, 2023 (withdrawn).

The following list identifies queries that financial institutions may consider when assessing their AI model data inputs for fair lending and other regulatory risks:

- Whether the data was obtained under appropriate authority, if any, and the source(s) of the data.
- Whether the data originates from a consumer reporting agency or amounts to consumer report information.
- Whether the data favors or disfavors lower-income/disadvantaged consumers.
- Whether the company is able to ascertain the output of the AI model to describe the reason or reasons for an adverse action accurately.
- Whether an adverse action notice citing this data as a reason or reasons for the adverse action could lead to reputational harm if made public, such as through publication in a national newspaper.
- Whether the data — alone or together with other data — is a proxy for distinguishing a protected class.
- Whether there is a credible explanation for the relationship of the data to creditworthiness.

This is not a comprehensive list; rather, these are just examples of some questions companies may consider when implementing an AI model.

Managing Risk When Using AI Models

If data fields are the ingredients for an AI model, the model's parameters are the recipe. The algorithms comprising AI models are similar to complex sets of instructions, and in many cases, these may change over time if the model is designed to "learn" through inferences drawn by comparing its predictions to actual outcomes. Model risk management procedures are important in understanding and evaluating the effectiveness of, and risk presented by, AI models. While many institutions already have model risk management policies and programs that have been in place for some time, these policies may need to be updated to account for the proliferation of AI models offering financial products and services.

Notably, federal laws have not yet imposed specific requirements or restrictions on the development and maintenance of AI models. However, federal financial regulators have issued longstanding guidance on how entities should approach model risk management generally.[17] The various federal guidance materials focus on the use of a "model," which the FRB has defined as "a quantitative method, system, or approach that applies statistical, economic, financial, or mathematical theories, techniques, and assumptions to process input data into quantitative estimates," a term broad enough to apply to a wide spectrum of AI tools.[18]

Even though much of this guidance was issued before the proliferation of AI, and there is debate as to whether such guidance is even appropriate for AI-driven models,[19] certain high-level principles may be useful for financial services companies that use AI in their operations to consider:

- *Identify the goals of the AI model.* Regulators have noted that financial services companies should consider creating "disciplined model development and implementation processes" that align with the financial services company's situation, goals, and internal policy.[20] This means that, at implementation, the financial services company may consider documenting the model's purpose and intended use, taking into consideration the underlying sources of data, methodologies, and processing concepts. In addition, financial services companies should understand the underlying theories that drive AI models and their results to ensure compliance with federal and state laws and regulations. This may include testing the model's performance in advance of implementation, with an eye towards replicating real-world circumstances under varying market conditions.
- *Test the AI model.* As with any technology, model validation is an important tool for determining how

[17] FRB, *Supervisory Guidance on Model Risk Management*, Supervision and Regulation Letter 11-7 (2011); OCC, *Sound Practices for Model Risk Management: Supervisory Guidance on Model Risk Management*, Bulletin 2011-12 (2011); FDIC, *Adoption of Supervisory Guidance on Model Risk Management*, Financial Institution Letter-22-2017 (2017); See FRB, *Commercial Bank Examination Manual* (Nov. 28, 2023); FDIC, *Risk Management Manual of Examination Policies* (Oct. 28, 2024); OCC, *Comptroller's Handbook: Safety and Soundness: Model Risk Management* (Aug. 2021).

[18] SR 11-7, April 4, 2011, https://www.federalreserve.gov/supervisionreg/srletters/sr1107.htm.

[19] Bank Policy Institute, The Most Damaging "Guidance" in Banking (April 2025), https://bpi.com/the-most-damaging-guidance-in-banking/.

[20] *Id*.

the algorithm is functioning and whether the end product aligns with the financial services company's goals, operates as intended, and produces accurate results. Fundamental errors that result in inaccurate outputs or inappropriate use of an otherwise sound model can increase the risk that a financial services company may not be compliant with applicable laws. To address these challenges, before putting models into effect, financial services companies often test their models to ensure that their results both comply with the law and further the business's objectives. Testing is not a one-and-done activity; rather, regulators may expect ongoing testing of models with a particular focus on re-evaluation when there are changes to the model or consumers' economic circumstances.

- *Implement specific policies and controls.* The best AI models are only as good as the people who use them, such that AI models should be accompanied by detailed guidance on how to use, maintain, and test them. Such guidance should be understandable, explain how the model operates and its intended use, and identify any assumptions underlying its development or limitations inherent in the model. Model guidance should not be a cut-and-paste process; any guidance should be tailored to the model, the business case, the financial institution using the model, and should be kept up to date.

- *Ongoing monitoring.* Any AI models should undergo ongoing monitoring to determine whether there have been changes to the product, activities, clients, market conditions, law, or other factors that may require recalibration or review of the model. Monitoring should begin when the financial services company tests the model and continues throughout the model's lifecycle.

In addition to the foregoing principles, regular bias reviews of models, particularly in areas such as credit scoring, fraud detection, and underwriting, can promote consistency in results and may mitigate a company's risk of a discrimination claim under the FHA, ECOA, or other federal or state anti-discrimination laws. The implementation of fairness guidelines and the inclusion of diverse data sets during training — as well as ongoing retrieval augmented generation ("grounding" the model using external data sources such as statutes and regulations) and reinforcement learning from human feedback (incorporating manual evaluation of the AI model's performance into the AI model) — may also limit regulatory risk.

Explainability

In addition to scrutinizing *how* AI models are used, regulators have been focused on the *why* behind the outputs these models generate, particularly where the results impact consumers.

The CFPB issued two circulars during the Biden administration regarding the use of complex algorithms in credit decisioning.[21] While the Bureau withdrew these circulars in 2025, they serve to illustrate the fact that regulators are thinking about the importance of explainability in connection with the use of AI in making credit decisions. Specifically, in those now-withdrawn circulars, the CFPB clarified that ECOA's adverse action notice requirements (which remain in force) apply regardless of the technology used to make the underlying decision and that a specific reason for the adverse action must be provided. The reason may not be overly broad or vague such that it obscures the specific and accurate reasons relied upon, and a sample form may not be used where it does not "specifically and accurately indicate the principal reason(s) for the adverse action."[22]

Unfortunately, these circulars did little more than restate longstanding requirements under ECOA and Regulation B, giving little meaningful guidance to financial services companies attempting to comply with these laws.[23] Still, even under a less-active CFPB, state regulators and private litigants may look to the principles in the CFPB's circulars in reviewing the explainability of credit decisions made by regulated entities. As such, financial services companies that use AI for underwriting should be aware of this guidance and factor it into their regulatory compliance programs.

The National Institute of Standards and Technology ("NIST"), under the U.S. Department of Commerce, has

[21] CFPB, Consumer Financial Protection Circular 2022-03 (Withdrawn); CFPB, Consumer Financial Protection Circular 2023-03 (Withdrawn).

[22] *Id.*

[23] 12 C.F.R. 1002.9(b)(2) ("The statement of reasons for adverse action required by . . . this section must be specific and indicate the principal reason(s) for the adverse action. Statements that the adverse action was based on the creditor's internal standards or policies or that the applicant, joint applicant, or similar party failed to achieve a qualifying score on the creditor's credit scoring system are insufficient"); 12 C.F.R. Part 1002 cmt. 9(b)(2)-2 ("The specific reasons . . . must relate to and accurately describe the factors actually considered or scored by a creditor.").

been seen as a leader in setting forth principles for the use of AI, though it is not a financial regulator. NIST's mandate is to establish cybersecurity standards for federal agencies and contractors, and in that capacity, it has issued a number of AI-related frameworks that companies have adopted in a number of industries, including financial services. With respect to explainability, NIST has issued four principles discussing explainable artificial intelligence to which it proposes that AI systems adhere, which may be helpful to consider while navigating AI tools:[24]

- *Explanation*: According to NIST, companies should ensure that they can identify the evidence, support, or reasoning behind an outcome from, or a process of, an AI system. The AI system should be able to deliver or contain "accompanying evidence or reason(s) for outputs and/or processes." These explanations may vary depending on the customer as well as the reason for using the AI model, and may serve different purposes. For example, explanations may be tailored to explain a result to a consumer, or it may take a different approach and include additional information beyond the result.

- *Meaningful*: NIST postulates that an AI system should provide "explanations that are understandable to the intended consumer(s)." Considerations as to what would be a "good" explanation may vary by recipient and require trial and error as well as regular review. Such considerations may include, among many other factors, determining the purpose of the explanation, identifying the recipient, and determining the recipient's prior knowledge related to the AI model and explanations, as well as the expectation for the explanation.

- *Explanation Accuracy*: The third principle relates to the accuracy of the explanation (as opposed to the accuracy of the decision explained) regardless of the level of detail (*e.g.*, whether the explanation accurately describes how the AI model came to the conclusion identified in the explanation). NIST notes that "[s]ometimes, if an explanation does not have 100 percent explanation accuracy, it can be exploited by adversaries who manipulate a classifier's output on small perturbations of an input to hide the biases of a system." Lack of accuracy in the explanation may result in generating misleading explanations.

- *Knowledge Limits*: The final principle NIST identifies relates to the limitations of AI models as well as the AI model's ability to identify such limitations. If an AI model is able to recognize or identify instances that it is not equipped to handle or scenarios where the model's responses may be inaccurate, it may be able to limit misleading results. Recognizing, identifying, and flagging these knowledge limits will help companies identify potential guardrails that may mitigate some risk.

While NIST does not have the authority to issue binding regulations for financial services companies, a recent GAO report on the use of AI in financial services referenced NIST as a source of standards in connection with AI governance and oversight in financial services.[25]

ON THE HORIZON

AI is a multi-disciplinary area that requires companies and their advisors to adopt a similarly multi-disciplinary approach. Notwithstanding the proposed 10-year moratorium on state AI legislation, state and federal regulators are likely to continue introducing and considering legislation that applies to AI generally; predictive, generative, and agentic AI specifically; AI oversight, task forces, and other regulatory commissions, and information gatherings; chatbots; and injury or liability in connection with AI. In addition to AI-specific legislation, ancillary areas such as privacy, conceptual fairness in lending, and black-letter financial regulatory concerns must all be taken into consideration.

AI is rapidly evolving, and as such, the use of any AI model or tool is a moving target with respect to compliance. However, given that a raft of AI laws are presently moving through various legislatures, making it uncertain when specific laws may impact the use of AI in financial services, financial services companies should consider being proactive with respect to analyzing AI models and use cases, with a focus on data considerations, model risk management, and explainability. As industry, government, and everyone else continue to explore AI and its risks and possibilities, more guidance — if not more clarity — is sure to come. ■

[24] Draft NIST IR 8312, *Four Principles of Explainable Artificial Intelligence*, September 2021, https://nvlpubs.nist.gov/nistpubs/ir/2021/NIST.IR.8312.pdf; *see also* NIST AI 100-1 Artificial Intelligence Risk Management Framework (AI RMF 1.0), January 2023, https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf (referring to the 2021 principles).

[25] GAO-25-107197, *Artificial Intelligence: Use and Oversight in Financial Services*, May 19, 2025.