Senate Committee on Banking, Housing, and Urban Affairs

"Artificial Intelligence in Financial Services"

Melissa Koide, CEO, FinRegLab

Good morning. Thank you for the opportunity to testify before you today at the hearing, "Artificial Intelligence in Financial Services."

I am the founder and CEO of FinRegLab. FinRegLab is a DC-based independent, nonpartisan research organization that evaluates the use of new technologies and data to drive the financial services sector toward a responsible and inclusive marketplace. Through our research and policy discourse, we facilitate collaboration across the financial ecosystem to inform public policy and market practices.

FinRegLab has been working on topics relating to machine learning and artificial intelligence (ML/AI) in financial services since 2019, focusing most intensively on the potential benefits and risks of adopting machine learning models for credit underwriting. Our work includes a groundbreaking empirical analysis of techniques for managing explainability and fairness concerns with ML underwriting models, conducted with researchers from the Stanford Graduate School of Business.[1] We have also interviewed and convened dozens of thought leaders to discuss evolving market practice and public policy issues, as reflected in a series of reports that will culminate with a detailed policy analysis of the adoption of machine learning techniques in credit underwriting later this fall.[2] Our research found some explainability techniques can generate useful information about key aspects of model operations and that more automated approaches hold promise for identifying model alternatives that reduce demographic disparities compared to more traditional fair lending strategies.

Looking more broadly, we co-hosted a symposium on "Artificial Intelligence and the Economy" last year with the U.S. Department of Commerce, National Institute of Standards and Technology, and Stanford Institute for Human-Centered Artificial Intelligence.[3]

---

[1] FinRegLab, Laura Blattner, & Jann Spiess, Machine Learning Explainability & Fairness: Insights from Consumer Lending (updated June 2023) (hereinafter Machine Learning Empirical White Paper). As described below, the study evaluated proprietary explainability tools offered by seven technology companies–Arthur AI, H20.ai, Fiddler AI, Relational AI, Solas AI, Stratyfy, and Zest AI–as well as open-source techniques in identifying information relevant to fair lending, consumer disclosures, and model risk management compliance.

[2] For reports to date, see FinRegLab, Explainability & Fairness in Machine Learning for Credit Underwriting: Policy & Empirical Findings Overview (2023) (hereinafter Machine Learning Policy & Empirical Findings Overview), and FinRegLab, The Use of Machine Learning for Credit Underwriting: Market & Data Science Context (2021) (hereinafter Machine Learning Market & Data Science Context Report). For FAQs, podcasts, and webinars on AI in financial services, see https://finreglab.org/ai-machine-learning/.

[3] See https://finreglab.org/artificial-intelligence-and-the-economy-charting-a-path-for-responsible-and-inclusive-ai-2.

Although public interest in ML/AI topics has skyrocketed since ChatGPT was released last November, some applications of machine learning and artificial intelligence in financial services date back decades. Like their peers in other sectors, financial services stakeholders are grappling with the potential benefits and risks of specific use cases and with defining standards for responsible use as the technologies and applications continue to evolve.

Federal regulatory requirements, including model risk management guidance that apply to banks and credit-related laws that apply to lenders more generally, are an extremely helpful starting point in these discussions, prompting stakeholders to identify and manage risks *before* deployment rather than simply reacting after damage has been done on the back end. However, these frameworks are not uniform across the entire financial ecosystem and warrant adjusting for the AI era.

To address these topics, my testimony starts with a general overview of the state of ML/AI adoption across various financial services use cases; potential benefits and risks to customers, providers, and the broader economy; and the ways that federal financial regulatory frameworks are shaping ML/AI adoption in this sector. I then discuss the adoption of machine learning models in credit underwriting, including our research on explainability and fairness issues, and recent reactions to generative AI as two case studies in how financial services providers are responding to ML/AI innovations. I conclude with thoughts on ways in which federal financial regulatory frameworks could be strengthened to encourage responsible use in financial services.

1. Adoption of ML/AI in Financial Services

Artificial intelligence is a term coined in 1956 to describe computers that perform processes or tasks that "traditionally have required human intelligence." Machine learning is often used to refer to the subset of artificial intelligence that gives "computers the ability to learn without being explicitly programmed."[4] In practice, these terms are often used to describe a broad range of models and applications that may be applied to a variety of data sources (e.g., tabular financial data, text, images, and voice) for a wide range of purposes (including predictive analytics, web and information searches, autofill functions, and generating new content in response to queries). The level of human involvement also varies depending on the particular application and technique. At a minimum human oversight is critical in determining what data the algorithms are exposed to, choosing among different techniques and model architectures at the start, and validating, monitoring, and managing the models depending on the desired goals and the risks to address.

---

[4] See, e.g., Financial Stability Board, Artificial Intelligence and Machine Learning in Financial Services (2017); Ting Huang et al., The History of Artificial Intelligence, University of Washington (2006). Machine learning concepts were used by Alan Turing in the 1930s and studied intensively by the Department of Defense beginning in the 1950s. Increases in computing power and digital information sources in the 1990s sparked broader testing and adoption across a broad range of commercial contexts. FinRegLab, Frequently Asked Questions, AI in Financial Services: Key Concepts (2020).

It is important to emphasize that ML/AI is not monolithic and that terminology can vary in different contexts (e.g., pitching to investors versus data science convenings and public policy debates), which can sometimes make it difficult to understand how much new innovations differ from older forms of statistical analysis and automation. There also can be tremendous differences between sectors, use cases, and individual users as to how these ML/AI techniques are deployed and managed.

The financial services sector has used various forms of statistical analysis and automation for decades, including machine learning models in some contexts. Other deployments of machine learning and artificial intelligence techniques are relatively new. Prominent ML/AI use cases include:

- **Fraud detection, anti-money laundering, and related functions:** This use case is one of the oldest and broadest examples of machine learning deployment in financial services. Machine learning models are often used to monitor bank account, credit card, or other financial transactions data to flag suspicious patterns, and may be combined with more traditional rules-based tools. ML models may also be used in connection with initial account opening and application processes. In light of the nature and volume of the data, rapidly evolving fraud patterns, and other considerations, these machine learning models are often more complex and are updated more frequently than those used in other financial services contexts, although they are not necessarily fully dynamic (in the sense of continuously adjusting themselves based on incoming data without developer initiation and validation prior to deployment).[5]

- **Investment and finance uses:** Machine learning models have also been used in trading and other investment contexts for decades.[6] These models analyze markets to inform human investment decisions and the automated trading of securities. They are attracting increased scrutiny where used to recommend investments to smaller investors.[7]

---

[5] Maghsoud Amiri & Siavash Hekmat, Banking Fraud: A Customer-Side Overview of Categories and Frameworks of Detection and Prevention, 2 Journal of Applied Intelligent Systems & Information Services 58 (2021); Bonnie G. Buchanan, Artificial Intelligence in Finance, The Alan Turing Institute (2019); PK Doppalapudi et al., The Fight Against Money Laundering: Machine Learning Is a Game Changer, McKinsey & Co. (Oct. 7, 2022); T.J. Horan, Blog, Evolution of Fraud Analytics—An Inside Story, KD Nuggets (March 14, 2014).

[6] Buchanan; I.V. Dwaraka Srihith et al., Trading on Autopilot: The Rise of Algorithmic
Trading, 3 International Journal of Advanced Research in Science, Communication & Technology (May 2023); Oliver Wyman et al., Artificial Intelligence Applications in Financial Services: Asset Management, Banking and Insurance (2019).

[7] While the same regulations and fiduciary obligations apply to financial advisors whether they leverage machine learning or not, there are concerns that machine learning models may optimize the financial gain of a firm over their client. Recent proposed rulemaking by the Securities and Exchange Commission would apply to any predictive analytics, including more traditional statistical techniques. However, they pose a particular challenge for machine learning models because their complexity makes it generally more difficult to prove they do not exhibit a conflict of interest than more traditional methods. U.S. Securities & Exchange Commission, Press Release, SEC Proposes New Requirements to Address Risks to Investors from Conflicts of Interest Associated with the Use of Predictive Data Analytics by Broker-Dealers and Investment Advisers (July 26, 2023).

- **Marketing:** AI tools are also used to make marketing more efficient and effective, for instance by predicting which potential customers are most likely to be interested in a particular product or service, to respond to specific channels or messages, and ultimately to be profitable for the provider. These models may rely upon a somewhat different and broader range of inputs and different programming parameters than underwriting or other account screening models.[8]

- **Credit decisioning:** Some lenders are increasingly using machine learning to make decisions about loan applications, pricing, and credit line adjustments. Due to regulatory considerations and the seriousness of consequences for both lenders and borrowers, these underwriting models tend to be trained on more heavily curated data, involve simpler architectures, and are subject to more extensive validation in both initial development and updating than models deployed for some other use cases. Machine learning models can also be used in the loan servicing, portfolio management, and collections contexts, for instance to assess which borrowers are likely to stabilize their finances or respond to particular outreach.[9]

- **Insurance:** Somewhat similarly to credit underwriting, machine learning models are being used to forecast risks and manage ongoing customer relationships in the context of insurance. The pace of adoption is expected to increase as the industry builds its technical expertise and solves explainability and regulatory compliance challenges. [10]

- **Customer service and communication functions:** ML/AI tools are utilized in a variety of customer service contexts and to facilitate various forms of customer communication. These include speech-to-text tools to help process customer inquiries over the phone, voice recognition for authentication, natural language processing to classify and process email inquiries, and simple chatbots that provide standardized content in response to customer questions.[11]

The release of ChatGPT in November 2022 sparked broad interest in large language models and other forms of generative AI that can be used to create new content (including text and images) by relating prompts to learned patterns in training data.[12] These models are typically trained on

---

[8] For example, marketing activities sometimes rely on "unsupervised" machine learning models that screen for similarities and patterns in data without having a developer define a specific target variable to predict. In contrast, most machine learning models used for underwriting purposes are considered "supervised" models because developers have defined some measure of default as the target variable to be predicted. FinRegLab, Machine Learning Market & Data Science Context Report 56-58.

[9] FinRegLab, Machine Learning Market & Data Science Context Report § 2.

[10] Financial Stability Board; Geneva Association, Promoting Responsible Artificial Intelligence in Insurance (2020); Oliver Wyman et al. See CAS Machine Learning Working Party, Machine Learning in Insurance, Casualty Actuarial Society E-Forum (Winter 2022) for a discussion of use cases and barriers to machine learning adoption in the insurance industry.

[11] Buchanan; Consumer Financial Protection Bureau, Issue Spotlight, Chatbots in Consumer Finance (2023); Financial Stability Board.

[12] See, e.g. McKinsey & Company, What Is Generative AI? (2023); Mark Riedl, A Very Gentle Introduction to Large Language Models without the Hype, Medium (2023); Foley & Lardner LLP, ChatGPT: Herald of Generative AI in 2023?, JD Supra (2023); David De Cremer et al., How Generative AI Could Disrupt Creative Work, Harvard Business

massive datasets scraped from large portions of the internet and follow many-layered, complex architectures.[13] They have sparked a flurry of interest in their potential applications across a broad range of sectors, but also deep concerns about reliability/accuracy, bias, intellectual property rights, privacy issues, and other risks. In financial services specifically, many providers are testing these technologies particularly for their potential to augment human-based back-office activities and customer service and communication functions. However, our discussions to date with the financial sector indicate they are taking a relatively cautious approach to implementation and deployment as discussed further in Section 3.b. below.

2. Potential Benefits, Risks, and Regulatory Frameworks

Across these and other use cases, the adoption of machine learning and artificial intelligence techniques offer the potential for substantial benefits for financial services providers, customers, and the broader economy—as well as substantial risks. Federal regulatory frameworks encourage industry to identify and mitigate these risks throughout the model development, validation, and deployment process, rather than operating from a reactive posture after implementation. However, the application of these frameworks depends upon the product and nature of the provider, as well as the nature of the particular ML/AI application.

A.   Potential Benefits

As financial services providers are deciding whether and how to adopt ML/AI techniques for various use cases, they are often incentivized by the prospect of some combination of the following potential benefits:

o **Reducing losses:** More accurate predictive tools to identify fraudulent activity, which credit applicants will not be able to repay their loans, and emerging trends in securities trading have obvious intuitive appeal to financial services providers. Particularly where circumstances are evolving rapidly—such as in detecting new forms of fraud or emerging signals in times of substantial economic turmoil—the ability of more powerful algorithms to ingest large and diverse pools of data, detect more nuanced relationships within that information, and to iterate and refine models more quickly has substantial appeal. Financial services firms have used statistical models and automated techniques for decades in these contexts and are naturally evaluating where newer ML/AI applications may further improve their functions.

o **Increasing operational efficiency and consistency:** In some settings, ML/AI techniques have the potential to improve the speed and efficiency of various functions, strengthen

---

Review (2023). The content creation process in generative AI relies on models that predict words or images based on patterns learned in large amounts of sequential data. For instance, auto-fill functions are a low-level version of generative AI that predict the most likely letters or phrases that follow the initial content.

[13] See Nvidia, Large Language Models Explained, https://www.nvidia.com/en-us/glossary/data-science/large-language-models/, visited Sept. 18, 2023; Riedl.

the consistency of human-based knowledge management and customer service systems, and provide various other types of operational efficiencies that can help providers increase the scale and consistency of their operations and reduce costs.

o **Tailoring products and services to particular market segments:** Some ML/AI deployments also offer the potential to tailor financial products and services to better meet the needs of particular customer populations, allowing financial services providers to expand their customer bases and revenues in various ways.

o **Keeping pace with other market actors:** Adoption and advocacy of more powerful technologies by fraudsters, rival financial services providers (both traditional and emerging, such as big tech companies), investors, and other market actors can further accelerate interest in ML/AI applications. While larger institutions and fintech startups tend to have the most resources to devote to exploring and deploying ML/AI techniques, a number of vendors are working to facilitate adoption particularly among smaller, more resource-constrained providers.

The extent to which these potential benefits carry over to individual customers and the broader economy depends on the particular use case and the way in which adoption is carried out by individual providers, but could be substantial in some circumstances. For example, improving the effectiveness and efficiency of anti-money laundering and anti-fraud functions through so-called "federated machine learning" techniques could help reduce losses for customers and prompt financial institutions to restore services, such as remittances, to areas that they consider too high-risk to engage with today.[14]

Similarly, increasing financial services providers' confidence in working with customer segments who are more risky or costly to serve and better tailoring of financial products and services to meet customers' individual circumstances could have particular benefits for populations who have struggled to access safe and affordable financial services historically. For example, despite the benefits of using more standardized data and automated underwriting systems over the last 50 years,[15] about 20 percent of U.S. adults today lack sufficient traditional credit history to

---

[14] Federated machine learning models effectively bring the model to different pools of data rather than consolidating data in the first instance. They are particularly helpful where confidentiality needs are extremely high due to law enforcement, privacy, or other considerations. See, e.g., FinRegLab, Frequently Asked Questions, Federated Machine Learning in Anti-Financial Crime Processes (2020).

[15] Research suggests that these changes have tended to reduce lenders' costs, improve the consistency of treatment of similarly situated applicants, and increase competition for borrowers. Board of Governors of the Federal Reserve System, Report to Congress on Credit Scoring and Its Effects on the Availability and Affordability of Credit S-2 to S-4, O-2 to O-4, 32-49 (2007); Allen N. Berger & W. Scott Frame, Small Business Credit Scoring and Credit Availability, 47 J. of Small Bus. Mgmt. 5 (2007); Susan Wharton Gates et al., Automated Underwriting in Mortgage Lending: Good News for the Underserved?, 13 Housing Policy Debate 369 (2002); FinRegLab, The Use of Cash-Flow Data in Underwriting Credit: Market Context & Policy Analysis 11 n.16 (2020) (hereinafter, Cash-Flow Market Context & Policy Analysis).

generate credit scores under the most widely used models,[16] and roughly a third may struggle to access affordable credit because their scores are "non-prime."[17] Small business owners also often struggle to access commercial credit in part because of information gaps.[18] In each of these cases, communities of color and low-income populations are substantially more likely to be affected by these information barriers than other applicants.[19] Accordingly, machine learning techniques particularly in combination with more inclusive data sources have the potential to make substantial improvements in lenders' ability to predict repayment risks among underserved populations.

Given the importance of consumer spending and small business job formation to the broader U.S. economy, improvements in access to financial services can also have important multiplier effects for economic participation levels more broadly. Credit plays a particularly important role because it can both help bridge short-term gaps, and fund long-term investments in housing, transportation, education, and small business formation. The credit system thus both reflects and influences the ability of families, small businesses, and communities to participate in the broader economy.[20]

B. Risks and Concerns

In addition to considering the potential benefits of ML/AI adoption for particular use cases, it is also important to identify and manage potential risks and concerns. ML/AI applications are often substantially more complex than prior generations of automated systems. This is due not only to the fact human developers play a different role in the model creation process, but also due to the amount and nature of the data that the models are built on, the fact that many of them use more complex architectures that are harder for humans to understand, and the ways in which they trace more nuanced relationships in the underlying data.[21] These and other technical issues

---

[16] Consumer Financial Protection Bureau, Data Point, Credit Invisibles 4-6, 17 (2015); Mike Hepinstall et al., Financial Inclusion and Access to Credit, Oliver Wyman (2022); FinRegLab, Cash-Flow Market Context & Policy Analysis § 2.2.

[17] In lower score bands, the majority of applicants may be likely to repay, but lenders cannot determine which particular applicants are lower risk without additional information. Lenders may choose not to lend to that cohort or may impose higher prices because default risks for the group as a whole are relatively high. FinRegLab, Cash-Flow Market Context & Policy Analysis §§ 2.1, 2.2. The number of consumers with non-prime scores shrank during early stages of the pandemic but delinquencies and other signs of distress are now returning to pre-pandemic levels. Gina Heeb, Credit Scores Went Up in Pandemic. Now, More Borrowers Are Slipping, Wall Street Journal (July 12, 2023).

[18] FinRegLab,The Use of Cash-Flow Data in Underwriting Credit: Small Business Spotlight §§ 2.1, 2.2 (2019).

[19] For example, nearly 30 percent of African-Americans and Hispanics cannot be scored using the most widely adopted credit scoring models, compared to about 16 percent of whites and Asians. Racial disparities regarding access to credit are far greater than for more basic transaction accounts. FinRegLab, Cash-Flow Market Context & Policy Analysis §§ 2.1, 2.2.

[20] FinRegLab, Debt Resolution Options: Market & Policy Context § 2.1 (2022); FinRegLab, The Use of Cash-Flow Data in Underwriting Credit: Small Business Spotlight; FinRegLab, Cash-Flow Market Context & Policy Analysis §§ 2.1, 2.2.

[21] For example, depending on what parameters are set, machine learning models can detect relationships are non-linear (meaning that each incremental increase in an input feature may not change the likelihood of a predicted

raise fundamental questions about our ability to understand, manage, and rely upon the models to perform a variety of tasks. Many of these risks and concerns fall within the following five categories: performance, fairness and inclusion, privacy and other consumer protections, security, and transparency.

Performance

There is little reason to adopt machine learning models if they cannot be relied upon to produce accurate and reliable information and (in many cases) to improve upon the accuracy and reliability of incumbent systems. In addition to evaluating whether a particular model's predictions meet the accuracy needs for its use case and comparing them to the accuracy levels of incumbent models, a second key aspect of performance relates to the robustness of the model's performance where data inputs start to shift due to changes in economic conditions, customer behavior, or other factors.

On this second aspect of reliability, machine learning models' ability to identify a wider range of relationships in training data than incumbent models may increase their susceptibility to performance problems due to two issues: (1) overfitting, or the risk that the machine learning algorithm fits the predictive model too narrowly to the specific characteristics of training data; and (2) data drift, which can occur when conditions in deployment start to differ from the data on which a model was trained, for instance due to shifts in consumer behavior, populations, or economic conditions.

Fairness and Inclusion

Concerns about whether and how machine learning underwriting models could negatively impact populations who have historically been subject to discrimination, exclusion, or other disadvantage are prevalent. These issues are broader than establishing compliance with anti-discrimination laws and include more fundamental questions about gaps in training data, modeling decisions, and other issues that can affect the performance of models for particular groups.

Machine learning models' ability to identify a wider range of relationships in training data also heightens concerns about the risk of replicating or even amplifying historical disparities, particularly in the context of credit underwriting. For instance, some models rely on "latent features" that are identified by the learning algorithms from relationships in the input data rather than intentionally programmed into the models by developers. This raises concerns that the models could reverse engineer applicants' race or gender from correlations in input data or create complex variables that have disproportionately negative effects on particular groups, but

---

outcome by an equal amount) and non-monotonic (meaning that increasing the value of an input feature may reduce the predicted likelihood a particular outcome in some circumstances and increase it in others). The use of salt in cooking illustrates both concepts. The first increment of salt added to a dish may improve the flavor by a different amount than the second increment, and at some point, additional increments will start to make the dish taste worse. Such a relationship is neither linear nor monotonic. FinRegLab, Machine Learning Market & Data Science Context 33-34.

that developers would have difficulty diagnosing or mitigating such problems due to the complexity of the models.

At the same time, the data science community is debating a wide variety of ways to define, measure, and manage for different notions of fairness, including the use of debiasing techniques to mitigate fairness concerns through data curation and through various approaches at different stages of the model development process. These developments have triggered an additional set of debates about whether and how to deploy such definitions and techniques consistent with existing laws and compliance approaches.[22]

Privacy and Other Consumer Protections

The ability of machine learning underwriting models to analyze large, diverse datasets and create deeper, more personalized profiles of consumers is closely tied to their potential benefits for accuracy and inclusion but can also raise significant questions about privacy, fairness, and data protections. This is especially true where the models use elements that feel personally intrusive or lack an intuitive link to what models are predicting. Other concerns include what constitutes informed consumer consent to data access, rules regarding data retention and use, and explaining decisions made by automated systems to help identify errors and educate consumers.

While these issues are not unique to machine learning models, they receive heightened attention in the machine learning context due to strong interest in pairing advanced analytical techniques with diverse data sources. Challenges with regard to the complexity and explainability of machine learning models also increase concern about what data elements are being used for what purposes in order to evaluate privacy and fairness concerns and accurately educate consumers about how to improve the likelihood of future approvals.

Security

The potential for machine learning models to rely on more granular and sensitive data also can heighten concerns about information security.[23] In addition to increasing the potential consequences of security breaches, for instance, stakeholders have identified novel risks in some other machine learning contexts. For example, research suggests that AI systems can be manipulated without direct access to their code,[24] for example by maliciously embedding signals

---

[22] For a detailed discussion of different approaches, see FinRegLab, Machine Learning Market & Data Science Context Report § 5. Two techniques that can be used at the model development phase are joint optimization and adversarial debiasing. In joint optimization, the developer instructs the learning algorithm to optimize for both predictive accuracy and another goal such as minimizing demographic disparities as it builds successive iterations of the model. In adversarial debiasing, the algorithm is instructed to optimize for predictive accuracy and to minimize the accuracy of a separate, "adversarial" model that is designed to predict demographic disparities. The alternative models themselves do not factor in demographic information but the spectrum of choices potentially allows a financial services provider to choose options that will reduce disparities overall.

[23] See generally Andrew Burt & Patrick Hall, What to Do When AI Fails, O'Reilly Radar (2020); Sophie Stalla-Bourdillon et al., Warning Signs: The Future of Privacy and Security in an Age of Machine Learning, Future of Privacy Forum (2019).

[24] Nicolas Papernot et al., Practical Black-Box Attacks against Machine Learning, Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security 506-519 (2017).

in social network feeds or newsfeeds that are not detectable by humans.[25] Further, because machine learning models encode aspects of training data into the mechanisms by which they operate, they have the potential to expose private or sensitive information from the training data to users.[26]

Transparency

Model transparency—the ability of stakeholders in a particular model to access various kinds of information about its design, use, and performance—increases stakeholders' confidence in procedural fairness, consistency, and accountability.[27] In consumer credit, laws requiring lenders to provide applicants with a list of the principal reasons for adverse decisions serve this aim by enabling error correction in underlying credit information and educating consumers about what factors may affect their ability to access credit over time.[28]

Transparency is also often critical for helping to diagnose and manage other concerns about the responsible use and reliability of ML/AI applications. For example, understanding which variables are driving model outcomes can be important to assessing a model's potential sensitivity to changing conditions and to diagnosing and mitigating the sources of demographic disparities in model outcomes. In the financial services context, several regulatory regimes implicitly rely on multiple concepts of transparency to manage various risk concerns.

However, depending on their structure, data sources, and other factors, machine learning models can raise additional transparency challenges relative to incumbent models because of their size, complexity, and reliance on unintuitive data relationships. Stakeholders in financial services and other contexts are debating whether these concerns are best managed by upfront constraints on model complexity, *post hoc* techniques that can be applied to analyze various aspects of model operations, or a combination of those two approaches. These debates focus on the potential accuracy tradeoffs of constraining model architectures and on how to determine whether particular *post hoc* techniques are adequate for particular purposes.[29]

---

[25] Valeriia Cherepanova et al., LowKey: Leveraging Adversarial Attacks to Protect Social Media Users from Facial Recognition, published as a conference paper at the 2021 International Conference on Learning Representations, arXiv:2101.07922 (2021).

[26] See Patrick Hall, Proposals for Model Security: Fair and Private Models, Whitehat and Forensic Model Debugging, and Common Sense (2019) (highlights how surrogate models can be used to extract unauthorized information from a model through inversion or a membership inference attack).

[27] Some stakeholders use terms such as interpretability and explainability to refer to similar concepts, while others use interpretability specifically to refer to the use of up-front constraints on model architecture that are designed to make its operation more inherently understandable and explainability to refer to *post hoc* techniques that applied to better understand model operations. See FinRegLab, Machine Learning Market & Data Science Context Report § 3.

[28] 15 U.S.C. § 1691(d)(6); 12 CFR § 1002.9; 15 U.S.C. § 1681a(k)(1); 12 CFR § 1022.72. See Section 2.c for further discussion.

[29] For a detailed discussion of different approaches, see FinRegLab, Machine Learning Market & Data Science Context Report § 3. Two broad categories of *post hoc* techniques include surrogate models such as Local Interpretable Model-Agnostic Explanations (LIME) that use simpler structures to try to approximate the outcome of more complex models and feature importance metrics such as Shapley Values (SHAP) that calculate the effects of omitting various input features in order to measure the features' aggregate importance to model operations.

Cross-cutting themes

Across these different risk areas, concerns about data quality, scale, and the means of human oversight are recurring themes. ML/AI applications require relatively large amounts of data and computational resources to train effectively, which means they are generally only developed for relatively large-scale applications. However, where data quality is poor, mistakes are made in the development process, or performance deteriorates rapidly in changing data conditions, these scale factors can often increase negative consequences for both customers and providers relative to prior systems. For instance, while underwriting by human loan officers can also be subject to various forms of error and bias, the impact of any one individual officer may tend to be relatively smaller than where underwriting is performed via a single automated system.

Questions about human oversight are deeply interwoven with the various risk and data concerns, particularly transparency. ML/AI models in certain ways are substantially easier to audit and analyze than subjective decision making by individual humans, but they are requiring financial services providers, regulators, and other stakeholders to adjust the ways in which they analyze and monitor models relative to prior generations of automated systems. Many stakeholders view this as a substantial risk and are questioning whether oversight mechanisms can be sufficiently trusted if they do not yield exactly the same kinds of information that can be generated for incumbent models. Others suggest that it may also be an opportunity to revisit existing risk frameworks and practices to consider whether there may be more effective ways to achieve business and policy goals for both incumbent and newer ML/AI systems going forward.

C. Regulatory Frameworks

These broad conceptual risks and concerns are implicated in a number of federal regulatory requirements that apply to various portions of the financial services ecosystem, regardless of the technologies used to deliver those financial products or services.[30] These frameworks are most extensive as applied to retail banking operations and to consumer credit underwriting in particular, and have generally encouraged financial services providers to spend significant upfront resources on risk identification and mitigation before deploying AI/ML applications. They include:

- **Model risk management (MRM) guidance:** To protect the safety and soundness of banks and the broader financial system, banks are expected to implement risk-based governance mechanisms for the development, deployment, and monitoring of models based on the degree of business, compliance, and other reputational risks presented by the particular use case. These expectations govern management of data used to build models, initial model development processes, validation and testing, and monitoring and adjustment after deployment, and include analyzing whether models are relying on relationships in the data that are "conceptually sound" and assessing models' stability in

---

[30] For more background on these regimes, see FinRegLab, Machine Learning Market & Data Science Context Report § 2.3 & Appendix B and Financial Health Network, Flourish, FinRegLab & Mitchell Sandler, Consumer Financial Data: Legal and Regulatory Landscape (2020).

changing data conditions.[31] While the framework does not formally apply to nonbank financial services providers, elements may be adopted as a matter of best practice or required by investors, bank partners, and other counterparties.

- **Vendor due diligence expectations:** When banks outsource functions to outside vendors, they are still accountable for compliance with substantive requirements that would apply if they were to conduct the vendor's activity directly. The Consumer Financial Protection Bureau has issued similar guidance with regard to the nonbanks that it supervises.[32] Given these expectations, covered financial services providers typically create risk management programs to conduct due diligence and ongoing monitoring of vendor relationships, including those involving the provision of ML/AI applications or related services.

- **Information security and data privacy requirements:** Both bank and nonbank financial institutions are required to safeguard customer information and comply with certain privacy requirements under the Gramm-Leach-Bliley Act (GLBA). For example, banks and other providers of consumer financial products and services are required to maintain information security programs that identify reasonably foreseeable risks, assess current safeguards, implement ongoing testing and monitoring, and provide for periodic updates. GLBA also generally prohibits financial institutions from sharing consumers' nonpublic personal information with non-affiliated companies without first providing notice and an opportunity to opt out, although there are exceptions for various business functions.[33]

- **Prohibitions on unfair, deceptive, and abusive acts and practices (UDAAP):** Federal law prohibits businesses from engaging in unfair and deceptive acts and practices, and providers of consumer financial products and services from abusive acts and practices as well.[34] These prohibitions trigger particular scrutiny of ML/AI application that affect customer communications.

---

[31] Board of Governors of the Federal Reserve, Supervisory & Regulatory Letter 11-7: Supervisory Guidance on Model Risk Management (Apr. 4, 2011); Office of the Comptroller of the Currency, Bulletin 2011-12: Sound Practices for Model Risk Management: Supervisory Guidance on Model Risk Management (Apr. 4, 2011); Federal Deposit Insurance Corporation, Financial Institution Letter 22-2017: Adoption of Supervisory Guidance on Model Risk Management (Jun. 7, 2017). MRM expectations for credit unions are generally only focused on interest rate risk. National Credit Union Administration, Interest Rate Risk Measurement Systems, Model Risk (2016).

[32] Bank Service Company Act, 12 U.S.C. §§ 1661–1867(c); 12 U.S.C. §§ 5514–5516; CFPB, Compliance Bulletin & Policy Guidance 2016-02: Service Providers (2016).

[33] 15 U.S.C. §§ 6801–6809; 15 U.S.C. §§ 6801(b), 6805(b)(2). The Consumer Financial Protection Bureau is currently engaged in a rulemaking to implement a 2010 law that provides consumers with the right to access their own information relating to consumer financial products or services that they have obtained in the past, but the law does not apply to other types of financial products and services. 12 U.S.C. §5533; Consumer Financial Protection Bureau, Outline of Proposals and Alternatives under Consideration (2022).

[34] 12 U.S.C. §§ 5531, 5536(a)(1)(B); 15 U.S.C. § 45. Deception focuses on whether material representations or omissions mislead or are likely to mislead a reasonable consumer. Unfairness is generally defined to include activities that (1) cause or are likely to cause substantial injury (usually monetary) to consumers; (2) cannot be reasonably avoided by consumers; and (3) are not outweighed by countervailing benefits to consumers or to competition. Abusiveness focuses on activities that materially interfere with the ability of consumers to understand a term or condition of a consumer financial product or service or take unreasonable advantage of

- **Fair lending compliance:** Federal fair lending laws generally prohibit both (1) the use of race, gender, or other protected characteristics in credit underwriting models ("disparate treatment"); and (2) the use of facially neutral criteria that have a disproportionately adverse impact on protected groups, unless those criteria further a legitimate business need that cannot reasonably be achieved through less impactful means ("disparate impact").[35] These requirements thus focus on both the nature of the inputs and the outputs of underwriting models in evaluating potential bias concerns.

- **Adverse action reporting:** Federal laws require lenders to provide individualized disclosures to credit applicants of the "principal reasons" for rejecting an application and the "key factors" that are negatively affecting consumers' credit scores if the lender charges higher prices based on credit report information.[36] Lenders who are considering use of machine learning underwriting models are grappling with how to produce accurate and reliable disclosures where models rely upon larger numbers of features, more complex structures, and more complex relationships in the data.

However, it is important to note that even within the financial services ecosystem, coverage varies. Many of the frameworks and requirements described above are focused on retail banking, and even within that sphere protections are most robust for credit as compared to other consumer financial products and services.[37] Customer protections for small business owners are often less clearly defined and often subject to less robust monitoring than for consumers.[38] Regulatory scrutiny also varies depending on the type of financial services providers, for instance with banks subject to regular supervision and nonbanks often subject to more limited monitoring.[39]

3. Adoption of ML Underwriting Models and Generative AI

Because data sources, technologies, business and compliance practices, and policy issues can vary significantly across different use cases within the financial services sector, it is helpful to examine ML/AI adoption in different contexts to both identify broader themes and important distinctions in context. My testimony addresses the adoption of machine learning in credit underwriting as the context that FinRegLab has made its recent focus, as well as generative AI as a quite recent example that is attracting significant stakeholder attention. It is important to note

---

consumers in various ways. Financial Health Network, Flourish, FinRegLab & Mitchell Sandler, Consumer Financial Data: Legal and Regulatory Landscape at 135-151.

[35] 15 U.S.C. § 1961(a); 12 C.F.R. § 1002.6(a); 42 U.S.C. § 3605.

[36] 15 U.S.C. § 1691(d)(6); 12 CFR § 1002.9; 15 U.S.C. § 1681a(k)(1); 12 CFR § 1022.72.

[37] The CFPB issued an examination manual update in March 2022 indicating that discrimination in the provision of other consumer financial products and services may constitute an unfair act or practice, but a federal district court ruled in September 2023 that the agency exceeded its authority. Consumer Financial Protection Bureau, Press Release, CFPB Targets Unfair Discrimination in Consumer Finance (March 16, 2022); Chamber of Commerce vs. Consumer Financial Protection Bureau, No. 6:22-cv-00381, Opinion and Order (E.D. Tex. Sept. 8, 2023).

[38] FinRegLab, The Use of Cash-Flow Data in Underwriting Credit: Small Business Spotlight § 5.

[39] FinRegLab, Cash-Flow Market Context & Policy Analysis at 39.

that both examples are considered by financial services stakeholders to be relatively high-stakes applications that are receiving particularly rigorous attention.

A. Machine Learning in Credit Underwriting

The adoption of machine learning in credit underwriting is one of the highest stakes use cases for ML/AI in financial services given the potential impacts on borrowers, lenders, and the larger economy as well as the extensive regulatory frameworks that apply to consumer credit. Relative to other use cases and sectors, adoption of ML underwriting models has been relatively slow as stakeholders have worked to evaluate broad concerns about the models' reliability, fairness, and transparency and uncertainty about meeting related compliance requirements. Where individual market actors have decided to proceed, they are generally relying on heavily curated data (even if it extends beyond traditional credit bureau records[40]) and techniques that require substantial involvement from human developers. Validation and monitoring procedures for initial models and updates can also be extensive, particularly among bank users.

Yet while ML underwriting models are determining the outcomes of millions of credit applications submitted by consumers and small business owners, there are still substantial questions about the pace and breadth of adoption going forward. Adoption is concentrated mainly among large banks and fintech lenders, while smaller bank lenders face substantial technology and resource constraints. Stakeholders are also grappling with how to define best practices and revise policy frameworks as the technologies continue to evolve. Questions about managing explainability and fairness concerns for ML underwriting models are particularly important, including the use of certain secondary techniques that help to analyze how complex models are making their predictions and debiasing approaches that can be used to identify alternative models that may reduce disparities on the basis of race, ethnicity, and gender.

To interrogate these issues, FinRegLab worked with Laura Blattner and Jann Spiess of the Stanford Graduate School of Business to evaluate proprietary tools offered by seven technology companies–Arthur AI, H20.ai, Fiddler AI, Relational AI, Solas AI, Stratyfy, and Zest AI–as well as open-source data science techniques as applied to various tasks relating to model risk management, adverse action disclosures, and fair lending compliance.[41] We also conducted extensive interviews and stakeholder engagement to explore evolving market practice and policy issues. Our findings include the following:

---

[40] For example, a number of lenders are exploring the use of cash-flow data from digital bank account records and other sources as a means of gathering additional information about applicants' finances. FinRegLab, Cash-Flow Market Context & Policy Analysis.

[41] The tools were applied to a variety of models built by the research team using credit bureau data, including a logistic regression model, a constrained neural network, an XGBoost model, and a more complex neural network, as well as to models built by some of the participating companies using the same data. FinRegLab, Blattner & Spiess, Machine Learning Empirical White Paper.

- **Some *post hoc* explainability techniques can provide reliable information about key aspects of model behavior, but stakeholders are still debating their appropriate use and sufficiency.** We found that some techniques provided reliable information about key aspects of model behavior, though there was no "one size fits all" technique or tool that performed the best across all regulatory tasks. Our results emphasize the importance of choosing the right explainability tool for the particular ML model and task, deploying it in a thoughtful way, and interpreting the outputs with an understanding of the underlying data.

  As stakeholders work to develop standards for the appropriate deployment of explainability techniques, the analytical framework we developed for our research project may be useful for evaluating the performance of tools in different settings. Beyond methodological and process questions, stakeholders are grappling with broader questions about whether being able to produce exactly the same types of information that can be generated about more traditional models is critical for various business and compliance functions.

- **The transition to machine learning has the potential to improve fairness and inclusion, in part by giving lenders a more robust toolkit for mitigating disparities.** Despite the focus on transparency as a threshold issue for ML models, in our empirical research the most powerful approaches to managing fairness did not necessarily hinge upon explaining the inner workings of the model as an initial step. Instead, we found that automated approaches that generated a range of alternative models produced options that had greater predictive accuracy and smaller demographic disparities than traditional strategies that assessed which input features made the biggest contribution to disparities and then omitted or made narrow adjustments to those individual features.

  As stakeholders deepen their understanding of various debiasing tools and implementation choices, public policy questions regarding fair lending compliance have taken on additional urgency in light of the adoption of ML models. Lenders have been hesitant to deploy certain debiasing techniques in particular ways because the techniques use data about race, gender, and other protected characteristics in different ways than traditional mitigation approaches. The availability of new debiasing approaches has also highlighted outstanding questions about the nature and extent of lenders' obligations to search for fairer models during the development process.

- **Defining basic concepts and expectations could be a useful first step toward updating regulatory frameworks for the machine learning era.** While ML technologies and our understanding of them are evolving rapidly, regulators can take steps now to encourage responsible use. For instance, defining the key qualities of explainability tools and clarifying expectations about how and when lenders should search for fairer alternative underwriting models would increase consistency of practice and shape how lenders use their expanded toolkits in the ML context. And

while existing model risk management guidance provides a strong principles-based framework for governance that many other sectors lack, it does not apply to the full spectrum of lenders. Stakeholders see potential value in addressing governance concerns for particular subgroups of lenders and articulating basic elements that should be considered in developing and deploying ML underwriting models.

Our research helps to identify topics for additional empirical and policy analyses as stakeholders work to determine how ML underwriting models can be deployed in ways that benefit both borrowers and lenders alike. While the risks merit serious consideration, the research highlights the potential benefits of the technologies, as well as opportunities to improve market practices and regulatory frameworks governing automated underwriting models going forward.

B. Generative AI Adoption in Financial Services

"Generative AI," such as ChatGPT and other large language models, is used to create new content (including text and images) by relating prompts to learned patterns in training data.[42] These models are typically trained on massive datasets scraped from large portions of the internet and follow a many-layered, complex architecture.[43] They are requiring more training data and becoming more complex as models become more advanced. For example, GPT (2018) has 117 million parameters and was trained on five gigabytes of data, GPT-2 (2019) has 1.5 billion parameters and was trained on 40 gigabytes of data, and GPT-3 (2020) has 175 billion parameters and was trained on 45 terabytes of data.[44]

Generative AI appears to be attracting significant interest within the financial services sector as measured by job postings and interviews,[45] but financial services providers are generally taking a cautious approach to implementation. The most immediate application of these models in the financial services context appears to be in connection with facilitating customer service functions, helping developers to write code, and other activities involving substantial amounts of textual analysis and generation. While many financial services providers have already been working for some time with natural language models and chatbots, generative AI has the potential to go beyond identifying patterns in complaints and servicing files and beyond delivering canned content to provide more responsive content. However, while a variety of financial services

---

[42] See, e.g. Riedl; McKinsey & Company; Foley & Lardner LLP; De Cremer et al. The content creation process in generative AI relies on models that predict words or images based on patterns learned in large amounts of sequential data. For instance, auto-fill functions are a low-level version of generative AI that predict the most likely letters or phrases that follow the initial content.

[43] See Nvidia; Riedl.

[44] Min Zhang & Juntao Li, A Commentary of GPT-3 in *MIT Technology Review 2021*, Fundamental Research 1-6, 831-833 (2021).

[45] Generative AI in Finance: Opportunities & Challenges, Gradient Flow Newsletter, https://gradientflow.com/generative-ai-in-finance-opportunities-challenges/, visited Sept. 18, 2023; see also William Shaw & Aisha S Gani, Wall Street Banks Are Using AI to Rewire the World of Finance, Bloomberg (May 31, 2023); Evident: J.P. Morgan, Capital One lead on AI talent in 46k-strong global workforce, Finadium (June 12, 2023).

providers have begun testing versions of these technologies, many banks have banned the use of ChatGPT and similar general-use versions due to a broad range of risk concerns.[46]

Both federal regulatory expectations and broader business incentives to protect financial services' providers data resources are influencing how they approach these new technologies. The computational and data intensity of developing these models suggests that all but the largest financial services providers will have to rely on third party providers of generative AI at least for initial implementation. Notably, providers do not want employees to input sensitive information into third-party models like OpenAI's ChatGPT or Google's Bard because of privacy and security concerns. It is conceivable, for instance, that the information processed by these models could be retrieved by an external actor, putting customer data and proprietary information at risk.[47]

Financial services providers are also expressing concerns that as these models are trained on the data from the wider internet, inaccurate and biased training data could contaminate the models' outputs that lead to obvious or subtle distinctions in what information is provided in response to different queries. This risk can be mitigated by carefully curating and filtering training data. However, as training datasets become even larger to train more accurate models, the management of that data will become more and more resource intensive. Further refining a model by providing feedback to its responses can also somewhat mitigate the risk of biased or inaccurate responses.[48]

Public announcements and our conversations with financial services providers to date suggest that they are mitigating these risks primarily by testing (1) open source models trained at least in part on the companies' own information behind firewalls to prevent data sharing; and (2) applications that are designed to assist coders, customer services representatives, and other knowledge-based employees, rather than to replace them in performing various functions.[49] More broadly, our conversations with financial institutions suggest generative AI is being considered for various other forms of back-office search and document summarization and for coding support—activities that require human review and validation.

As financial services providers are testing the technologies for general reliability, they are also considering the potential application of various federal regulatory frameworks. Customer-facing

---

[46] See, e.g., Brian Bushard, Workers' ChatGPT Use Restricted At More Banks—Including Goldman, Citigroup, Forbes (Feb. 24, 2023).

[47] See, e.g., Jaydeep Borkar, What Can We Learn from Data Leakage and Unlearning for Law? (July 19, 2023), https://arxiv.org/abs/2307.10476.

[48] Extensive adjustments to combat inaccurate and biased responses is very resource intensive, however, as it generally requires humans to continually verify and provide feedback on responses. This can be done at scale by asking users whether they like the response they received, yet relying on user feedback for large language models in the context of financial services would generally be too high-risk due to user error and other factors. An external bad actor could bombard the model with incorrect feedback, for example.

[49] For instance, Morgan Stanley is launching a chatbot to help their financial advisors quickly access relevant Morgan Stanley research and to provide them with administrative support. Tatiana Bautzer & Lananh Nguyen, Morgan Stanley to Launch AI Chatbot to Woo Wealthy, Reuters (Sept. 7, 2023). Intuit has recently announced a new generative AI financial assistant to provide small businesses and consumers with personalized information to make more informed financial decisions. Intuit, Press Release, Introducing Intuit Assist: The Generative AI-Powered Financial Assistant for Small Businesses and Consumers (Sept. 6, 2023).

applications are especially high-risk due to the potential for generative AI models to provide inaccurate information, discriminatory or otherwise harmful outputs, and expose sensitive information in training data. Risk management and fairness expectations make financial services providers hesitant to develop or use generative AI trained on external data as well. Similarly, regulatory compliance demands a level of explainability and transparency of models that many providers are not confident they can attain with generative AI in the near term. For example, the potential utility of the kinds of *post hoc* techniques that FinRegLab evaluated in the credit underwriting context is unclear as applied to generative AI, and early efforts toward explainability are in some cases focusing on asking models to identify specific source materials relied upon in generating new content.

In light of these considerations, a number of financial services providers report that they have imposed special review policies, processes, and thresholds for any generative AI applications, even beyond their normal model risk management and compliance procedures. They are balancing competitive considerations and a desire to improve services and functions with anticipating substantial regulator and public scrutiny of uses of these new technologies.

4. Adjusting Market Practices and Regulatory Frameworks Going Forward

As described above, adoption of machine learning and artificial intelligence applications in financial services has tended to be slower than in some other sectors, yet holds both substantial promise and risk going forward. Federal regulatory frameworks have played an important role in how financial services providers approach testing, development, and implementation processes, particularly by encouraging a more proactive and comprehensive approach to risk management as compared to other sectors. In this respect, the frameworks may provide a useful model for other contexts, and in fact have been cited by the National Institute of Standards & Technology and other stakeholders focusing on ML/AI adoption across a wide variety of markets and use cases.[50]

Yet as we consider the path forward, it is also important to consider how the financial services sector can achieve consistently responsible, fair, and inclusive implementation of ML/AI applications. Given the diversity and speed of evolution in technologies, use cases, market practices, and policy debates, this requires particularly careful balancing between the advantages of consistent baseline standards and the need for tailored approaches. Diversity of backgrounds and disciplines is also critical, given that the issues posed by ML/AI are not simply technical in nature but rather implicate a range of broader economic, policy, and dignitary considerations.

---

[50] See, e.g., NIST AI Risk Management Framework Playbook (2023); U.S. Chamber of Commerce Technology Engagement Center, Comment on Artificial Intelligence Risk Management Framework Request for Information to NIST (Sept. 15, 2021).

While some technologies and use cases are still in very early days, several actions could help the financial services ecosystem move toward more rapid identification and implementation of best practices and regulatory safeguards:

- o **Increasing resources to support the production of public research, engagement by historically underrepresented and under-resourced actors, and broad intra-and inter-sector dialogue:** Technology resources and expertise have a significant impact on the extent which stakeholders can adopt ML/AI applications and engage in the process of refining related market practices and policy frameworks. Considering ways to facilitate greater engagement by smaller financial services providers, historically underserved communities and their advocates, civil society and academic organizations, and government agencies (both regulators and law enforcement) is critical to ensuring that ML/AI adoption operates to the benefit of broader populations and the general economy rather than narrower groups of providers or customers. For instance, increasing resources for public research and facilitating the creation of high quality datasets could help to increase the knowledge base of all stakeholders about evolving technologies and market practices. Additional guidance and supervisory activity by regulators could help to address the particular challenges that smaller providers face in both adopting new technologies and performing due diligence.[51] Creating opportunities for a broad range of financial services stakeholders to dialogue about emerging issues and to draw upon broader debates about data science and standards in other sectors is also critically important at a time of rapid change.

- o **Careful consideration of data governance practices and standards:** Questions regarding data quality, accessibility, and protection are fundamental to both the potential benefits and risks of ML/AI applications. Many of the worst headlines concerning ML/AI applications gone wrong relate to flaws in the nature or treatment of underlying data,[52] and many of the most promising use cases of ML/AI for financial inclusion also hinge in significant part on the ability to access new data sources.[53] Yet data bias and governance issues also arise in contexts that do not involve machine learning in the first instance, and therefore require continuing direct attention in their own right to further refine best practices and regulatory expectations. While federal law provides more detailed and robust protections for consumer financial data than for many other categories of

---

[51] FinRegLab, Machine Learning Policy & Empirical Findings Overview § 5.3.

[52] See, e.g., Gianluca Mauro & Hilke Schellmann, 'There Is No Standard': Investigation Finds AI Algorithms Objectify Women's Bodies, The Guardian (Feb. 8, 2023); Janus Rose, Facebook's New AI System Has a 'High Propensity' for Racism and Bias, Vice (May 9, 2022); Leonardo Nicoletti & Dina Bass, Humans Are Biased and Generative AI is Even Worse, Bloomberg (2023). Steve Lohr, Facial Recognition Is Accurate, If You're a White Guy, N.Y. Times (Feb. 9, 2018); Ed Yong, A Popular Algorithm Is No Better at Predicting Crimes than Random People, The Atlantic (Jan. 17, 2018); Starre Vartan, Racial Bias Found in a Major Health Care Risk Algorithm, Scientific American (Oct. 24, 2019).

[53] For discussion of the potential combination of new data sources and ML/AI applications in credit underwriting, see Sections 2.a and 3.a and FinRegLab, Machine Learning Market & Data Science Context Report § 2.1.2. The combination of more inclusive data sources and ML/AI applications also holds promise in identity verification for purposes of other financial products and services. Kathleen Yaworksy et al., Unlocking the Promise of (Big) Data to Promote Financial Inclusion, Accion (2017).

consumer information, many of the consumer financial laws have not been substantially updated in several decades. They also vary substantially as to what products and services are covered, how they extend to small business owners (if at all), and the level of regulatory monitoring and enforcement.[54]

o **Review of other risk management and customer protection frameworks that apply to automated decisionmaking:** As described above, model risk management expectations and fair lending requirements have pushed financial services providers to manage affirmatively for performance and fairness risks when adopting various ML/AI applications. However, those regulatory regimes do not apply equally to the entire financial services sector, prompting concerns both about concentration of risk and about level playing fields between competitors. Accordingly, some stakeholders have suggested that imposing basic governance expectations on nonbank financial services providers could be beneficial to the broader ecosystem. Stakeholders are also pointing to ways in which the existing regulatory guidance could be updated and expanded to address topics that are becoming more urgent in the ML/AI era, such as standards for evaluating *post hoc* explainability tools and for evaluating multiple underwriting models to determine whether they constitute a "less discriminatory alternative" for purposes of disparate impact compliance.[55]

In considering such frameworks, one important consideration is whether to focus on machine learning and artificial intelligence specifically, on a broader range of statistical models and automated systems, or on the performance of the general function regardless of the extent it is executed by humans, computers, or some combination of the two.[56] Consistent high-level principles or standards may be particularly useful at this stage of evolution, given the definitional challenges discussed above, the fact that technologies are evolving rapidly, and the universal importance of qualities such as accuracy and fairness.

o **Broader efforts to increase opportunity, fairness, and economic participation:** It is also critical to note that while filling information gaps and adopting more predictive models could help substantial numbers of consumers and small business owners access more affordable credit, such actions will not by themselves erase longstanding disparities in income and assets or recent hardships imposed by the pandemic. These factors will continue to shape whether and how customers access financial services, for instance by affecting the number of loan applicants who are assessed as presenting significant risk of default, which will in turn continue to affect whether they are granted credit and at what price. This underscores the importance of using many initiatives and policy levers to address the deep racial disparities in income and assets at the same time that stakeholders in the financial system continue to explore and implement promising data

---

[54] See generally Financial Health Network, Flourish, FinRegLab & Mitchell Sandler, Consumer Financial Data: Legal and Regulatory Landscape.

[55] See generally FinRegLab, Machine Learning Policy & Empirical Findings Overview.

[56] See generally Jonas Schuett, Defining the Scope of AI Regulations (Nov. 20, 2022), https://arxiv.org/abs/1909.01095.

and modeling technique innovations. While there is reason to believe that the financial system can enhance its ability to provide fair and inclusive products and services, relying solely on it to address these cumulative, structural issues would produce too little change too slowly.

Thank you again for the opportunity to speak with you today.